Transactions of the ASME®

Technical Editor, **LEWIS T. WHEELER**

APPLIED MECHANICS DIVISION

Executive Committee
(Chair) **T. J. R. HUGHES**
D. KRAJCINOVIC
S. KYRIAKIDES
P. D. SPANOS
M. C. BOYCE
Associate Technical Editors
J. R. BARBER (2003)
R. C. BENSON (2003)
A. A. FERRI (2003)
H. GAO (2003)
J. W. JU (2001)
V. K. KINRA (2002)
D. A. KOURIS (2002)
A. K. MAL (2001)
B. M. MORAN (2002)
A. NEEDLEMAN (2001)
M. ORTIZ (2001)
N. C. PERKINS (2002)
M.-J. PINDERA (2003)
K. R. RAJAGOPAL (2003)
K. T. RAMESH (2003)
K. RAVI-CHANDAR (2003)
W. S. SARIC (2003)
D. A. SIGINER (2003)
T. E. TEZDUYAR (2003)
N. TRIANTAFYLLIDIS (2003)

BOARD ON COMMUNICATIONS

Chairman and Vice-President
R. K. SHAH

OFFICERS OF THE ASME

President, **J. R. PARKER**
Executive Director, **D. L. BELDEN**
Treasurer, **J. A. MASON**

PUBLISHING STAFF

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

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published bimonthly (Jan., Mar., May, July, Sept., Nov.) The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing office. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300. **CHANGES OF ADDRESS** must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address. **STATEMENT from By-Laws.** The Society shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B7.1, Para. 3).

Journal of Applied Mechanics

Published Bimonthly by The American Society of Mechanical Engineers

VOLUME 68 • NUMBER 2 • MARCH 2001

(**Contents continued on inside back cover**)

This journal is printed on acid-free paper, which exceeds the ANSI Z39.48-1992 specification for permanence of paper and library materials. ♾™
♻ 85% recycled content, including 10% post-consumer fibers.

Journal of Applied Mechanics Volume 68, Number 2 MARCH 2001

BRIEF NOTES

ANNOUNCEMENTS AND SPECIAL NOTES

Finite Amplitude Azimuthal Shear Waves in a Compressible Hyperelastic Solid

J. B. Haddow
Department of Mechanical Engineering,
University of Victoria,
Victora, BC V8W 3P6, Canada
Mem ASME

L. Jiang
Martec Ltd.,
1888 Brunswick Street
Suite 400,
Halifax Nova Scotia B36 3J8
Canada

Lagrangian equations of motion for finite amplitude azimuthal shear wave propagation in a compressible isotropic hyperelastic solid are obtained in conservation form with a source term. A Godunov-type finite difference procedure is used along with these equations to obtain numerical solutions for wave propagation emanating from a cylindrical cavity, of fixed radius, whose surface is subjected to the sudden application of a spatially uniform azimuthal shearing stress. Results are presented for waves propagating radially outwards; however, the numerical procedure can also be used to obtain solutions if waves are reflected radially inwards from a cylindrical outer surface of the medium. A class of strain energy functions is considered, which is a compressible generalization of the Mooney-Rivlin strain energy function, and it is shown that, for this class, an azimuthal shear wave can not propagate without a coupled longitudinal wave. This is in contrast to the problem of finite amplitude plane shear wave propagation with the neo-Hookean generalization, for which a shear wave can propagate without a coupled longitudinal wave. The plane problem is discussed briefly for comparison with the azimuthal problem.
[DOI: 10.1115/1.1334862]

1 Introduction

Static azimuthal shear of a compressible hyperelastic solid has been considered by various authors and recently Polignone and Horgan [1], Beatty and Jiang [2] and Jiang and Ogden [3] have obtained conditions that compressible strain energy functions must satisfy so that pure azimuthal strain is possible. A related dynamic problem of small azimuthal oscillations superimposed on a finite static pure azimuthal shear of a particular compressible hyperelastic solid has been considered by Vandyke and Wineman [4], who found that the small oscillations induce radial motion. The purpose of this paper is to consider propagation of finite amplitude azimuthal shear waves in a compressible isotropic elastic solid. This is a plane-strain problem in which a spatially uniform azimuthal shearing stress is suddenly applied to the surface of a circular cylindrical cavity in an unbounded medium, with the radius of the cavity held constant. A purely mechanical theory is considered and a study by Haddow and Jiang [5] provides some justification for this. The numerical procedure used can also be applied when there is a fixed outer radius which results in reflected waves. Numerical results are obtained for a class of strain energy functions which is a compressible generalization of the Mooney-Rivlin form and is given by

$$W = W_D + H(I_3), \tag{1.1}$$

where

$$W_D = \frac{\mu}{2}\{f(I_1 - 3I_3^{1/3}) + (1-f)(I_2/I_3 - 3I_3^{-1/3})\}, \tag{1.2}$$

and $H(I_3)$ satifies the conditions $H(1) = H'(I) = 0$. In (1.2), μ is the shear modulus for infinitesimal deformation from the natural reference state, $1 \geqslant f \geqslant 0$, I_1, I_2, and I_3 are the principal invariants of $\mathbf{F}^T\mathbf{F}$ or $\mathbf{F}\mathbf{F}^T$, $\mathbf{F}$ is the deformation gradient tensor and the superposed T denotes the transpose. For pure dilatation $W_D = 0$, and for isochoric deformation $H = 0$. The generalized Blatz and Ko strain energy function ([6]) and the polynomial strain energy function proposed by Levinson and Burgess [7] are examples of functions which can be expressed in the form (1.1).

It may be shown that

$$H(I_3) = \frac{1}{2}K(I_3^{1/2} - 1)^2 + O(I_3^{1/2} - 1)^3, \tag{1.3}$$

where K is the bulk modulus for infinitesimal deformation from the natural reference state, and the approximation

$$H(I_3) = \frac{1}{2}K(I_3^{1/2} - 1)^2, \tag{1.4}$$

is applicable for sufficiently small values of $|I_3^{1/2} - 1|$, which decrease as K/μ increases.

Numerical results are presented for $K/\mu = 100$, which gives Poisson's ratio $\nu = 0.495$ for infinitesimal deformation from the natural reference state. This is close to the value $\nu = 0.493$, obtained experimentally by Beatty and Stalnaker [8] for urethane. For the problems considered, the volume strain, $I_3^{1/2} - 1$, is a second-order effect with its maximum absolute value less than 0.0035. Results were obtained using approximation (1.4), and

$$H = \frac{\mu}{2}\left[f\left\{ 3I_3^{1/3} - \frac{(1+\nu)}{\nu} + \frac{(1-2\nu)}{\nu} I_3^{-\nu/(1-2\nu)} \right\} + (1-f)\left\{ 3I_3^{-1/3} - \frac{(1+\nu)}{\nu} + \frac{(1-2\nu)}{\nu} I_3^{\nu/(1-2\nu)} \right\} \right], \tag{1.5}$$

for the generalized Blatz and Ko model, and

$$H = \frac{\mu}{2}\left(-3 + 3fI_3^{1/3} + 3(1-f)I_3^{-1/3} + 2(1-2f)(I_3^{1/2} - 1) + \left(2f + \frac{4\nu - 1}{1-2\nu} \right)(I_3^{1/2} - 1)^2 \right), \tag{1.6}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 22, 1999; final revision, June 1, 2000. Associate Technical Editor: A. K. Mal. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

for the Levinson and Burgess model. The results obtained using (1.4), (1.5), and (1.6) were in very close agreement. Consequently it is reasonable to assume the results are valid for any admissible form of H if approximation (1.4) is applicable.

Strain energy functions of the class (1.1) do not satisfy the necessary conditions for the existence of static pure azimuthal shear, and the propagation of finite amplitude azimuthal waves involves coupled radial longitudinal waves.

A plane shear wave can propagate without a coupled plane longitudinal wave if a generalization of the neo-Hookean strain energy function, that is (1.1) with $f=1$, is assumed. For (1.1), with $1>f\geqslant 0$, a plane shear wave cannot exist without a coupled longitudinal wave. A brief discussion of the propagation of plane shear and plane longitudinal waves is now given, as a preliminary to the consideration of the azimuthal problem.

2 Plane Wave Problem

In this section we consider plane longitudinal and transverse waves propagating in the X_1 direction of a half-space which is defined in the natural reference state by $X_1\geqslant 0$, where OX_α, $\alpha\in\{1,2,3\}$, is a rectangular Cartesian coordinate system. The dynamic deformation is plane strain and is given by

$$x_1=x_1(X_1,t),\quad x_2=\lambda X_2+q(X_1,t),\quad x_3=X_3, \tag{2.1}$$

where λ is a constant, t is time, X_α, $\alpha\in\{1,2,3\}$ and x_i, $i\in\{1,2,3\}$ are the coordinates of a material particle in the natural reference and current states, respectively. The special case with $\lambda=1$, is of most interest; however, the more general case of constant λ is considered for comparison with the azimuthal shear problem.

Components of the deformation gradient tensor are given by

$$[\mathbf{F}]=\begin{bmatrix}\delta & 0 & 0\\ \gamma & \lambda & 0\\ 0 & 0 & 1\end{bmatrix}, \tag{2.2}$$

where $\delta=\partial x_1/\partial X_1$, $\gamma=\partial q/\partial X_1$. The principal invariants of $\mathbf{F}^T\mathbf{F}$ or $\mathbf{F}\mathbf{F}^T$ are then given by

$$\begin{aligned} I_1&=\delta^2+\gamma^2+\lambda^2+1,\\ I_2&=\lambda+\gamma^2+\delta^2\delta^2\lambda^2,\\ I_3&=\lambda^2\delta^2. \end{aligned} \tag{2.3}$$

A Lagrangian approach is used and the nominal stress tensor, $\mathbf{S}$, is a function of $\mathbf{F}(X_1,t)$,

$$\mathbf{S}=\hat{\mathbf{S}}(\mathbf{F})\quad\text{and}\quad \mathbf{S}(X_1,t)=\hat{\mathbf{S}}(\mathbf{F}(X_1,t)). \tag{2.4}$$

The boundary conditions are

$$S_{11}(0,t)=\phi(t),\quad S_{12}(0,t)=\psi(t), \tag{2.5}$$

where S_{11} and S_{12} are the normal and tangential components, respectively, of the stress vector acting on a material plane surface normal to the X_1-axis. The initial conditions are

$$S_{11}(X_1,0)=S_{12}(X_1,0)=0\quad\text{and}\quad V_1(X_1,0)=V_2(X_1,0)=0, \tag{2.6}$$

where V_1 and V_2 are the nonzero velocity components and are in the X_1 and X_2 directions, respectively. The nominal stresses, S_{11} and S_{12}, are identical to the corresponding Cauchy stresses, σ_{11} and σ_{12} when $\lambda=1$.

It is necessary to express (1.1), as a function of δ, γ, and λ, and it follows from (1.1), (1.2), and (2.3) that

$$\begin{aligned} &\hat{W}(\delta,\gamma,\lambda)\\ &=\frac{\mu}{2}\left\{\begin{aligned} &f(\delta^2+\gamma^2+\lambda^2+1-3\lambda^{2/3}\delta^{2/3})\\ &+(1-f)(\lambda^{-2}+\gamma^2\lambda^{-2}\delta^{-2}+\delta^{-2}+1-3\lambda^{-2/3}\delta^{-2/3})\end{aligned}\right\}\\ &\quad+H(\lambda^2\delta^2). \end{aligned} \tag{2.7}$$

The nominal stresses of interest are then given by

$$\begin{aligned} \hat{S}_{11}&=\frac{\partial\hat{W}}{\partial\delta}=\mu\{f(\delta-\delta^{-1/3}\lambda^{2/3})\\ &\quad-(1-f)(\lambda^2\delta^{-3}\gamma^2+\delta^{-3}-\lambda^{-2/3}\delta^{-5/3})\}\\ &\quad+\frac{\partial H}{\partial\delta}, \end{aligned} \tag{2.8}$$

$$\hat{S}_{12}=\frac{\partial\hat{W}}{\partial\gamma}=\mu\{f\gamma+(1-f)\gamma\lambda^2\delta^{-2}\}. \tag{2.9}$$

If $f=1$, $\hat{S}_{11}$ depends only on δ and the constant value of λ, and $\hat{S}_{12}$ depends only on γ, so that there is a limited form of superposition.

Governing equations for the wave propagation problem are the compatibility equations,

$$\frac{\partial\delta}{\partial t}-\frac{\partial V_1}{\partial X_1}=0, \tag{2.10a}$$

$$\frac{\partial\gamma}{\partial t}-\frac{\partial V_2}{\partial X_1}=0, \tag{2.10b}$$

and the equations of motion,

$$\frac{\partial V_1}{\partial t}-\frac{1}{\rho_0}\frac{\partial S_{11}}{\partial X_1}=0, \tag{2.11a}$$

$$\frac{\partial V_2}{\partial t}-\frac{1}{\rho_0}\frac{\partial S_{12}}{\partial X_1}=0, \tag{2.11b}$$

where ρ_0 is the density in the natural reference configuration. Equations (2.10) and (2.11) are a totally hyperbolic system of quasi-linear partial differential equations. The wave velocities $\pm c_L$ and $\pm c_T$, where the $+(-)$ signs denote waves propagating in the $+(-)$ X_1 directions, are obtained as indicated in the text by Whitham [9], and are given by

$$\begin{aligned} c_L&=\sqrt{\frac{(c_{11}+c_{22})+\sqrt{(c_{11}-c_{22})^2+4c_{12}^2}}{2}},\\ c_T&=\sqrt{\frac{(c_{11}+c_{22})-\sqrt{(c_{11}-c_{22})^2+4c_{12}^2}}{2}}, \end{aligned} \tag{2.12}$$

where

$$\begin{aligned} c_{11}&=\frac{1}{\rho_0}\frac{\partial S_{11}}{\partial\delta}=\frac{1}{\rho_0}\frac{\partial^2\hat{W}}{\partial\delta^2},\quad c_{22}=\frac{1}{\rho_0}\frac{\partial S_{12}}{\partial\gamma}=\frac{1}{\rho_0}\frac{\partial^2\hat{W}}{\partial\gamma^2},\\ c_{12}&=\frac{1}{\rho_0}\frac{\partial^2\hat{W}}{\partial\gamma\partial\delta}. \end{aligned} \tag{2.13}$$

It is evident from (2.12) that $c_L>c_T$, consequently c_L and c_T are the propagation speeds of what are essentially longitudinal and shear waves, respectively. If $f=1$ it follows from (2.7) and (2.13) that $c_{12}=0$, c_{11} is a function of δ and the constant value of λ only, and c_{22} is a constant, so that the wave speeds $c_L=\sqrt{c_{11}}$ and $c_T=\sqrt{c_{22}}$ are uncoupled. It also follows that, if $f=1$, the system of equations (2.10*a*), (2.11*a*), governing longitudinal wave propagation, and the system (2.10*b*), (2.11*b*), governing shear wave propagation, are uncoupled, consequently a plane shear wave can propagate without a coupled longitudinal wave.

3 Formulation of Azimuthal Shear Problem

The surface of a circular cylindrical cavity of fixed radius, A, in an unbounded medium, is subjected to a spatially uniform sudden application of azimuthal shearing stress. The numerical scheme used to obtain solutions is applicable to consider reflected waves

propagating radially inwards from an outer cylindrical surface; however, in this paper we consider only waves traveling radially outwards.

Cylindrical polar coordinates of a material point are denoted by R,Θ,Z in the natural reference configuration and by r,θ,z in the deformed configuration. The deformation is plane strain and is given by

$$r=r(R,t), \quad \theta=\Theta+g(R,t), \quad z=Z, \tag{3.1}$$

so that the components of the deformation gradient tensor are

$$[\mathbf{F}]=\begin{bmatrix} r' & 0 & 0 \\ rg' & r/R & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.2}$$

where a prime denotes the partial derivative with respect to R. The following notation is introduced:

$$\dot{g}=\omega, \quad \dot{r}=V_r, \quad \omega r=V_\theta, \tag{3.3}$$

and

$$r'=\delta, \quad r/R=\lambda, \quad g'=\alpha \quad \text{and} \quad rg'=\gamma, \tag{3.4}$$

where a superposed dot denotes partial differentiation with respect to t. If $r/R=1$ in (3.2) the resulting isochoric deformation is pure azimuthal shear, and (3.2) is of the same form as (2.2) with $\lambda=1$. Certain compressible strain energy functions admit static pure azimuthal shear deformation; however, it is probable that there is no compressible strain energy function which will admit dynamic pure azimuthal shear.

Again a Lagrangian approach is adopted and, similar to (2.4),

$$\mathbf{S}=\hat{\mathbf{S}}(\mathbf{F}) \quad \text{and} \quad \mathbf{S}(R,t)=\hat{\mathbf{S}}(\mathbf{F}(R,t)).$$

The components of $\mathbf{S}$ are

$$[\mathbf{S}]=\begin{bmatrix} S_{Rr} & S_{R\theta} & 0 \\ S_{\Theta r} & S_{\Theta\theta} & 0 \\ 0 & 0 & S_{Zz} \end{bmatrix}.$$

However, only S_{Rr}, $S_{R\theta}$, and $S_{\Theta\theta}$ are of interest for the wave propagation problem and these are given by

$$\hat{S}_{Rr}=\frac{\partial\hat{W}}{\partial\delta}, \quad S_{\Theta\theta}=\frac{\partial\hat{W}}{\partial\lambda}, \quad S_{R\theta}=\frac{\partial\hat{W}}{\partial\gamma}, \tag{3.5}$$

where the strain energy function is a function $W=\hat{W}(\delta,\lambda,\gamma)$ of δ, λ, and γ.

The boundary conditions are

$$S_{R\theta}(A,t)=S_0U(t), \quad r(A,t)=A\Rightarrow V_r(A,t)=0, \quad \lambda(A,t)=1, \tag{3.6}$$

where $U(t)$ is the unit step function and S_0 is a constant, and the initial conditions are

$$\mathbf{S}(R,0)=0, \quad V_r(R,0)=V_\theta(R,0)=0. \tag{3.7}$$

4 Governing Equations for Azimuthal Shear Problem

The following compatibility equations can be deduced from (3.3) and (3.4)

$$\frac{\partial\lambda}{\partial t}-\frac{V_r}{R}=0, \tag{4.1a}$$

$$\frac{\partial\delta}{\partial t}-\frac{\partial V_r}{\partial R}=0, \tag{4.1b}$$

$$\frac{\partial\alpha}{\partial t}-\frac{\partial\omega}{\partial R}=0, \tag{4.1c}$$

and the equation of motion in vector form is

$$\text{Div}\mathbf{S}=\rho_0\mathbf{a}, \tag{4.2}$$

where $\mathbf{a}$ is the acceleration with nonzero components,

$$a_r=\dot{V}_r-\lambda R\omega^2, \tag{4.3a}$$

$$a_\theta=\lambda R\dot{\omega}+2V_r\omega, \tag{4.3b}$$

and ρ_0 is the density in the natural reference configuration. The nontrivial components of (4.2) are

$$\frac{\partial S_{Rr}}{\partial R}+\frac{S_{Rr}-S_{\Theta\theta}}{R}-\alpha S_{R\theta}=\rho_0(\dot{V}_r-\lambda R\omega^2), \tag{4.4}$$

and

$$\frac{\partial S_{R\theta}}{\partial R}+\frac{S_{R\theta}}{R}\left(1+\frac{\delta}{\lambda}\right)=\rho_0(\lambda R\dot{\omega}+2V_r\omega). \tag{4.5}$$

In order to apply the numerical scheme it is necessary to put the system of equations given by (4.1), (4.4), and (4.5) in the conservation form

$$\frac{\partial\mathbf{Q}}{\partial t}+\frac{\partial\mathbf{H}(\mathbf{Q})}{\partial R}+\mathbf{b}(\mathbf{Q})=\mathbf{0}, \tag{4.6}$$

where $\mathbf{Q}=(\lambda,\delta,\gamma,V_r,V_\theta)^T$, the superposed T denotes the transpose and $\mathbf{b}$ is a source term.

Equations (4.1a) and (4.1b) are in a form suitable for (4.6); however, (4.1c) must be put in the form

$$\frac{\partial\gamma}{\partial t}-\frac{\partial V_\theta}{\partial R}+\left(\frac{V_\theta\delta-V_r\gamma}{R\lambda}\right)=0, \tag{4.7}$$

which is obtained by using the relations $\gamma=\lambda R\alpha$ and $V_\theta=\lambda R\omega$. Also the acceleration components (4.3) must be put in the form

$$a_r=\frac{\partial V_r}{\partial t}-\frac{V_\theta^2}{\lambda R}, \quad a_\theta=\frac{\partial V_\theta}{\partial t}+\frac{V_rV_\theta}{\lambda R}. \tag{4.8}$$

At this stage it is convenient to adopt the following nondimensionalization scheme:

$$(\bar{r},\bar{R})=(r,R)/A, \quad (\bar{W},\bar{\mathbf{S}})=(W,\mathbf{S})/\mu, \quad \bar{t}=t(\mu/\rho_0)^{1/2}/A,$$
$$(\bar{V}_r,\bar{V}_\theta)=(V_r,V)_\theta/(\mu/\rho_0)^{1/2}, \quad \bar{\omega}=\omega A/(\mu/\rho_0)^{1/2}, \tag{4.9}$$
$$\bar{\mathbf{a}}=\mathbf{a}\rho_0A/\mu.$$

Henceforth nondimensional quantities are used, with the overbars omitted, and primes and superposed dots now denote partial differentiation with respect to nondimensional R and t, respectively. Since the overbars are omitted (4.1a,b) and (4.7) are unchanged by the nondimensionalization.

The nondimensional forms of the matrices in (4.6),

$$\mathbf{Q}=\begin{bmatrix} \lambda \\ \delta \\ \gamma \\ V_r \\ V_\theta \end{bmatrix}, \quad \mathbf{H}=\begin{bmatrix} 0 \\ -V_r \\ -V_\theta \\ -S_{Rr} \\ -S_{R\theta} \end{bmatrix}$$

and

$$\mathbf{b}=\begin{bmatrix} -V_r/R \\ 0 \\ (V_\theta\delta-V_r\gamma)/(\lambda R) \\ -(S_{Rr}-S_{\Theta\theta})/R+\gamma S_{R\theta}/(\lambda R)-V_\theta^2/(\lambda R) \\ -S_{R\theta}(1+\delta/\lambda)/R+V_rV_\theta/(\lambda R) \end{bmatrix}, \tag{4.10}$$

are formed from (4.1a), (4.1b), (4.7) and the nondimensional forms of (4.4) and (4.5) with the acceleration components given by (4.8). In order to determine the wave velocities, (4.6) must be put in the form

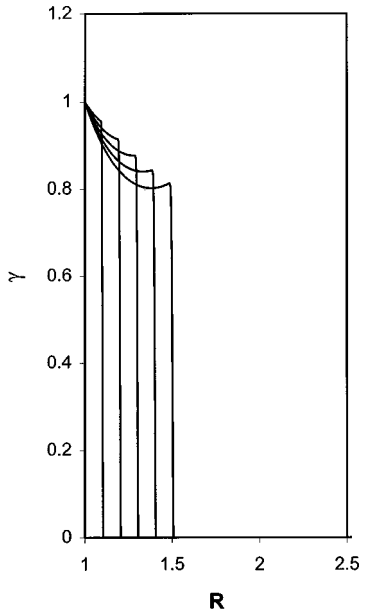


Fig. 1 Relationships between γ and nondimensional radius R for nondimensional times 0.1, 0.2, 0.3, 0.4, 0.5 and $f=1$ and $f=0.6$

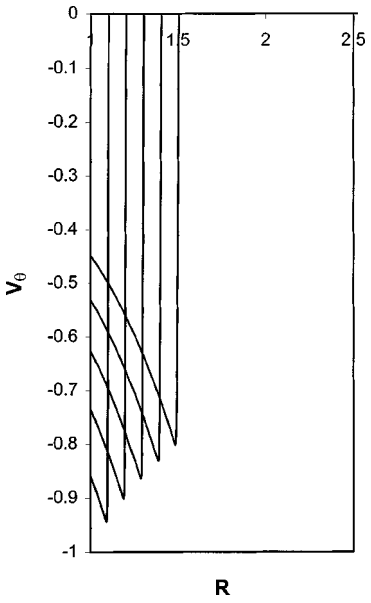


Fig. 2 Relationships between nondimensional V_θ and nondimensional radius R for nondimensional times 0.1, 0.2, 0.3, 0.4, 0.5 and $f=1$ and $f=0.6$

$$\frac{\partial \mathbf{Q}}{\partial t}+\mathbf{A}(\mathbf{Q})\frac{\partial \mathbf{Q}}{\partial t}+\mathbf{b}'=\mathbf{0}, \tag{4.11}$$

where

$$\mathbf{A}=\partial \mathbf{H}(\mathbf{Q})/\partial \mathbf{Q}, \tag{4.12}$$

and $\mathbf{b}'=\mathbf{b}$, since $\mathbf{A}$ is not an explicit function of R and/or t. Equation (4.11) represents a totally hyperbolic system for the present problem and the wave velocities are given by the eigenvalues of $\mathbf{A}$. The numerical scheme used does not involve the relations along the characteristics consequently the eigenvectors of $\mathbf{A}^T$ are not required. It follows from (4.10) and (4.12) that

$$\mathbf{A}=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ A_{41} & A_{42} & A_{43} & 0 & 0 \\ A_{51} & A_{52} & A_{53} & 0 & 0 \end{bmatrix}, \tag{4.13}$$

where

$$\begin{gathered} A_{41}=-\frac{\partial S_{Rr}}{\partial \lambda}=-\frac{\partial^2 \hat{W}}{\partial \delta \partial \lambda}, \quad A_{42}=-\frac{\partial S_{Rr}}{\partial \delta}=-\frac{\partial^2 \hat{W}}{\partial \delta^2}, \\ A_{43}=-\frac{\partial S_{Rr}}{\partial \gamma}=-\frac{\partial^2 \hat{W}}{\partial \delta \partial \gamma}, \\ A_{51}=-\frac{\partial S_{R\theta}}{\partial \lambda}=-\frac{\partial^2 \hat{W}}{\partial \lambda \partial \gamma}, \quad A_{52}=-\frac{\partial S_{R\theta}}{\partial \delta}=A_{43}, \\ A_{53}=-\frac{\partial S_{R\theta}}{\partial \gamma}=-\frac{\partial^2 \hat{W}}{\partial \gamma^2}. \end{gathered} \tag{4.14}$$

The eigenvalues of $\mathbf{A}$ are

$$0, \quad \pm\left\{\frac{1}{2}[-(A_{42}+A_{53})\pm\{(A_{42}-A_{53})^2+4A_{43}^2\}^{1/2}]\right\}^{1/2}, \tag{4.15}$$

and the nonzero eigenvalues are denoted by $\pm c_L$ and $\pm c_T$, where the $+(-)$ signs denote waves propagating in the radially outwards (radially inwards). For comparison with the plane wave problem c_L and c_T are expressed in the form

$$c_L=\sqrt{\frac{(c_{11}+c_{22})+\sqrt{(c_{11}-c_{22})^2+4c_{12}^2}}{2}}, \tag{4.16a}$$

$$c_T=\sqrt{\frac{(c_{11}+c_{22})-\sqrt{(c_{11}-c_{22})^2+4c_{12}^2}}{2}}, \tag{4.16b}$$

where

$$c_{11}=-A_{42}=\frac{\partial^2 \hat{W}}{\partial \delta^2}, \quad c_{22}=-A_{53}=\frac{\partial^2 \hat{W}}{\partial \gamma^2}, \quad c_{12}=-A_{43}=\frac{\partial^2 \hat{W}}{\partial \delta \partial \gamma}. \tag{4.17}$$

In (4.16) $c_L>c_T$ and c_L and c_T represent an essentially longitudinal wave speed and an essentially shear wave speed, respectively, as for the plane case.

The nondimensional form of (2.7) is

$$\begin{aligned} \hat{W}(\delta,\gamma,\lambda) &= \frac{1}{2}\left\{\begin{matrix} f(\delta^2+\gamma^2+\lambda^2+1-3\lambda^{2/3}\delta^{2/3}) \\ +(1-f)(\lambda^{-2}+\gamma^2\lambda^{-2}\delta^{-2}+\delta^{-2}+1-3\lambda^{-2/3}\delta^{-2/3}) \end{matrix}\right\} \\ &\quad +\frac{1}{\mu}H(\lambda^2\delta^2), \end{aligned} \tag{4.18}$$

where δ, λ, and γ are given by (3.4) for the azimuthal shear problem. It follows from (4.17) and (4.18) that

$$\begin{aligned} c_{11}&=f(1+1/3\lambda^{2/3}\delta^{-4/3})+(1-f)(3\delta^{-4}\lambda^{-2}\gamma^2+3\delta^{-4} \\ &\quad -5/3\lambda^{-2/3}\delta^{-8/3})+\frac{1}{\mu}\frac{\partial^2 H}{\partial \delta^2} \end{aligned} \tag{4.19a}$$

$$c_{22}=f+(1-f)\delta^{-2}\lambda^{-2}, \tag{4.19b}$$

$$c_{12}=-2(1-f)\gamma\delta^{-3}\lambda^{-2}. \tag{4.19c}$$

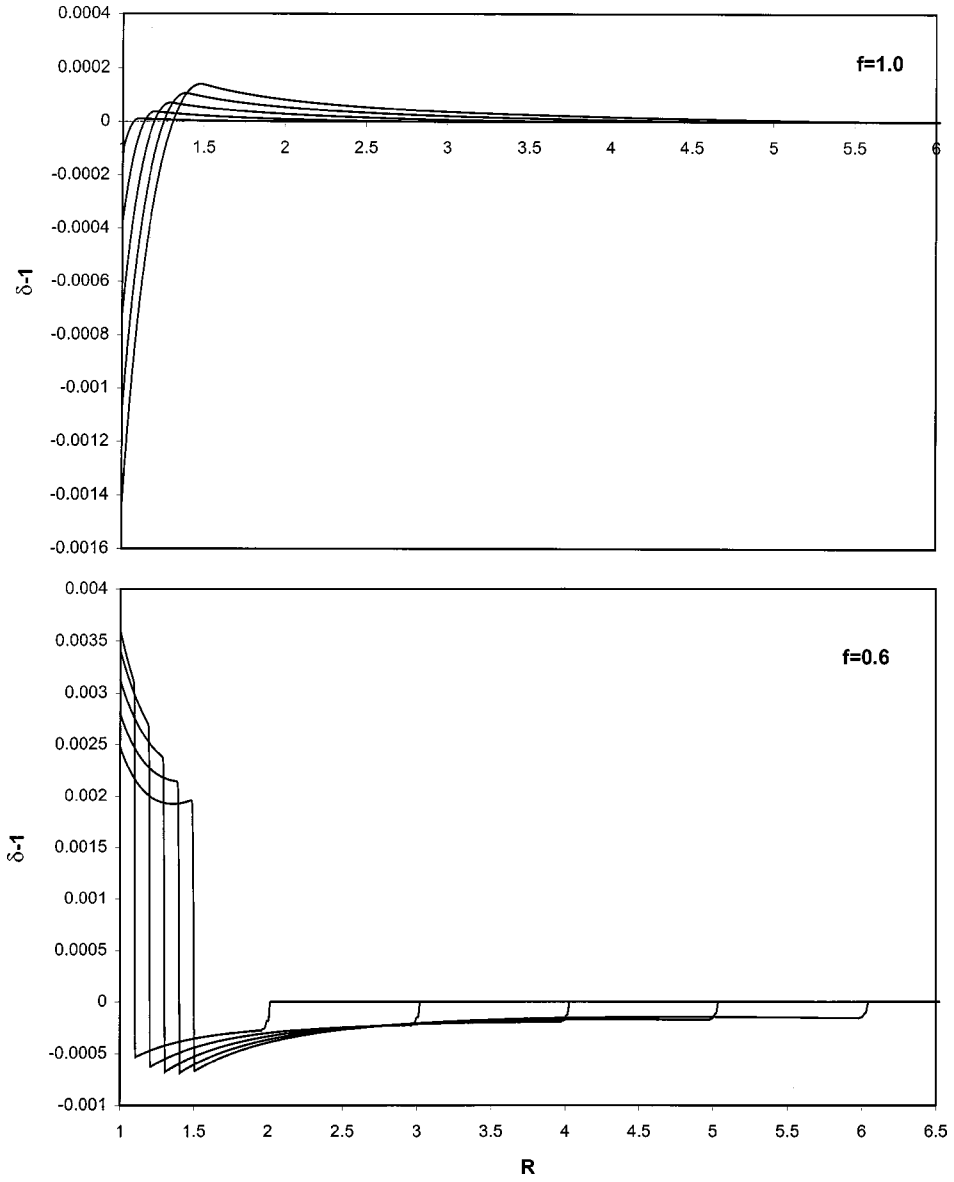


Fig. 3 Relationships between $\delta-1$ and nondimensional radius R for nondimensional times 0.1, 0.2, 0.3, 0.4, 0.5

If $f=1$ in (4.18), it follows from (4.19) that $c_{12}=0$, c_{11} is a function of δ and λ only, and $c_{22}=1$, so that the wave speeds are

$$c_L=\sqrt{c_{11}}=\left(1+1/3\lambda^{2/3}\delta^{-4/3}+\frac{1}{\mu}\frac{\partial^2 H}{\partial\delta^2}\right)^{1/2} \quad \text{and} \quad c_T=\sqrt{c_{22}}=1,$$

and are uncoupled. However, unlike the plane problem, the governing equations for the essentially longitudinal wave propagation and those for essentially shear wave propagation are coupled, consequently an azimuthal shear wave cannot exist without a coupled longitudinal wave for $f=1$.

5 Numerical Method

Numerical results were obtained using a modified second-order Godunov-type finite difference scheme. Application of this scheme requires that the governing system of equations be expressed in the conservation form (4.6). In order to implement the scheme a thick-walled cylinder with nondimensional inner and outer radii 1 and B, respectively, is considered. The interval $[1,B]$ is discretized into m equal cells of length $\Delta R=(B-1)/m$ and $R_i=1+(i-1)\Delta R$, $i\in\{1,m+1\}$. For a typical time step $t\in[t^n,t^{n+1}]$ and jth cell $R\in[R_j,R_{j+1}]$ the weak form of the governing equation can be formulated as

$$\int_{t^n}^{t^{n+1}}\int_{R_j}^{R_{j+1}}\left(\frac{\partial\mathbf{Q}}{\partial t}+\frac{\partial\mathbf{H}}{\partial R}+\mathbf{b}\right)dR\,dt=0, \tag{5.1}$$

using the finite volume method. Integration of (5.1) by parts gives

$$\overline{\mathbf{Q}}_{j+1/2}^{n+1}=\overline{\mathbf{Q}}_{j+1/2}^{n}-\frac{\Delta t}{\Delta R}(\hat{\mathbf{H}}_{j+1}-\hat{\mathbf{H}}_j)-\Delta t\tilde{\mathbf{b}}_{j+1/2}, \tag{5.2}$$

where the superposed bar, caret, and tilde denote spatial, time, and volume-averaged quantities, respectively, whereas the subscripts and superscripts indicate time and space discretizations. Further approximations for the time-averaged values give

$$\overline{\mathbf{Q}}_{j+1/2}^{n+1}=\overline{\mathbf{Q}}_{j+1/2}^{n}-\frac{\Delta t}{\Delta R}[\mathbf{H}(\hat{\mathbf{Q}}_{j+1})-\mathbf{H}(\hat{\mathbf{Q}}_j)]-\Delta t\mathbf{b}\left[\frac{1}{2}(\hat{\mathbf{Q}}_{j+1}+\hat{\mathbf{Q}}_j)\right], \tag{5.3}$$

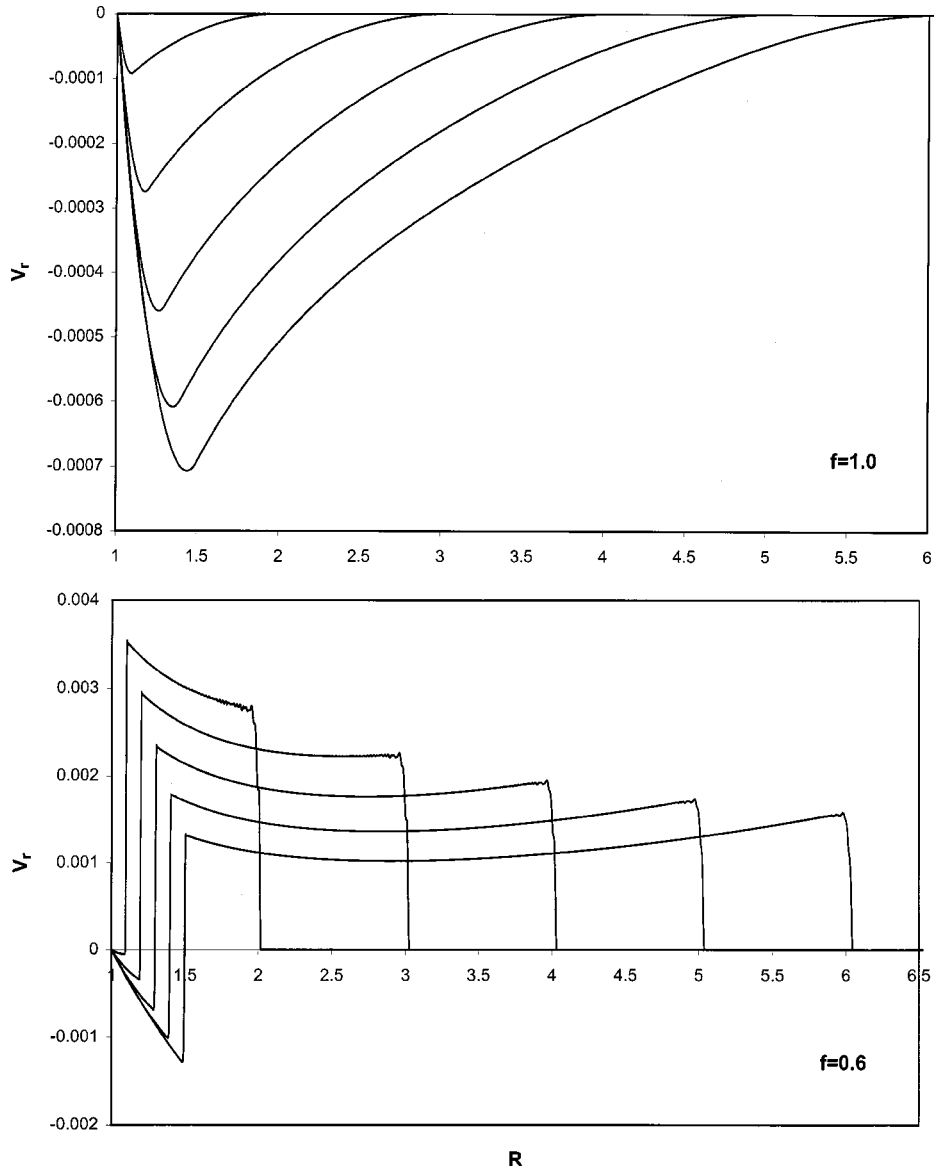


Fig. 4 Relationships between nondimensional V_r and nondimensional radius R for nondimensional times 0.1, 0.2, 0.3, 0.4, 0.5

where the $\hat{\mathbf{Q}}_j$, $j \in \{1,2,\ldots,m+1\}$, are obtained by solving a Riemann problem at each cell interface and by considering the appropriate boundary conditions.

The present implementation of the second-order Godunov-type finite difference method follows the procedure proposed by van Leer [10] and represents the state variables at a given time-step, t^n, by piecewise linear functions of the form

$$\mathbf{Q}(R,t^n)=\overline{\mathbf{Q}}^n_{j+1/2}+\frac{(\Delta\mathbf{Q}^n_{j+1/2})_{av}}{\Delta R}(R-R_{j+1/2}),$$

$$\text{for } R\in[R_j,R_{j+1}], \tag{5.4}$$

where $(\Delta\mathbf{Q}^n_{j+1/2})_{av}$ denotes the spatial-averaged slopes of the state variables which can be evaluated from the spatial-averaged values, $\overline{\mathbf{Q}}^n$, in adjacent cells using a monotonic condition defined as

$$(2\Delta q^n_{j+1/2})_{av}=\begin{cases}\min(|a^n_{j+1/2}|/2,\ 2|b^n_j|,\ 2|c^n_{j+1}|)\times\text{sgn}(a^n_{j+1/2}),\\ \text{if } \text{sgn}(b^n_j)=\text{sgn}(c^n_{j+1}),\\ 0,\ \text{otherwise},\end{cases} \tag{5.5}$$

where

$$a^n_{j+1/2}=\overline{q}^n_{j+3/2}-\overline{q}^n_{j-1/2},\quad b^n_j=\overline{q}^n_{j+1/2}-\overline{q}^n_{j-1/2},$$
$$c^n_{j+1}=\overline{q}^n_{j+3/2}-\overline{q}^n_{j+1/2}$$

and

$$q\in\{\lambda,\delta,\gamma,V_r,V_\theta\}.$$

The state variables at the cell boundaries at a half time-step, $t^{n+1/2}$, can be calculated using Taylor series expansions

$$(\mathbf{Q}^{n+1/2}_j)^+=\overline{\mathbf{Q}}^n_{j+1/2}+\frac{\Delta t}{2}\left(\frac{\partial\mathbf{Q}}{\partial t}\right)^n_{j+1/2}-\frac{\Delta R}{2}\left(\frac{\partial\mathbf{Q}}{\partial R}\right)^n_{j+1/2}, \tag{5.6a}$$

$$(\mathbf{Q}^{n+1/2}_{j+1})^-=\overline{\mathbf{Q}}^n_{j+1/2}+\frac{\Delta t}{2}\left(\frac{\partial\mathbf{Q}}{\partial t}\right)^n_{j+1/2}+\frac{\Delta R}{2}\left(\frac{\partial\mathbf{Q}}{\partial R}\right)^n_{j+1/2}, \tag{5.6b}$$

where the superscript $+(-)$ denotes the state variables at the right (left) side of the cell interface. Approximating the spatial deriva-

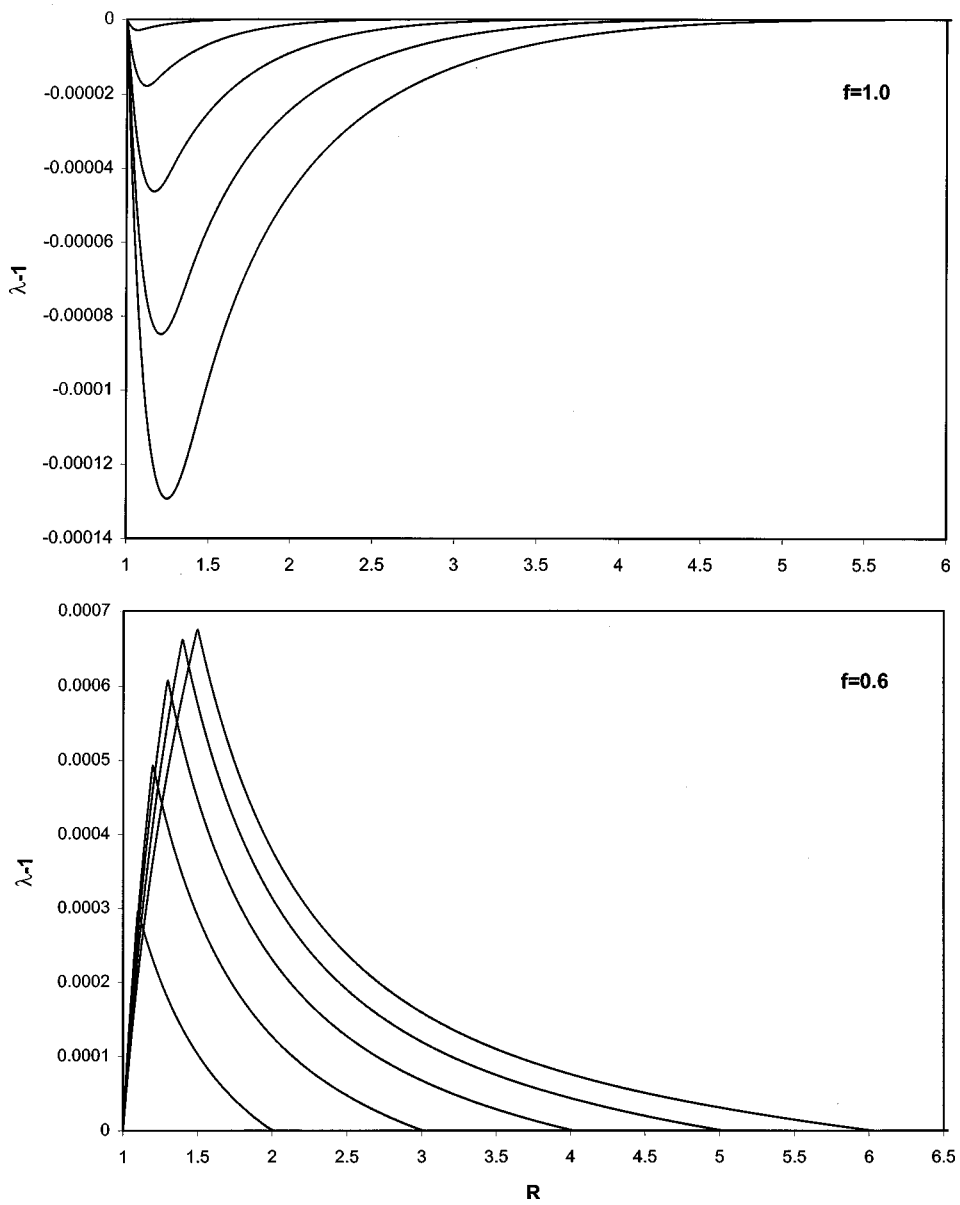


Fig. 5 Relationships between λ−1 and nondimensional radius *R* for nondimensional times 0.1, 0.2, 0.3, 0.4, 0.5

tives in (5.6) by the spatial averaged slopes and eliminating time derivatives by using the linearized form of the system of governing equations (4.11) gives

$$(\mathbf{Q}_j^{n+1/2})^+ = \overline{\mathbf{Q}}_{j+1/2}^n - \frac{1}{2}\left[\mathbf{I} + \frac{\Delta t}{\Delta R}\mathbf{A}(\overline{\mathbf{Q}}_{j+1/2}^n)\right](\Delta \mathbf{Q}_{j+1/2}^n)_{av} - \frac{\Delta t}{2}\mathbf{b}(\overline{\mathbf{Q}}_{j+1/2}^n), \tag{5.7a}$$

$$(\mathbf{Q}_{j+1}^{n+1/2})^- = \overline{\mathbf{Q}}_{j+1/2}^n + \frac{1}{2}\left[\mathbf{I} - \frac{\Delta t}{\Delta R}\mathbf{A}(\overline{\mathbf{Q}}_{j+1/2}^n)\right](\Delta \mathbf{Q}_{j+1/2}^n)_{av} - \frac{\Delta t}{2}\mathbf{b}(\overline{\mathbf{Q}}_{j+1/2}^n), \tag{5.7b}$$

where **A** and **b** are the Jacobian matrix and source term, as defined earlier in this paper, and **I** denotes the unit matrix.

In order to calculate the time-averaged state variables, $\hat{\mathbf{Q}}_j$, a Riemann problem must be solved at each cell interface to resolve discontinuities between the right and left states $(\mathbf{Q}_j^{n+1/2})^+$ and $(\mathbf{Q}_j^{n+1/2})^-$. Since the overall accuracy of the Godunov-type finite difference scheme is usually controlled by the order of the spatial discretizations, an approximate Riemann problem solver has been developed in the present work based on the jump relations

$$\pm U_L[\mathbf{Q}] = [\mathbf{H}], \tag{5.8a}$$

$$\pm U_T[\mathbf{Q}] = [\mathbf{H}], \tag{5.8b}$$

which follow from (4.6). In (5.8*a*,*b*), U_L and U_T are the essentially longitudinal and essentially transverse discontinuity speeds, respectively, and the square brackets denote the discontinuity of the enclosed quantity. Approximating U_L and U_T by the corresponding wave speeds c_L and c_T given in (4.16) leads to

$$(c_{Lj}^{n+1/2})^+(\hat{\delta}_j - (\delta_j^{n+1/2})^+) + (\hat{V}_r - (V_{rj}^{n+1/2})^+) = 0, \tag{5.9a}$$

$$(c_{Lj}^{n+1/2})^-(\hat{\delta}_j - (\delta_j^{n+1/2})^-) - (\hat{V}_r - (V_{rj}^{n+1/2})^-) = 0, \tag{5.9b}$$

and

$$(c_{Tj}^{n+1/2})^+(\hat{\gamma}_j-(\gamma_j^{n+1/2})^+)+(\hat{V}_\theta-(V_{\theta j}^{n+1/2})^+)=0, \tag{5.10a}$$

$$(c_{Tj}^{n+1/2})^-(\hat{\gamma}_j-(\gamma_j^{n+1/2})^-)-(\hat{V}_\theta-(V_{\theta j}^{n+1/2})^-)=0. \tag{5.10b}$$

The time-averaged value of λ is approximated by

$$\hat{\lambda}_j=\frac{1}{2}((\lambda_j^{n+1/2})^+ +(\lambda_j^{n+1/2})^-). \tag{5.11}$$

The boundary conditions must be considered in order to complete the numerical algorithm. For the inner surface the boundary conditions (3.6) give

$$\begin{gathered}\hat{\lambda}_1=1, \quad \hat{V}_{r1}=0,\\ \hat{\delta}_1=(\delta_1^{n+1/2})^+-(\hat{V}_{r1}-(V_{r1}^{n+1/2})^+)/(c_{L1}^{n+1/2})^+,\\ \hat{\gamma}_1=S_0/(f+(1-f)\hat{\lambda}_1^{-2}\hat{\delta}_1^{-2}),\\ \hat{V}_{\theta 1}=(V_{\theta 1}^{n+1/2})^+-(\hat{\gamma}_1-(\gamma_1^{n+1/2})^+)\times(c_{T1}^{n+1/2})^+.\end{gathered} \tag{5.12}$$

The boundary conditions at $R=B$ can also be treated using a similar approach. If it is assumed that both radial and tangential displacement components are zero at $R=B$ then

$$\begin{gathered}\hat{\lambda}_{m+1}=1, \quad \hat{V}_{rm+1}=\hat{V}_{\theta m+1}=0,\\ \hat{\delta}_{m+1}=(\delta_{m+1}^{n+1/2})^-+(\hat{V}_{rm+1}-(V_{rm+1}^{n+1/2})^-)/(c_{Lm+1}^{n+1/2})^-,\\ \hat{\gamma}_{m+1}=(\gamma_{m+1}^{n+1/2})^-+(\hat{V}_{\theta m+1}-(V_{\theta m+1}^{n+1/2})^-)/(c_{Tm+1}^{n+1/2})^-.\end{gathered} \tag{5.13}$$

Although numerical results for times before the longitudinal wave reaches $R=B$ are presented in this paper, it is clear that the finite difference scheme can be used to obtain solutions which involve multiple wave reflection and interactions.

6 Numerical Results

In Figs. 1–5, results for, γ, δ, λ, V_r, and V_θ are shown graphically, as functions of R, for times 0.1 to 0.5, $f=1$ and $f=0.6$, $K/\mu=100$, boundary conditions (3.6) with $S_0=1$, and initial conditions (3.7). These results were obtained using approximation (1.4) or the Levinson and Burgess relation (1.6). Results obtained using (1.4) differ negligibly from those obtained using (1.6) and any differences can not be shown in the figures. Results obtained using the Blatz and Ko relation (1.5) are in very close agreement with those obtained using (1.4) or (1.6).

Results for γ and V_θ exhibit negligible dependence on f and are shown in Figs. 1 and 2, respectively, for both values of f, since any difference between results for $f=1$ and $f=0.6$ can not be shown graphically with scales of Figs. 1 and 2. Qualitative differences between the results for $f=1$ and $f=0.6$ are evident for δ, λ and V_r, which result from the coupled longitudinal wave. However, it is also evident from the figures that the coupled longitudinal wave is a second-order effect.

First consider the numerical results for $f=1$. There is a discontinuity at the transverse wave front across which there is a jump in γ and $S_{R\theta}$ but no jump in δ and S_{Rr}, and $\gamma=0$ ahead of the discontinuity. This discontinuity is not strictly a shock since $c_T=1$, so that the discontinuity speed is equal to the wave speed behind and ahead of the discontinuity. This type of discontinuity is sometimes called a contact discontinuity ([11]). The longitudinal wave is an acceleration wave and exhibits no discontinuities of δ or $\partial\lambda/\partial R$. Since the wave speeds are uncoupled the jump in γ across the transverse discontinuity does not affect the longitudinal wave speed so that it is continuous, along with δ, across the discontinuity.

Now consider the results for $f=0.6$. The results shown in Figs. 3 and 4 indicate that shocks occur. There is an essentially transverse shock with speed U_T. Across this shock there is a jump in γ, and V_θ, which coincides with a jump in δ and V_r so that there is also a jump in S_{Rr} and $S_{R\theta}$. Ahead of this shock $\gamma=0$ and $S_{R\theta}=0$. There is also an essentially longitudinal shock with shock speed $U_L>U_T$. Across this shock there is a jump in δ, V_r and S_{Rr}. The numerical scheme has produced some smearing for this shock. The two shock speeds U_T and U_L, which are taken to be positive since radially outward wave propagation is considered, satisfy the conditions

$$c_T^->U_T>c_T^+, \quad c_L^->U_L>c_L^+,$$

and the jumps satisfy the jump relations

$$U[\mathbf{Q}]=[\mathbf{H}], \tag{6.1}$$

across a shock where $[\phi]=\phi^- -\phi^+$, and ϕ^+ and ϕ^- are the values of the enclosed quantity ahead of and behind the shock, respectively, and U is the shock speed. The results shown in Fig. 5 indicate that there is also a jump in $\partial\lambda/\partial R$ across a jump in δ.

Results were also obtained for values of f, $1>f>0.6$, in order to examine the effect of incremental changes of f on the relations for λ, δ, and V_r. These results indicate that there is a gradual transition from the relations for $f=1$ to the qualitatively different relations for $f=0.6$, which involve shocks. These shocks become stronger as f is decreased and this is related to the coupling of the wave speeds for $f<1$.

7 Concluding Remarks

For the problems considered in this paper, the longitudinal and transverse wave speeds are uncoupled if, and only if, $\partial^2\hat{W}/\partial\gamma\partial\delta=0$. It is easily deduced that, for isotropic strain energy functions, this condition can be satisfied only if $W(I_1,I_2,I_3)$ is of the form

$$W=C_1(I_1-3)+C_2(I_2-3)+H_1(I_3), \tag{7.1}$$

where C_1 and C_2 are constants, $H_1(1)=0$ and $C_1+2C_2+H_1'(1)=0$. Strain energy function (1.1) with $f=1$ is a special case of (7.1). However, for azimuthal shear deformation, even if the wave speeds are uncoupled the governing equations can not be separated into two uncoupled systems which govern the propagation of a transverse wave and a longitudinal wave, respectively. It seems that dynamic pure azimuthal shear deformation is not possible in a compressible hyperelastic solid; however, this still awaits a rigorous proof.

References

[1] Polignone, D. A., and Horgan, C. O., 1994, ''Pure Azimuthal Shear of Compressible Nonlinear Elastic Circular Tubes,'' Q. Appl. Math., **52**, pp. 113–131.

[2] Beatty, M. F., and Jiang, Q., 1997, ''On Compressible Materials Capable of Sustaining Axisymmetric Shear Deformation. Part 2, Rotational Shear of Isotropic Hyperelastic Materials,'' Q. J. Mech. Appl. Math., **50**, pp. 211–237.

[3] Jiang, X., and Ogden, R. W., 1998, ''On Azimuthal Shear of a Circular Cylindrical Tube of Compressible Elastic Material,'' Q. J. Mech. Appl. Math., **51**, pp. 143–158.

[4] Vandyke, T. J., and Wineman, A. S., 1996, ''Small Amplitude Sinusoidal Disturbances Superimposed on Finite Circular Shear of a Compressible, Non-Linearly Elastic Material,'' Int. J. Eng. Sci., **34**, pp. 1197–1210.

[5] Haddow, J. B., and Jiang, L., 1996, ''A Study of Finite Amplitude Plane Wave Propagation in a Rubber-Like Solid,'' Wave Motion, **24**, pp. 211–225.

[6] Blatz, P. J., and Ko, W. L., 1962, ''Application of Finite Elasticity to the Deformation of Rubbery Materials,'' Trans. Soc. Rheol., **6**, pp. 223–251.

[7] Levinson, M., and Burgess, I. W., 1971, ''A Comparison of Some Simple Constitutive Relations for Slightly Compressible Rubberlike Materials,'' Int. J. Mech. Sci., **13**, pp. 563–572.

[8] Beatty, M. F., and Stalnacker, D. O., 1986, ''The Poisson Function of Finite Elasticity,'' ASME J. Appl. Mech., **53**, pp. 807–813.

[9] Whitham, G. B., 1974, *Linear and Nonlinear Waves*, John Wiley and Sons, New York.

[10] van Leer, B., 1979, ''Towards the Ultimate Conservative Difference Scheme, V. A Second-Order Sequel to Godunov's Method,'' J. Comput. Phys., **32**, pp. 101–136.

[11] Sod, G. A., 1985, *Numerical Methods in Fluid Mechanics*, Cambridge University Press, London.

A Dispersive Model for Wave Propagation in Periodic Heterogeneous Media Based on Homogenization With Multiple Spatial and Temporal Scales

W. Chen

J. Fish
Mem. ASME

Department of Civil Engineering and
Scientific Computation Research Center,
Rensselaer Polytechnic Institute,
Troy, NY 12180

A dispersive model is developed for wave propagation in periodic heterogeneous media. The model is based on the higher order mathematical homogenization theory with multiple spatial and temporal scales. A fast spatial scale and a slow temporal scale are introduced to account for the rapid spatial fluctuations as well as to capture the long-term behavior of the homogenized solution. By this approach the problem of secularity, which arises in the conventional multiple-scale higher order homogenization of wave equations with oscillatory coefficients, is successfully resolved. A model initial boundary value problem is analytically solved and the results have been found to be in good agreement with a numerical solution of the source problem in a heterogeneous medium.
[DOI: 10.1115/1.1357165]

1 Introduction

When the wavelength of a traveling signal in a heterogeneous medium is comparable to the characteristic length of the microstructure, successive reflection and refraction of the waves between the interfaces of the material lead to significant dispersion effect (see, for example, [1–3]). This phenomenon cannot be predicted by the classical homogenization theory and thus prompting a significant interest in the scientific community in attempt to develop a dispersive effective medium theory.

The use of multiple-scale expansions as a systematic tool of averaging for problems other than elastodynamics can be traced to Sanchez-Palencia [4], Benssousan, Lions, and Papanicoulau [5], as well as Bakhvalov and Panasenko [6]. The role of higher order terms in the asymptotic expansion has been investigated in statics by Gambin and Kroner [7] and Boutin [8]. In elastodynamics, Boutin and Auriault [9] demonstrated that the terms of a higher order successively introduce effects of polarization, dispersion and attenuation.

There is a substantial number of articles utilizing multiple-scale homogenization techniques for wave propagation problems in periodic media. Most often, a single-frequency time-dependence is assumed prior to the homogenization process [10]. A notable exception is a recent article of Fish and Chen [11], which investigated an initial boundary value problem with rapidly varying coefficients by employing the multiple-scale homogenization technique. It was shown that while higher order terms are capable of capturing dispersion effects, they introduce secular terms which grow unbounded with time. When the observation time is small, higher order terms introduce the necessary correction to the leading order term to resolve the dispersion effect. However, as the time window increases, the higher order terms become close to or larger than the leading order term owing to the existence of secularity. In this case the asymptotic expansion ceases to be valid. In the conventional multiple-scale homogenization technique, these secular terms cannot be cancelled out. To our best knowledge, the present manuscript represents a first attempt to resolve the problem of secularity within the framework of the multiple-scale homogenization for wave propagation in composites.

For dynamic problems, described by hyperbolic differential equations, there are at least four scales involved: (1) the scale of the microstructure, (2) the scale of the macrostructure, (3) the shortest wavelength of the signal traveling in the media, and (4) the time scale of observation. The dispersion phenomena become prominent when the time window is large enough. Therefore, in order to properly model the dispersion effect, it is desirable to construct uniformly valid asymptotic expansions.

The primary objective of the current manuscript is to study the problem of secularity introduced by the higher order multiple-scale approximation of the initial boundary value problem in periodic heterogeneous media. We first consider fast spatial and temporal scales in addition to the usual space-time coordinates. The resulting unit cell problem is shown to be hyperbolic giving rise to fast time-dependence of the solution in the unit cell domain, while the resulting macroscopic equation is the same as in the classical multiple spatial scale analysis and thus failing to resolve dispersion effects. The main contribution of the present paper is given in Sec. 3.2, where we introduce a fast spatial scale aimed to account for rapid spatial fluctuations of material properties and a slow temporal scale designated to capture the long-term behavior of the homogenized solution. The resulting macroscopic equations of motion are solved analytically in Sec. 4 for an illustrative initial boundary value problem.

2 Problem Description

We consider wave propagation normal to the layers of a periodic elastic bilaminate with Ω as the characteristic length (see Fig. 1). The governing equation for this elastodynamics problem is given by

$$\rho(x/\varepsilon)u_{;tt}-[E(x/\varepsilon)u_{;x}]_{;x}=0 \tag{1}$$

with appropriate boundary conditions on the domain boundary and initial conditions

$$u(x,0)=f(x), \quad u_{;t}(x,0)=q(x) \tag{2}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, February 3, 2000; final revision, August 15, 2000. Associate Editor: A. K. Mal. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

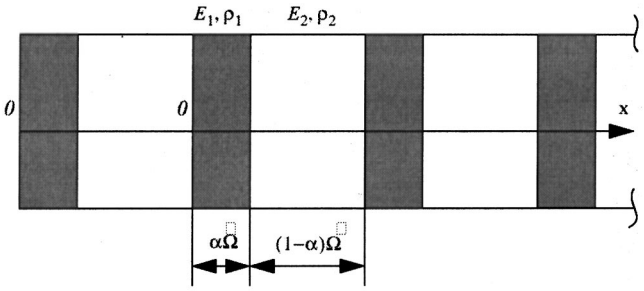


Fig. 1 A bilarninate with periodic microstructure

where $u(x,t)$ represents the displacement field; $\rho(x/\varepsilon)$ and $E(x/\varepsilon)$ are the mass density and elastic modulus, respectively; $(\,)_{;x}$ and $(\,)_{;t}$ denote differentiation with respect to x and time, respectively; and $0<\varepsilon\ll 1$ in Eq. (1) is used to express a rapid spatial variation of material properties.

The goal is to establish an effective homogeneous model in which the local fluctuations due to the heterogeneities do not appear explicitly and the response of the original heterogeneous material can be approximated by the response of the effective homogeneous medium. This is facilitated by the method of multiple-scale asymptotic expansion.

3 Asymptotic Analysis With Multiple Spatial and Temporal Scales

Under the premise that the composite macro-reference length $L=\lambda/(2\pi)$ (λ the macroscopic wavelength) ([9,12]) is much larger than the unit cell dimension Ω, i.e., $\Omega/L=(\omega\Omega)/c=k\Omega \ll 1$, where ω, k, and c are the circular frequency, wave number, and phase velocity of the macroscopic wave, respectively, it is convenient to introduce a microscopic spatial length variable y such that

$$y=x/\varepsilon. \tag{3}$$

In addition to this fast spatial variable, we will experiment with various time scales

$$\xi=\varepsilon^m t \tag{4}$$

where m is an integer. Since the response quantities u and σ depend on x, $y=x/\varepsilon$, t, and $\xi=\varepsilon^m t$, a two-scale asymptotic expansion is employed:

$$u(x,y,t,\xi)=\sum_{i=0}^{n}\varepsilon^i u_i(x,y,t,\xi),$$

$$\sigma(x,y,t,\xi)=\sum_{i=-1}^{n}\varepsilon^i \sigma_i(x,y,t,\xi). \tag{5}$$

The homogenization process consists of inserting the asymptotic expansions (5) into the governing Eq. (1), identifying the terms with equal power of ε, and then solving the resulting problems.

Following the aforementioned procedure and replacing the spatial derivative $(\,)_{;x}$ by $(\,)_{,x}+\varepsilon^{-1}(\,)_{,y}$ and the time derivative $(\,)_{;t}$ by $(\,)_{;t}+\varepsilon^m(\,)_{,\xi}$, we obtain a series of equations in ascending power of ε starting with ε^{-2}. We successively equate the factors of each of these powers to zero.

3.1 Fast Spatial-Temporal Scales. We first experiment with the case of $m=-1$. The resulting two time scales are related by

$$\xi=t/\varepsilon=\eta. \tag{6}$$

At $O(\varepsilon^{-2})$, we get

$$\rho(y)u_{0,\eta\eta}-[E(y)u_{0,y}]_{,y}=0 \tag{7}$$

from where it can be easily shown that u_0 is independent of y and η, thus

$$u_0=U_0(x,t). \tag{8}$$

For $O(\varepsilon^{-1})$ equation we get

$$\rho(y)u_{1,\eta\eta}-[E(y)(u_{0,x}+u_{1,y})]_{,y}=0. \tag{9}$$

Owing to linearity of the above equation, the solution to u_1 can be sought in the form

$$u_1(x,y,t,\eta)=U_1(x,t)+M(y,\eta)u_{0,x}. \tag{10}$$

Substituting Eq. (10) into (9) yields

$$\rho(y)M_{,\eta\eta}-[E(y)(1+M_{,y})]_{,y}=0. \tag{11}$$

Consider the unit cell in Fig. 1. The cell domain consists of subdomains $A^{(1)}$ and $A^{(2)}$, occupied by materials tagged by superscripts 1 and 2, respectively, such that

$$A^{(1)}=[y|0<y<\alpha\hat{\Omega}],\quad A^{(2)}=[y|\alpha\hat{\Omega}<y<\hat{\Omega}] \tag{12}$$

where $0\leqslant\alpha\leqslant 1$ is the volume fraction of the unit cell; Ω is the unit cell domain in the stretched coordinate system y, such that $\Omega/\hat{\Omega}=\varepsilon$. Since material properties are piece-wise constant over the unit cell, Eq. (11) can be written as

$$M_{j,\eta\eta}-c_j^2 M_{j,yy}=0,\quad (j=1,2) \tag{13}$$

where

$$c_1=\sqrt{E_1/\rho_1},\quad c_2=\sqrt{E_2/\rho_2}. \tag{14}$$

The boundary conditions for the unit cell problem described by Eq. (13) are

$$\text{(a) Periodicity:}\quad u_1(y=0)=u_1(y=\hat{\Omega}),$$
$$\sigma_0(y=0)=\sigma_0(y=\hat{\Omega})$$
$$\text{(b) Continuity:}\quad [u_1(y=\alpha\hat{\Omega})]=0,\quad [\sigma_0(y=\alpha\hat{\Omega})]=0 \tag{15}$$

where [ ] is the jump operator and

$$\sigma_i=E(y)(u_{i,x}+u_{i+1,y}),\quad i=0,1,\ldots,n. \tag{16}$$

For simplicity, initial conditions are taken as

$$M_j(y,0)=M_{j,\eta}(y,0)=0,\quad (j=1,2). \tag{17}$$

We solve the unit cell problem defined by Eqs. (13), (15)–(17) using the method of Laplace transform. The detailed solutions are given in the Appendix.

From the solutions for $M(y,\eta)$, it can be observed that $M(y,\eta)$ consists of two parts. The first part is fast time-independent whereas the second part is fast time-dependent.

Finally, for $O(1)$ equation, we get

$$\rho(y)(u_{0,tt}+2u_{1,t\eta}+u_{2,\eta\eta})-[E(y)(u_{0,x}+u_{1,y})]_{,x}$$
$$-[E(y)(u_{1,x}+u_{2,y})]_{,y}=0. \tag{18}$$

For a Ω-periodic function $g=g(x,y,t,\xi)$, we define an averaging operator

$$\langle g\rangle=\frac{1}{|\hat{\Omega}|}\int_{\hat{\Omega}} g(x,y,t,\xi)dy. \tag{19}$$

Applying the above averaging operator to Eq. (18) and making use of the solution for u_1, we arrive at

$$\langle\rho(y)\rangle u_{0,tt}+\langle\rho(y)u_{2,\eta\eta}\rangle-\langle E(y)(1+M_{,y})\rangle u_{0,xx}=0. \tag{20}$$

We assume that fast time-average

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T u_i(x,y,t,\eta)\,d\eta \tag{21}$$

exists and is finite. Following Francfort [13], we suppose that

$$\lim_{s\to 0}\overline{su_i} \tag{22}$$

exists and is finite, where $\overline{u_i}$ is the Laplace transform of u_i with respect to the fast time η. Taking the Laplace transform of Eq. (20) with respect to η and performing the averaging in the fast time, we get the macroscopic equation of motion at $O(1)$:

$$\rho_0 u_{0,tt} - E_0 u_{0,xx} = 0 \tag{23}$$

where

$$\rho_0 = \langle\rho\rangle = \alpha\rho_1 + (1-\alpha)\rho_2, \quad E_0 = \frac{E_1E_2}{(1-\alpha)E_1 + \alpha E_2}. \tag{24}$$

It can be seen that the macroscopic equation of motion at $O(1)$ is nondispersive. It can be also derived without introduction of the fast time-scale (see Sec. 3.2). Proceeding with the derivation of the higher order terms reveals that the fast time-dependence of the displacement field introduces secular terms at $O(\varepsilon^2)$ and higher. Hence the fast time-scaling does not capture the dispersion effect.

3.2 Fast Spatial and Slow Temporal Scales. In this section we experiment with a fast spatial scale aimed at accounting for the rapid spatial fluctuations of material properties and a slow temporal scale intended for the long-term behavior of the homogenized solution. We select $m=2$, i.e.,

$$\xi = \varepsilon^2 t = \tau. \tag{25}$$

At $O(\varepsilon^{-2})$, we have

$$[E(y)u_{0,y}]_{,y} = 0. \tag{26}$$

The general solution to the above equation is

$$u_0 = a_1(x,t,\tau)\int_{y_0}^{y_0+y}\frac{1}{E(z)}\,dz + a_2(x,t,\tau) \tag{27}$$

where $a_1(x,t,\tau)$ and $a_2(x,t,\tau)$ are integration constants. Due to periodicity of u_0, $a_1(x,t,\tau)$ vanishes, implying that the leading-order displacement depends only on the macroscale, i.e.,

$$u_0 = u_0(x,t,\tau). \tag{28}$$

At the next order $O(\varepsilon^{-1})$, the perturbation equation is

$$[E(y)(u_{0,x}+u_{1,y})]_{,y} = 0. \tag{29}$$

Due to linearity, the general solution of u_1 becomes

$$u_1(x,y,t,\tau) = U_1(x,t,\tau) + N(y)u_{0,x}. \tag{30}$$

Substituting Eq. (30) into (29) yields

$$[E(y)(1+N_{,y})]_{,y} = 0. \tag{31}$$

Equation (31) together with the periodicity and continuity conditions of u_1 and σ_0 over the unit cell domain as well as the normalization condition $\langle N(y)\rangle = 0$ define the unit cell boundary value problem, from which $N(y)$ can be uniquely determined as

$$N_1(y) = \frac{(1-\alpha)(E_2-E_1)}{(1-\alpha)E_1+\alpha E_2}\left[y - \frac{\alpha\hat{\Omega}}{2}\right],$$

$$N_2(y) = \frac{\alpha(E_1-E_2)}{(1-\alpha)E_1+\alpha E_2}\left[y - \frac{(1+\alpha)\hat{\Omega}}{2}\right]. \tag{32}$$

It is interesting to note that $N(y)$ is the same as the fast time-independent part of $M(y,\eta)$ in the previous section.

At $O(1)$, the perturbation equation is

$$\rho(y)u_{0,tt} - [E(y)(u_{0,x}+u_{1,y})]_{,x} - [E(y)(u_{1,x}+u_{2,y})]_{,y} = 0. \tag{33}$$

Applying the averaging operator defined in Eq. (19) to the above equation and taking into account the periodicity of σ_1, we get the nondispersive macroscopic equation of motion at $O(1)$, which is identical to Eq. (23). In order to capture the dispersion effect, we proceed to higher order terms.

3.2.1 $O(\varepsilon)$ Homogenization. Higher order correction, u_2, can be determined from $O(1)$ perturbation Eq. (33). Substituting Eqs. (30) and (23) into (33), yields

$$[E(y)(u_{2,y}+U_{1,x}+Nu_{0,xx})]_{,y} = E_0[\rho(y)/\rho_0 - 1]u_{0,xx}. \tag{34}$$

Linearity suggests that u_2 may be sought in the form

$$u_2(x,y,t,\tau) = U_2(x,t,\tau) + N(y)U_{1,x} + P(y)u_{0,xx}. \tag{35}$$

Substituting the above expression into Eq. (34) yields

$$[E(y)(N+P_{,y})]_{,y} = E_0[\rho(y)/\rho_0 - 1]. \tag{36}$$

The boundary conditions for the above equation are: periodicity and continuity of u_2 and σ_1 as well as the normalization condition $\langle P(y)\rangle = 0$. Here we only provide general ideas. For detailed solution of the unit cell boundary value problem we refer to [11]. Once $P(y)$ is found, we can calculate

$$\langle\rho N\rangle = 0, \quad \langle E(N+P_{,y})\rangle = 0, \quad \langle E(u_{1,x}+u_{2,y})\rangle = E_0U_{1,x}. \tag{37}$$

Consider the equilibrium equation of $O(\varepsilon)$:

$$\rho(y)u_{1,tt} - [E(y)(u_{1,x}+u_{2,y})]_{,x} - [E(y)(u_{2,x}+u_{3,y})]_{,y} = 0. \tag{38}$$

Applying the averaging operator to the above equation, exploiting the periodicity of σ_2 and making use of (37), we arrive at

$$\rho_0 U_{1,tt} - E_0 U_{1,xx} = 0. \tag{39}$$

3.2.2 $O(\varepsilon^2)$ Homogenization. Substituting Eqs. (30), (35), and (39) into the $O(\varepsilon)$ equilibrium Eq. (38) yields

$$[E(y)(u_{3,y}+Pu_{0,xxx}+NU_{1,xx}+U_{2,x})]_{,y} = E_0(\rho(y)/\rho_0 - 1)U_{1,xx} + [E_0N\rho(y)/\rho_0 - E(y)(N+P_{,y})]u_{0,xxx}. \tag{40}$$

Due to linearity of the above equation, the general solution to u_3 is as follows:

$$u_3(x,y,t,\tau) = U_3(x,t,\tau) + N(y)U_{2,x} + P(y)U_{1,xx} + Q(y)u_{0,xxx}. \tag{41}$$

Substituting the above expression into Eq. (40) gives

$$[E(y)(P+Q_{,y})]_{,y} = E_0N\rho(y)/\rho_0 - E(y)(N+P_{,y}). \tag{42}$$

The above equation, together with the periodicity and continuity of u_3 and σ_2 over the unit cell domain as well as the normalization condition $\langle Q(y)\rangle = 0$, fully determines $Q(y)$. After $Q(y)$ is solved for, we can calculate

$$\langle\rho P\rangle = \frac{[\alpha(1-\alpha)]^2(\rho_2-\rho_1)(E_1\rho_1 - E_2\rho_2)E_0\hat{\Omega}^2}{12\rho_0E_1E_2}. \tag{43}$$

$$\langle E(P+Q_{,y})\rangle = -\frac{\alpha(1-\alpha)E_0\hat{\Omega}^2}{12\rho_0} \times\left\{\frac{(E_2-E_1)[\alpha^2\rho_1 - (1-\alpha)^2\rho_2] + E_0\rho_0}{(1-\alpha)E_1+\alpha E_2} - \rho_0\right\}. \tag{44}$$

Finally, consider the equilibrium equation of $O(\varepsilon^2)$:

$$\rho(y)[u_{2,tt}+2u_{0,t\tau}] - [E(y)(u_{2,x}+u_{3,y})]_{,x} - [E(y)(u_{3,x}+u_{4,y})]_{,y} = 0. \tag{45}$$

Applying the averaging operator to the above equation, taking into account the periodicity of σ_3 and making use of Eqs. (43) and (44) lead to

$$\rho_0 U_{2,tt} - E_0 U_{2,xx} = \frac{1}{\varepsilon^2} E_d u_{0,xxxx} - 2\rho_0 u_{0,t\tau} \tag{46}$$

where

$$E_d = \frac{[\alpha(1-\alpha)]^2 (E_1\rho_1 - E_2\rho_2)^2 E_0 \Omega^2}{12\rho_0^2 [(1-\alpha)E_1 + \alpha E_2]^2}. \tag{47}$$

E_d characterizes the effect of the microstructure on the macroscopic behavior. It can be seen that it is proportional to the square of the dimension of the unit cell Ω. For a homogeneous material, $\alpha=0$ or $\alpha=1$, and in the case of identical impedance of the two material constituents ($r=z_1/z_2=1$, $z=\sqrt{E\rho}$), E_d vanishes.

Remark 1: In absence of the slow time-scale, the macroscopic equation of motion at $O(\varepsilon^2)$ is

$$\rho_0 U_{2,tt} - E_0 U_{2,xx} = \frac{1}{\varepsilon^2} E_d u_{0,xxxx}. \tag{48}$$

In Sec. 4 we will show that the solution of this equation introduces secular terms.

Remark 2: Alternatively, we could have consider slow time scaling with $m=1$, i.e., $\xi=\varepsilon t=\zeta$. The homogenized equations of motion in this case are

$$\rho_0 U_{1,t\varepsilon} - E_0 U_{1,xx} = -2\rho_0 u_{0,t\zeta} \tag{49}$$

$$\rho_0 U_{2,tt} - E_0 U_{2,xx} = \frac{1}{\varepsilon^2} E_d u_{0,xxx} - 2\rho_0 U_{1,t\zeta} - \rho_0 u_{0,\zeta\zeta}. \tag{50}$$

It can be shown that in this case the forcing term of O(ε) macroscopic equation vanishes, and consequently this scaling does not lead to meaningful results.

3.3 Summary of Macroscopic Equations. In this section we summarize various order macroscopic equations of motion which have been derived in the previous section and prescribe initial and boundary conditions.

The macroscopic equations of motion are

$$O(1): \quad \rho_0 u_{0,tt} - E_0 u_{0,xx} = 0 \tag{23}$$

$$O(\varepsilon): \quad \rho_0 U_{1,tt} - E_0 U_{1,xx} = 0 \tag{39}$$

$$O(\varepsilon^2): \quad \rho_0 U_{2,tt} - E_0 U_{2,xx} = \frac{1}{\varepsilon^2} E_d u_{0,xxx} - 2\rho_0 u_{0,t\tau} \tag{46}$$

We consider the following problem: a domain composed of an array of bilaminates with fixed boundary at $x=0$ and free boundary at $x=l$ subjected to an initial disturbance $f(x)$ in the displacement field. At $O(1)$, the displacement field is determined by equation of motion (23) and the following initial and boundary conditions:

$$\text{ICs:} \quad u_0(x,0,0) = f(x), \quad u_{0,t}(x,0,0) = q(x) = 0 \tag{51}$$

$$\text{BCs:} \quad u_0(0,t,\tau) = 0, \quad u_{0,x}(l,t,\tau) = 0 \tag{52}$$

The calculation of the field $\varepsilon U_1(x,t,\tau)$ is performed by solving equation of motion (39). The initial and boundary conditions applied to $\varepsilon U_1(x,t,\tau)$ must be such that the global field $u_0(x,t,\tau) + \varepsilon U_1(x,t,\tau)$ meets macroscopic initial conditions and conditions imposed on the boundary, i.e.,

$$u_0(x,0,0) + \varepsilon U_1(x,0,0) = f(x),$$

$$u_{0,t}(x,0,0)\varepsilon U_{1,t}(x,0,0) = q(x) = 0$$

$$u_0(0,t,\tau) + \varepsilon U_1(0,t,\tau) = 0, \quad u_{0,x}(l,t,\tau) + \varepsilon U_{1,x}(l,t,\tau) = 0.$$

Taking into account Eqs. (51) and (52), the initial and boundary conditions for $\varepsilon U_1(x,t,\tau)$ are

$$\text{ICs:} \quad \varepsilon U_1(x,0,0) = 0, \quad \varepsilon U_{1,t}(x,0,0) = 0$$

$$\text{BCs:} \quad \varepsilon U_1(0,t,\tau) = 0, \quad \varepsilon U_{1,x}(l,t,\tau) = 0.$$

Similarly, the macroscopic field $\varepsilon^2 U_2(x,t,\tau)$ is determined by equation of motion (46), with initial and boundary conditions for $\varepsilon^2 U_2(x,t,\tau)$ constructed so that the global field, $u_0(x,t,\tau) + \varepsilon U_1(x,t,\tau) + \varepsilon^2 U_2(x,t,\tau)$, should satisfy macroscopic initial and boundary conditions.

With this in mind, we obtain the initial and boundary conditions for different order equations of motion

$$\text{ICs:} \quad u_0(x,0,0) = f(x), \quad u_{0,t}(x,0,0) = q(x) = 0$$

$$U_i(x,0,0) = 0, \quad U_{i,t}(x,0,0) = 0 \quad (i=1,2) \tag{53}$$

$$\text{BCs:} \quad U_i(0,t,\tau) = 0, \quad U_{i,x}(l,t,\tau) = 0 \quad (i=0,1,2). \tag{54}$$

From the above equations of motion and initial boundary conditions, we can observe that

$$U_1(x,t,\tau) \equiv 0. \tag{55}$$

We note that for the multidimensional case U_1 may not vanish. In general U_1 will vanish provided that the material is macroscopically isotropic.

3.4 Nonlocal Macroscopic Equations. We define the mean displacement $U(x,t)$ as

$$U(x,t) = \langle u(x,y,t,\tau) \rangle = u_0(x,t,\tau) + \varepsilon U_1(x,t,\tau) + \varepsilon^2 U_2(x,t,\tau) + \ldots. \tag{56}$$

Combing the macroscopic equations of motion (23), (39), and (46) and neglecting the terms higher than $O(\varepsilon^2)$, we obtain the equation of motion for the mean displacement

$$\rho_0 \ddot{U} - E_0 U_{,xx} - E_d U_{,xxxx} = 0 \tag{57}$$

where $\ddot{U} = U_{,tt}$ is the second-order full-time derivative. The above equation is fourth-order in space. It necessitates four boundary conditions to define a well-posed boundary value problem. However, for the problem under consideration, there are only two physically meaningful boundary conditions for the mean displacement. Equation (57) is often referred to as a ''bad'' Boussinesq equation (cf. [14,15]).

The ''bad'' Boussinesq equation can be reformulated as ''good'' Boussinesq equation ([14,15]) by exploiting the approximation

$$U_{,xx} = \frac{\rho_0}{E_0} \ddot{U} + O(\varepsilon^2) \tag{58}$$

which yields

$$\rho_0 \ddot{U} - E_0 U_{,xx} - \frac{E_d \rho_0}{E_0} \ddot{U}_{,xx} = 0. \tag{59}$$

The above equation is second-order in space, therefore the two physically meaningful boundary conditions are sufficient to define a well-posed initial boundary value problem. Similar equations to (57) and (59) arise in fluid dynamics of shallow water theory and crystal-lattice theory.

4 Solution of Macroscopic Equations

We begin with the zero-order equation of motion (23) and employ separation of variables to solve for this initial boundary value problem. Let

$$u_0(x,t,\tau) = X(x)T(t,\tau). \tag{60}$$

Substituting the above equation into (23) and dividing by the product $X \cdot T$ yield

$$\frac{1}{T}\frac{\partial^2 T}{\partial t^2}=c^2\frac{X''}{X}=-p^2 \tag{61}$$

where p is a separation constant and

$$c=\sqrt{E_0/\rho_0}. \tag{62}$$

The resulting differential equations and corresponding solutions are

$$\frac{\partial^2 T}{\partial t^2}+p^2T=0,\quad T(t,\tau)=S(\tau)\sin(pt)+R(\tau)\cos(pt) \tag{63}$$

$$X''+\frac{p^2}{c^2}X=0,\quad X(x)=h_1\sin\frac{px}{c}+h_2\cos\frac{px}{c} \tag{64}$$

where h_1 and h_2 are integration constants, $S(\tau)$ and $R(\tau)$ are undetermined functions. Substituting the above solutions into the boundary conditions (54) gives

$$h_2=0,\quad h_1\cos\frac{pl}{c}=0 \tag{65}$$

The second condition in the above equation leads to

$$p_n=(2n-1)\frac{\pi c}{2l},\quad (n=1,2,3,\ldots). \tag{66}$$

Due to linearity of the differential equation, the total solution can be written as the sum of individual solutions, i.e.,

$$u_0(x,t,\tau)=\sum_{n=1}^{\infty}\sin\frac{p_nx}{c}[S_n(\tau)\sin(p_nt)+R_n(\tau)\cos(p_nt)]. \tag{67}$$

Inserting the above solution into the second order macroscopic equation of motion (46) gives

$$\begin{aligned}U_{2,tt}-c^2U_{2,xx}=&\sum_{n=1}^{\infty}\frac{p_n}{c}\sin\frac{p_nx}{c}\\&\times\left\{\left[\frac{E_d}{\varepsilon^2\rho_0}\left(\frac{p_n}{c}\right)^3S_n(\tau)+2cR_n'(\tau)\right]\sin(p_nt)\right.\\&\left.+\left[\frac{E_d}{\varepsilon^2\rho_0}\left(\frac{p_n}{c}\right)^3R_n(\tau)-2cS_n'(\tau)\right]\cos(p_nt)\right\}.\end{aligned} \tag{68}$$

The forcing terms in (68) are solutions to the associated homogeneous equation and will generate secular terms. In order to eliminate the secular terms and avoid unbounded resonance of $U_2(x,t,\tau)$, the forcing terms are set to zero, i.e.,

$$\frac{E_d}{\varepsilon^2\rho_0}\left(\frac{p_n}{c}\right)^3S_n(\tau)+2cR_n'(\tau)=0,$$

$$\frac{E_d}{\varepsilon^2\rho_0}\left(\frac{p_n}{c}\right)^3R_n(\tau)-2cS_n'(\tau)=0. \tag{69}$$

Let

$$\omega_n=\frac{E_d}{2c\rho_0}\left(\frac{p_n}{c}\right)^3=\frac{[(2n-1)\pi]^3E_d}{16\rho_0cl^3} \tag{70}$$

then Eq. (69) can be written as

$$\varepsilon^2R_n'(\tau)+\omega_nS_n(\tau)=0,\quad \varepsilon^2S_n'(\tau)-\omega_nR_n(\tau)=0. \tag{71}$$

Differentiating the first equation in (71) and inserting the second equation into the resulting equation lead to

$$\varepsilon^4R_n''(\tau)+\omega_n^2R_n(\tau)=0. \tag{72}$$

Likewise, differentiating the second equation in (71) and inserting the first equation into the resulting equation yields

$$\varepsilon^4S_n''(\tau)+\omega_n^2S_n(\tau)=0. \tag{73}$$

Solutions to (72) and (73) are

$$\begin{aligned}R_n(\tau/\varepsilon^2)&=d_1\sin(\omega_n\tau/\varepsilon^2)+d^2\cos(\omega_n\tau/\varepsilon^2)\\S_n(\tau/\varepsilon^2)&=d_3\sin(\omega_n\tau/\varepsilon^2)+d_4\cos(\omega_n\tau/\varepsilon^2)\end{aligned} \tag{74}$$

where d_1, d_2, d_3, and d_4 are constants of integration. The above solutions must satisfy (71). Inserting (74) into (71) gives

$$d_1=-d_4,\quad d_2=d_3. \tag{75}$$

Substituting Eqs. (74) and (75) into (67) and utilizing initial conditions yield

$$d_1=0,\quad d_2=B_n=\frac{2}{l}\int_0^l f(x)\sin\frac{(2n-1)\pi x}{2l}\,dx \tag{76}$$

and thus the dispersive solution up to the second-order, denoted here as $u_d(x,t,\tau/\varepsilon^2)$, is given as

$$u_d(x,t,\tau/\varepsilon^2)=\sum_{n=1}^{\infty}B_n\sin\frac{p_nx}{c}\cos\left(\omega_n\frac{\tau}{\varepsilon^2}-p_nt\right). \tag{77}$$

For function evaluation we insert $t=\tau/\varepsilon^2$ which yields

$$u_d(x,t)=\sum_{n=1}^{\infty}B_n\sin\frac{p_nx}{c}\cos\left\{\left[\frac{E_d}{2E_0}\left(\frac{p_n}{c}\right)^2-1\right]p_nt\right\}. \tag{78}$$

Remark 3: In absence of the slow time scale, the forcing terms in Eq. (68) cannot be set to zero and hence, the solution to $U_2(x,t)$ will contain secular terms, which grow linearly with time. It has been shown in absence of slow time scaling [11] higher-order solution provides a reasonable approximation provided that the time window is very small.

Remark 4: The solution of the ''good'' Boussinesq Eq. (59) can also be found by separation of variables and is given as

$$U(x,t)=\sum_{n=1}^{\infty}B_n\sin\frac{p_nx}{c}\cos\left[\frac{p_nt}{\sqrt{1+\frac{E_d}{E_0}\left(\frac{p_n}{c}\right)^2}}\right]. \tag{79}$$

Using the binomial expansion

$$\frac{1}{\sqrt{1+\frac{E_d}{E_0}\left(\frac{p_n}{c}\right)^2}}=1-\frac{E_d}{2E_0}\left(\frac{p_n}{c}\right)^2+\frac{3E_d^2}{8E_0^2}\left(\frac{p_n}{c}\right)^4-\cdots. \tag{80}$$

It can be readily observed that solutions (78) and (79) are identical up to the second-order accuracy.

Remark 5: If we supplement the ''bad'' Boussinesq problem (57) with two boundary conditions

$$u_{,xx}(0,t)=0,\quad U_{,xxx}(l,t)=0 \tag{81}$$

satisfying (78) the solution can be found as

$$U(x,t)=\sum_{n=1}^{\infty}B_n\sin\frac{p_nx}{c}\cos\left\{\left[\sqrt{1-\frac{E_d}{E_0}\left(\frac{p_n}{c}\right)^2}\right]p_nt\right\}. \tag{82}$$

Even though the binomial expansion of (82) up to the second-order can be shown to be identical to (78), the above solution is considered to be ''bad'' since the higher order terms in the Fourier series result in a negative term under the square root of Eq. (82) giving rise to physically meaningless solution.

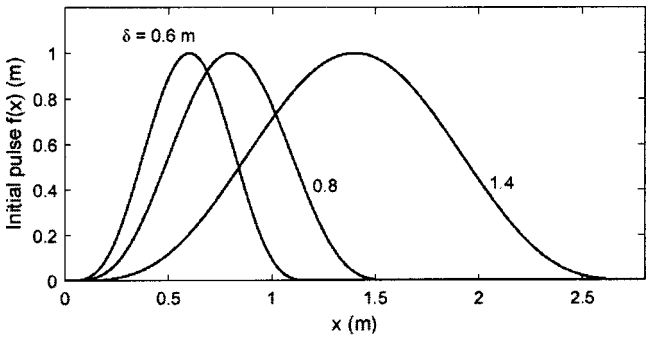


Fig. 2 The initial disturbance in displacement with different half-pulse widths

5 Numerical Results

To assess the accuracy of the proposed model, we construct a reference solution by utilizing a very fine finite element mesh and employ an explicit time integration scheme to solve the source problem in a heterogeneous medium. We consider the following initial disturbance in the displacement field:

$$f(x)=f_0 a_0[x-(x_0-\delta)]^4[x-(x_0+\delta)]^4\{1-H[x-(x_0+\delta)]\}\times[1-H(x_0-\delta-x)]$$

where $a_0=1/\delta^8$ and $H(x)$ is the Heaviside step function; f_0, x_0 and δ are the magnitude, the location of the maximum value and the half width of the initial pulse. Several pulses with $f_0=1m$ and different half-pulse width, δ, are plotted in Fig. 2.

It can be seen that this pulse is similar in shape to the Gaussian distribution function. Substituting the initial disturbance $f(x)$ into Eq. (76) and integrating it analytically, yields

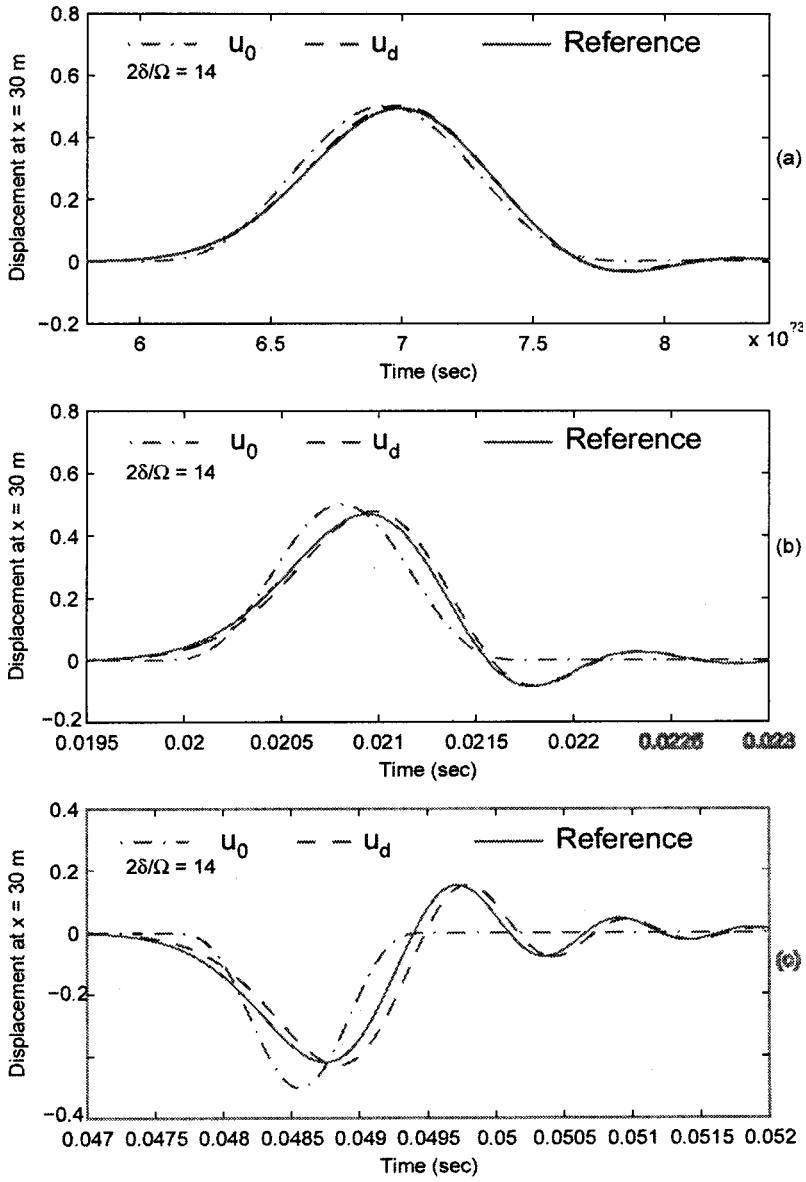


Fig. 3 Displacements at $x=30$ m for the normalized pulse width $2\delta/\Omega=14$

$$B_n=\frac{2}{l}\int_{x_0-\delta}^{x_0+\delta} f_0 a_0[x-(x_0-\delta)]^4 \times[x-(x_0+\delta)]^4 \sin\frac{(2n-1)\pi x}{2l}\,dx$$

$$=\frac{49152l^4 f_0}{\delta^8[(2n-1)\pi]^9}\left\{[1680l^4-180((2n-1)\pi\delta l)^2+((2n-1)\pi\delta)^4]\sin\frac{(2n-1)\pi x_0}{2l}\sin\frac{(2n-1)\pi\delta}{2l}+[20l((2n-1)\pi\delta)^3-840(2n-1)\pi\delta l^3]\times\sin\frac{(2n-1)\pi x_0}{2l}\cos\frac{(2n-1)\pi\delta}{2l}\right\}.$$

We choose material properties as $E_1=120$ GPa, $E_2=6$ GPa, $\rho_1=8000$ Kg/m^3, $\rho_2=3000$ Kg/m^3, and volume fraction $\alpha=0.5$. The dimension of the macro-domain and that of the unit cell are set as $l=40$ m and $\Omega=0.2$ m, respectively. The homogenized material properties are calculated as $E_0=11.43$ GPa, $\rho_0=5500$ Kg/m^3 and $E_d=1.76\times10^7$ N. In this case, $E_1/E_2=20$ and the ratio of the impedances of the two material constituents is $r=7.30$. The initial pulse is centered at the midpoint of the domain, i.e., $x_0=20$ m, with the magnitude $f_0=1.0$ m.

Figures 3–5 show the evolution of displacements at $x=30$ m for different values of pulse width: $\delta=1.4$ m, $\delta=0.8$ m, and $\delta=0.6$ m, respectively. The corresponding ratios between the pulse width and the unit cell dimension, $2\delta/\Omega$, are: 14, 8, and 6, respectively. In each of the Figs. 3–5, there are three responses in each graph (*a*)–(*c*), which correspond to the reference solution of the source heterogeneous problem, the analytical nondispersive solution $u_0(x,t)$ obtained by the classical homogenization and the analytical dispersive solution $u_d(x,t)$.

The dispersion phenomenon can be clearly observed from Figs. 3–5. In the low frequency case, depicted in Fig. 3, the pulse almost maintains its initial shape except for some small wiggles at the wavefront. In this case, the zeroth-order homogenization provides a reasonable approximation to the response of the heterogeneous media. However, when the pulse width of the initial disturbance is comparable to the dimension of the unit cell and the observation time is large, which are the cases shown in Figs. 4 and

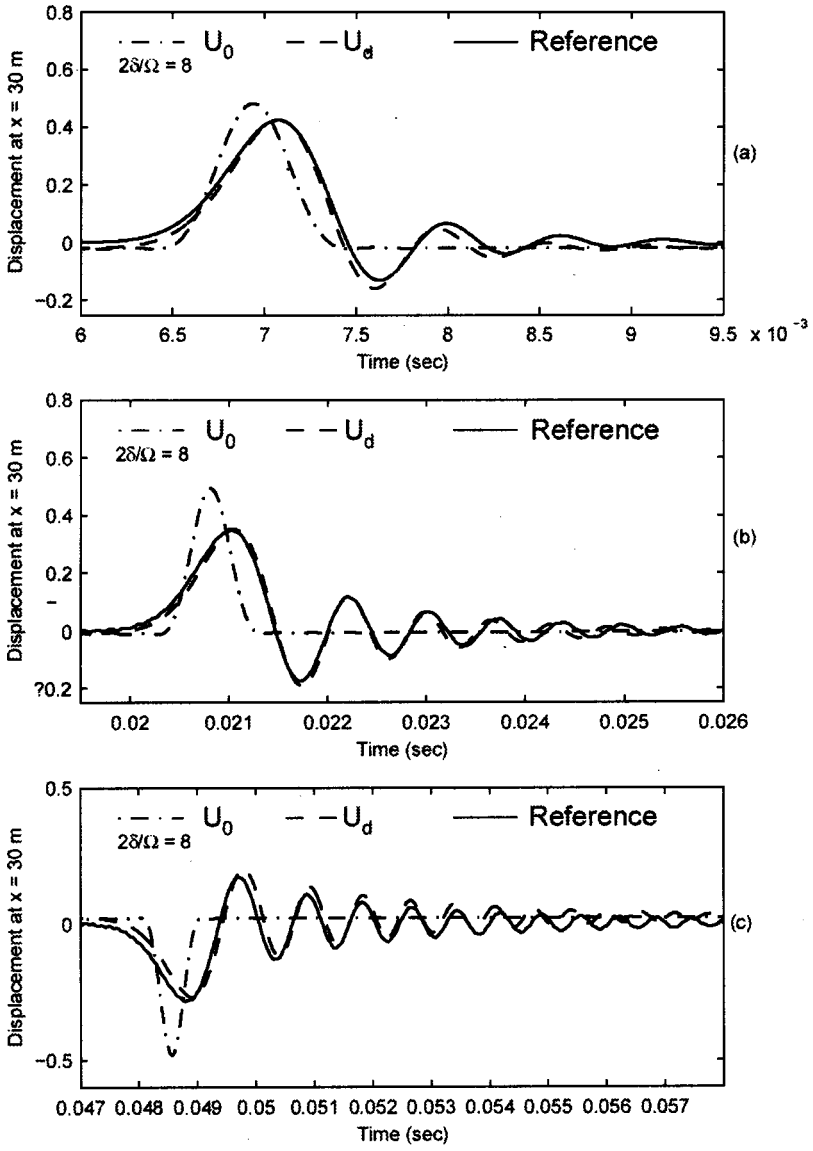


Fig. 4 Displacements at $x=30$ m for the normalized pulse width $2\delta/\Omega=8$

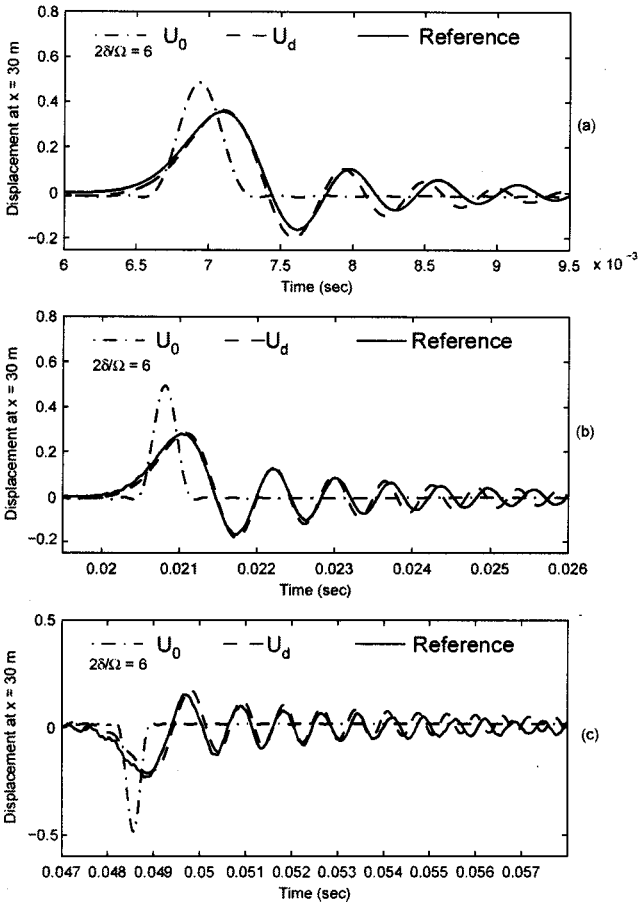


Fig. 5 Displacements at $x=30$ m for the normalized pulse width $2\delta/\Omega=6$

5, the wave becomes strongly dispersive and the zeroth-order homogenization errs badly. It can be also seen that our dispersive model is in good agreement with the reference solution of the heterogeneous medium.

6 Concluding Remarks

Homogenization approach with multiple spatial and temporal scales have been investigated for a model problem. This work is motivated by our recent study ([11]) which suggested that in absence of multiple time scaling, higher-order homogenization method gives rise to secular terms which grow unbounded with time. In the present manuscript we have experimented with various multiple time scalings.

We have found that the combination of fast spatial and slow temporal scales successfully captures dispersion effects. This development serves as a model problem from which we will extend to the general multidimensional cases in our future work. In the general three-dimensional case, the unit cell boundary value problem and macroscopic initial boundary value problem subjected to the secularity constraint will be solved using the finite element method.

Acknowledgment

The support of the National Science Foundation under Agreement number CMS-9713337 and SANDIA National Laboratory under contract number AX-8516 are gratefully acknowledged.

Appendix

Solution to the Fast Time-Dependent Unit Cell Problem Here we give the final results and omit all the details of the solution by the Laplace transform.

$$M_1(y,\eta)=\frac{(1-\alpha)(E_2-E_1)(y-\alpha\hat{\Omega}/2)}{(1-\alpha)E_1+\alpha E_2}+\frac{\alpha\ddot{\Omega}}{4}\left(1-\frac{E_2}{E_1}\right)\sum_{n=1}^{\infty}\left[\frac{W_1(\lambda_{1n})}{G_1(\lambda_{1n})}\cos\frac{2c_1\lambda_{1n}\eta}{\alpha\hat{\Omega}}+\frac{W_1(\lambda_{2n})}{G_2(\lambda_{2n})}\cos\frac{2c_1\lambda_{2n}\eta}{\alpha\hat{\Omega}}\right] \tag{83}$$

$$M_2(y,\eta)=\frac{\alpha(E_1-E_2)[y-(1+\alpha)\hat{\Omega}/2]}{(1-\alpha)E_1+\alpha E_2}+\frac{\alpha\ddot{\Omega}}{4}\left(1-\frac{E_2}{E_1}\right)\sum_{n=1}^{\infty}\left[\frac{W_2(\lambda_{1n})}{G_1(\lambda_{1n})}\cos\frac{2c_1\lambda_{1n}\eta}{\alpha\hat{\Omega}}+\frac{W_2(\lambda_{2n})}{G_2(\lambda_{2n})}\cos\frac{2c_1\lambda_{2n}\eta}{\alpha\hat{\Omega}}\right] \tag{84}$$

where

$$W_1(\lambda)=\sin(2\mu\lambda)\left[\cos\frac{2\lambda y}{\alpha\hat{\Omega}}-\cos\frac{2\lambda(y-\alpha\hat{\Omega})}{\alpha\hat{\Omega}}\right]+k[\cos(2\mu\lambda)-1]\left[\sin\frac{2\lambda y}{\alpha\hat{\Omega}}+\sin\frac{2\lambda(y-\alpha\hat{\Omega})}{\alpha\hat{\Omega}}\right] \tag{85}$$

$$W_2(\lambda)=[\cos(2\lambda)-1]\left[\sin\frac{2\mu\lambda(\hat{\Omega}-y)}{(1-\alpha)\hat{\Omega}}-\sin\frac{2\mu\lambda(y-\alpha\hat{\Omega})}{(1-\alpha)\hat{\Omega}}\right]+k\sin(2\lambda)\left[\cos\frac{2\mu\lambda(\hat{\Omega}-y)}{(1-\alpha)\hat{\Omega}}-\cos\frac{2\mu\lambda(y-\alpha\hat{\Omega})}{(1-\alpha)\hat{\Omega}}\right] \tag{86}$$

$$G_1(\lambda)=\lambda^2[k\sin\lambda\cos(\mu\lambda)+\cos\lambda\sin(\mu\lambda)]\times\{\mu[k\cos\lambda\cos(\mu\lambda)-\sin\lambda\sin(\mu\lambda)]-[k\sin\lambda\sin(\mu\lambda)-\cos\lambda\cos(\mu\lambda)]\} \tag{87}$$

$$G_2(\lambda)=\lambda^2[\sin\lambda\cos(\mu\lambda)+k\cos\lambda\sin(\mu\lambda)]\times\{\mu[-k\sin\lambda\sin(\mu\lambda)+\cos\lambda\cos(\mu\lambda)]+[k\cos\lambda\cos(\mu\lambda)-\sin\lambda\sin(\mu\lambda)]\} \tag{88}$$

$$k=\frac{c_1E_2}{c_2E_1}=\frac{\sqrt{E_2\rho_2}}{\sqrt{E_1\rho_1}},\quad \lambda=\frac{is\alpha\hat{\Omega}}{2c_1},\quad \mu=\frac{(1-\alpha)c_1}{\alpha c_2}. \tag{89}$$

λ_{1n} and λ_{2n} are the roots of

$$F_1(s)=\sin\lambda\cos(\mu\lambda)+k\cos\lambda\sin(\mu\lambda),$$

$$F_2(s)=k\sin\lambda\cos(\mu\lambda)+\cos\lambda\sin(\mu\lambda),$$

respectively.

References

[1] Sun, C. T., Achenbach, J. D., and Herrmann, G., 1968, ''Continuum Theory for a Laminated Medium,'' ASME J. Appl. Mech., **35**, pp. 467–475.

[2] Hegemier, G. A., and Nayfeh, A. H., 1973, ''A Continuum Theory for Wave

Propagation in Laminated Composites. Case I: Propagation Normal to the Laminates,'' ASME J. Appl. Mech., **40**, pp. 503–510.
[3] Bedford, A., Drumheller, D. S., and Sutherland, H. J., 1976, ''On Modeling the Dynamics of Composite Materials,'' *Mechanics Today*, Vol. 3, S. Nemat-Nasser, ed., Pergamon Press, New York, pp. 1–54.
[4] Sanchez-Palencia, E., 1980, *Non-homogeneous Media and Vibration Theory*, Springer, Berlin.
[5] Benssousan, A., Lions, J. L., and Papanicoulau, G., 1978, *Asymptotic Analysis for Periodic Structures*, North-Holland, Amsterdam.
[6] Bakhvalov, N. S., and Panasenko, G. P., 1989, *Homogenization: Averaging Processes in Periodic Media*, Kluwer, Dordrecht.
[7] Gambin, B., and Kroner, E., 1989, ''High Order Terms in the Homogenized Stress-Strain Relation of Periodic Elastic Media,'' Phys. Status Solidi, **51**, pp. 513–519.
[8] Boutin, C., 1996, ''Microstructural Effects in Elastic Composites,'' Int. J. Solids Struct., **33**, No. 7, pp. 1023–1051.
[9] Boutin, C., and Auriault, J. L., 1993, ''Rayleigh Scattering in Elastic Composite Materials,'' Int. J. Eng. Sci., **31**, No. 12, pp. 1669–1689.
[10] Kevorkian, J., and Bosley, D. L., 1998, ''Multiple-Scale Homogenization for Weakly Nonlinear Conservation Laws With Rapid Spatial Fluctuations,'' Stud. Appl. Math., **101**, pp. 127–183.
[11] Fish, J., and Chen, W., 1999, ''High-Order Homogenization of Initial/Boundary-Value Problem With Oscillatory Coefficients. Part I: One-Dimensional Case,'' ASOE J. Eng. Mech., submitted for publication.
[12] Boutin, C., and Auriault, J. L., 1990, ''Dynamic Behavior of Porous Media Saturated by a Viscoelastic Fluid: Application to Bituminous Concretes,'' Int. J. Eng. Sci., **28**, No. 11, pp. 1157–1181.
[13] Francfort, G. A., 1983, ''Homogenization and Linear Thermoelasticity,'' SIAM (Soc. Ind. Appl. Math.) J. Math. Anal., **14**, No. 4, pp. 696–708.
[14] Maugin, G. A., 1994, ''Physical and Mathematical Models of Nonlinear Waves in Solids,'' *Nonlinear Waves in Solids*, A. Jeffrey, and J. Engelbrecht, eds., Springer, Wien.
[15] Whitham, G. B., 1974, *Linear and Nonlinear Waves*, John Wiley and Sons, New York.
[16] Murakami, H., and Hegemier, G. A., 1986, ''A Mixture Model for Unidirectionally Fiber-Reinforced Composites,'' ASME J. Appl. Mech., **53**, pp. 765–773.
[17] Santosa, F., and Symes, W. W., , ''A Dispersive Effective Medium for Wave Propagation in Periodic Composites,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math., **51**, No. 4, pp. 984–1005.
[18] Mei, C. C., Auriault J. L., and Ng, C. O., 1996, ''Some Applications of the Homogenization Theory,'' *Advances in Applied Mechanics*, Vol. 32, J. W. Hutchinson and T. Y. Wu, eds., Academic Press, Boston, pp. 277–348.
[19] Mclachlan, N. W., 1947, *Complex Variable and Operational Calculus With Technical Applications*, Macmillan, New York.
[20] Thomson, W. T., 1950, *Laplace Transformation*, Prentice-Hall, New York.

Transient Green's Function Behavior for a Prestressed Highly Elastic Half-Space

L. M. Brock
Fellow ASME,
Mechanical Engineering,
University of Kentucky,
Lexington, KY 40506

A plane-strain study of a prestressed isotropic compressible neo-Hookean half-space subjected to shear and normal surface loads is performed. The loads are either stationary and applied for an instant, or travel at an arbitrary constant speed. The transient process is viewed as the superposition of infinitesimal deformations upon large, and exact expressions for the displacements, within and upon, the half-space are obtained. These, and the associated wave patterns, demonstrate the anisotropy induced by prestress. The wave speeds themselves are sensitive to prestress; in particular, Rayleigh waves disappear beyond a critical compressive prestress. A critical tensile prestress also exists, beyond which a negative Poisson effect occurs. [DOI: 10.1115/1.1357167]

Introduction

Wave propagation in prestressed solids has applications for seismology, nondestructive evaluation and material characterization ([1]) and, when the solid is highly elastic, finite deformations due to prestress can noticeably affect wave propagation properties ([2]). A convenient approach to the study of this situation is to view the deformation triggered by wave-inducing dynamic loading as infinitesimal, and to superpose it upon the existing (static) deformation due to prestress.

This approach, which allows the use of results by Green and Zerna [3] and Beatty and Usmani [4], is employed in this article to examine plane-strain transient Green's function problems of loads applied to the surface of a prestressed isotropic compressible neo-Hookean material. Shear and normal stationary loads that are applied suddenly, and those that translate at constant speeds over the surface, are both treated. In the latter case, the speed can be any finite constant value, i.e., can be sub, trans or supersonic.

Basic Equations

Consider an elastic body $\mathcal{R}$ that is homogeneous and isotropic relative to an undisturbed reference configuration κ_o. A smooth motion $\mathbf{x}=\mathbf{x}(\mathbf{X})$ then takes $\mathcal{R}$ to a deformed equilibrium configuration κ. The Cauchy stress $\mathbf{T}$ in κ is

$$\mathbf{T}=\alpha_0\mathbf{1}+\alpha_1\mathbf{B}+\alpha_2\mathbf{B}^2,\quad \mathbf{B}=\mathbf{F}\mathbf{F}^T,\quad \mathbf{F}=\frac{\partial\mathbf{x}}{\partial\mathbf{X}} \tag{1}$$

where $(\alpha_0,\alpha_1,\alpha_2)$ are scalar-valued response functions of the principal invariants (I, II, III) of $\mathbf{B}$, and body forces are absent. As noted in [4], experimentally based inequalities ([5]) tend to support the restrictions

$$\alpha_0-II\alpha_2\leqslant 0,\quad \alpha_1+I\alpha_2>0,\quad \alpha_2\leqslant 0. \tag{2}$$

An adjacent nonequilibrium deformed configuration κ^* is obtained by superposing an infinitesimal displacement $\mathbf{u}$, which depends on $\mathbf{x}$ and time. This requires an additional (incremental) Cauchy stress $\mathbf{T}'=\mathbf{T}^*-\mathbf{T}$, where $\mathbf{T}^*$ is the Cauchy stress in κ^*. To the first order in the displacement gradient $\mathbf{H}=\partial\mathbf{u}/\partial\mathbf{x}$, the components of $\mathbf{T}'$ in the principal reference system, i.e., $\mathbf{B}=\mathrm{diag}\{\lambda_1^2,\lambda_2^2,\lambda_3^2\}$ where λ_k are the principal stretches and

$$I=\lambda_1^2+\lambda_2^2+\lambda_3^2,\quad II=\lambda_1^2\lambda_2^2+\lambda_3^2\lambda_1^2+\lambda_2^2\lambda_3^2,\quad III=\lambda_1^2\lambda_2^2\lambda_3^2, \tag{3}$$

can be written as

$$\begin{bmatrix}T'_{11}\\T'_{22}\\T'_{33}\end{bmatrix}=\begin{bmatrix}\lambda'_{11}+2\mu'_{11} & \lambda'_{12} & \lambda'_{13}\\ \lambda'_{21} & \lambda'_{22}+2\mu'_{22} & \lambda'_{23}\\ \lambda'_{31} & \lambda'_{32} & \lambda'_{33}+2\mu'_{33}\end{bmatrix}\begin{bmatrix}H_{11}\\H_{22}\\H_{33}\end{bmatrix} \tag{4a}$$

$$T'_{12}=\mu'_{21}H_{21}+\mu'_{12}H_{12},\quad T'_{23}=\beta'_{32}H_{32}+\mu'_{23}H_{23},$$
$$T'_{31}=\mu'_{13}H_{13}+\mu'_{31}H_{31}. \tag{4b}$$

In (4), $(\lambda'_{ik},\mu'_{ik})$ are the generalized Lame' constants defined by

$$\Gamma'_{i1}=\Gamma_{i1}\lambda_1^2,\quad \Gamma'_{i2}=\Gamma_{i2}\lambda_2^2,\quad \Gamma'_{i3}=\Gamma_{i3}\lambda_3^2 \tag{5}$$

where $i=(1,2,3)$, the symbol Γ represents either λ or μ, and

$$\frac{1}{2}\lambda_{ik}=\frac{\partial\alpha_0}{\partial\lambda_k^2}+\lambda_i^2\frac{\partial\alpha_1}{\partial\lambda_k^2}+\lambda_i^4\frac{\partial\alpha_2}{\partial\lambda_k^2},\quad \mu_{ik}=\mu_{ki}=\alpha_1+\alpha_2(\lambda_i^2+\lambda_k^2). \tag{6}$$

In κ incremental traction boundary conditions on a surface with outwardly directed normal $\mathbf{n}$ can be written in terms of the vector

$$\mathbf{t}^{(n)}=\mathbf{T}'\mathbf{n}+\mathbf{T}\mathbf{n}(\mathbf{n}.\mathbf{H}\mathbf{n})-\mathbf{T}\mathbf{H}'\mathbf{n} \tag{7}$$

Finally, because κ_o is a homogeneous configuration, the incremental balance of linear momentum reduces to ([3])

$$\mathrm{div}\,\mathbf{T}'=\rho\ddot{\mathbf{u}} \tag{8}$$

where ρ is the mass density, $(\cdot)$ denotes (absolute) time differentiation, and a Cartesian basis is understood.

Compressible Neo-Hookean Material

Consider a Hadamard material which is characterized by the response functions

$$\alpha_0=2\sqrt{III}\frac{dG(III)}{dIII},\quad \alpha_1=\frac{1}{\sqrt{III}}(a_o-b_oI),\quad \alpha_2=\frac{b_o}{\sqrt{III}} \tag{9}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, February 23, 2000; final revision, August 28, 2000. Associate Editor: A. K. Mal. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

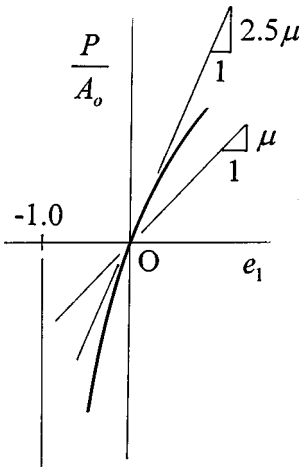


Fig. 1 Isotropic compressible neo-Hookean material in axial loading

where (a_o, b_o) are material constants such that $b_o = a_o - \mu$, μ is the shear modulus, and $G(1) = 0$. Setting $b_o = 0$ produces the subclass of isotropic compressible neo-Hookean materials ([4]), and a simple example of this arises when the form

$$G = \mu\left(\frac{1}{\sqrt{III}} - 1\right) \tag{10}$$

is chosen. This reduces (9) to the one-parameter model

$$\alpha_0 = -\frac{\mu}{III}, \quad \alpha_1 = \frac{\mu}{\sqrt{III}}, \quad \alpha_2 = 0 \tag{11}$$

that satisfies the restrictions (2). For illustration, consider $\mathcal{R}$ to be a cylindrical bar of circular cross-sectional area A_o in κ_o which is placed in a deformed equilibrium state κ under the uniaxial load P. If the bar axis is aligned with the X_1-direction, then the Cauchy stresses in κ are

$$T_{11} = \frac{P}{A}, \quad T_{22} = T_{33} = 0, \quad T_{ik} = 0 \quad (i \neq k) \tag{12}$$

where A is the cross-sectional area in κ, and uniform stress is assumed. Because $\mathbf{X}$ are the principal directions with stretches λ_1 and $\lambda_2 = \lambda_3 = \lambda_T$, Eqs. (1), (3), and (11) combine to give

$$\frac{P}{A} = \mu\left(\lambda_1^{3/2} - \frac{1}{\lambda_1}\right) \tag{13a}$$

$$\lambda_T = \frac{1}{\lambda_1^{1/4}}. \tag{13b}$$

Because $A = \lambda_T^2 A_o$ for homogeneous deformation and $\lambda_1 = 1 + e_1$, where e_1 is the axial unit extension of the bar, (13*a*) can be written as

$$\frac{P}{A_o} = \mu\left[1 + e_1 - \frac{1}{(1+e_1)^{3/2}}\right]. \tag{14}$$

Equation (14) relates a first Piola-Kirchoff stress to unit extension, which is a standard objective of the simple tension test ([6]). A schematic of (14) is given in Fig. 1, and the effective Young's modulus and Poisson's ratio (E^e, ν^e) follow from (13) and the slope of (14) as

$$E^e = \mu\left[1 + \frac{3}{2(1+e_1)^{5/2}}\right], \quad \nu^e = -\frac{\lambda_T - 1}{\lambda_1 - 1} = \frac{(1+e_1)^{1/4} - 1}{e_1(1+e_1)^{1/4}}. \tag{15}$$

Clearly $E^e \to \mu$ for large extensions, but $E^e \to 2.5\mu$ when they are small. This small e_1-behavior corresponds to a Young's modulus in an isotropic linear elastic solid with Poisson's ratio 1/4 ([6]) and, indeed, ν^e in (15) takes this value when $e_1 \to 0$.

Suddenly Applied Stationary Load Problem

Consider $\mathcal{R}$ in κ_o to occupy a half-space defined in terms of fixed Cartesian basis as the region $X_2 > 0$. The smooth motion

$$x_1 = \lambda_1 X_1, \quad x_2 = \lambda_2 X_2, \quad x_3 = X_3 \tag{16}$$

then takes $\mathcal{R}$ to the plane-strain equilibrium state κ defined as

$$T_{11} = \sigma, \quad T_{22} = 0, \quad \lambda_3 = 1 \tag{17}$$

where σ is a known constant stress. The region $\mathcal{R}$ now occupies the half-space $x_2 > 0$ and (x_k, λ_k) are principal directions and stretches. For the compressible neo-Hookean model (11), Eqs. (1), (3), (16), and (17) combine to give

$$\lambda_2 = \frac{1}{\omega^{1/4}}, \quad \lambda_1 = \omega\lambda_2, \quad \omega = \frac{\sigma}{2\mu} + \sqrt{1 + \left(\frac{\sigma}{2\mu}\right)^2} \tag{18a}$$

$$T_{33} = \mu\left(\frac{1}{\sqrt{\omega}} - \frac{1}{\omega}\right), \quad T_{ik} = 0 \quad (i \neq k) \tag{18b}$$

where $0 < \omega \leqslant 1 (\sigma \leqslant 0)$ and $\omega \geqslant 1 (\sigma \geqslant 0)$. Equations (16)–(18) describe $\mathcal{R}$ in κ.

For any superposed infinitesimal deformation $\kappa \to \kappa^*$ the incremental stresses are given by (4) where, in view of (5), (6), (17), and (18),

$$\lambda'_{1k} = \mu\left(\frac{2}{\omega} - \omega\right), \quad \lambda'_{2k} = \mu'_{2k} = \frac{\mu}{\omega}, \quad \lambda'_{3k} = \mu\left(\frac{2}{\omega} - \frac{1}{\sqrt{\omega}}\right) \tag{19a}$$

$$\mu'_{k1} = \mu\omega, \quad \mu'_{k3} = \frac{\mu}{\sqrt{\omega}}. \tag{19b}$$

Here $k = (1,2,3)$ and it is noted that all the constants in (19) are positive so long as

$$0 < \omega < \sqrt{2} \quad \left(\sigma < \frac{\mu}{\sqrt{2}}\right). \tag{20}$$

In this instance the superposed infinitesimal deformation preserves the plane-strain nature of κ by the application of surface shear and normal line loads aligned parallel to the x_3-direction. Application takes place for an instant at $(x_1, x_2) = 0$. This is depicted schematically in Fig. 2(*a*), where (S, N) are the shear and

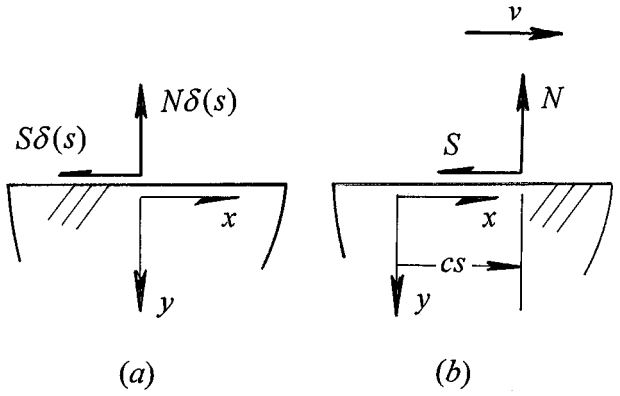


Fig. 2 (*a*) Stationary loads suddenly applied to surface, (*b*) surface loads translating on surface

normal load magnitudes, and $\delta(\,)$ represents the Dirac function. For convenience the variable $s=v_r\times(\text{time})$ has been introduced, where

$$v_r=\sqrt{\frac{\mu}{\rho}} \tag{21}$$

is the classical ([7]) rotational wave speed. Similarly, the variables (x, y) are introduced in place of (x_1,x_2). The boundary conditions for $\mathcal{R}$ in κ^* along $y=0$ are

$$t_3^{(-2)}=0,\quad t_1^{(-2)}=S\,\delta(x)\,\delta(s),\quad t_2^{(-2)}=N\,\delta(x)\,\delta(s). \tag{22}$$

Because the process is plane-strain, the superposed displacements are (u_1,u_2), and depend only on (x, y, s). Thus, in light of (4), (8), and (19), the relevant field equations in $y>0$ for the superposed deformation are

$$\left[\frac{1}{\omega}\frac{\partial^2}{\partial y^2}+\left(\frac{2}{\omega}+\omega\right)\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial s^2}\right]u_1+\frac{2}{\omega}\frac{\partial^2 u_2}{\partial x\,\partial y}=0 \tag{23a}$$

$$\frac{2}{\omega}\frac{\partial^2 u_1}{\partial x\,\partial y}+\left(\frac{3}{\omega}\frac{\partial^2}{\partial y^2}+\omega\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial s^2}\right)u_2=0 \tag{23b}$$

where the nonzero elements of $\mathbf{T}'$ obey the constitutive formulas

$$\frac{1}{\mu}T'_{11}=\left(\frac{2}{\omega}+\omega\right)\frac{\partial u_1}{\partial x}+\left(\frac{2}{\omega}-\omega\right)\frac{\partial u_2}{\partial y},\quad \frac{1}{\mu}T'_{22}=\frac{1}{\omega}\frac{\partial u_1}{\partial x}+\frac{3}{\omega}\frac{\partial u_2}{\partial y} \tag{24a}$$

$$\frac{1}{\mu}T'_{33}=\left(\frac{2}{\omega}-\frac{1}{\sqrt{\omega}}\right)\left(\frac{\partial u_1}{\partial x}+\frac{\partial u_2}{\partial y}\right),\quad \frac{1}{\mu}T'_{12}=\omega\frac{\partial u_2}{\partial x}+\frac{1}{\omega}\frac{\partial u_1}{\partial y}. \tag{24b}$$

Equation (24*a*) indicates that extensional strain associated with the *x*-direction is independent of transverse loading when ω achieves the critical value defined by (20). For ω exceeding this value, (24a) implies a negative Poisson effect. Equation (24) as a set also show that tr$\mathbf{T}'$ and tr$\mathbf{H}$ are not proportional. This indicates the typical ([3,4]) result that the superposed deformations are governed by equations analogous to those for an anisotropic body, even though $\mathcal{R}$ in κ_o is isotropic.

The boundary conditions along $y=0$ for the superposed deformation can in view of (7), (17), and (18*b*) be extracted from (22) as

$$T'_{12}-\sigma\frac{\partial u_2}{\partial x}=S\,\delta(x)\,\delta(s),\quad T'_{22}=N\,\delta(x)\,\delta(s) \tag{25}$$

while the initial $(s<0)$ conditions are

$$\left(u_k,\frac{\partial u_k}{\partial s}\right)\equiv 0 \tag{26}$$

and $k=(1,2)$. In addition u_k should be bounded as $\sqrt{x^2+y^2}\to\infty$, $y>0$ for finite $s>0$, and should be finite and continuous everywhere except perhaps at $(x,y)=0$ and certain wavefronts.

Integral Transform Solution

To obtain the superposed displacements, the unilateral [8] and bilateral [9] Laplace transforms

$$\hat{F}=\int_0^\infty F e^{-ps}ds, \tag{27a}$$

$$\tilde{F}=\int_{-\infty}^{\infty}\hat{F}e^{-pqx}dx \tag{27b}$$

are introduced, along with their corresponding inverse operators

$$F=\frac{1}{2\pi i}\int\hat{F}e^{ps}dp, \tag{28a}$$

$$\hat{F}=\frac{p}{2\pi i}\int\tilde{F}e^{pqx}dq. \tag{28b}$$

In (27) the variable p can be taken as real, positive and large enough to guarantee existence of (27*a*), while q is in general imaginary. In (28*a*,*b*), the integrations are taken along Bromwich contours in, respectively, the complex p and q-planes. Application of (27) to (23) in view of (24)–(26) reduces the superposed deformation problem to two coupled linear ordinary differential equations in $y>0$ that can readily be solved to yield the double transforms

$$\mu p\tilde{u}_k=A_k(q)e^{-\sqrt{\omega}pay}+B_k(q)e^{-\sqrt{\omega}pby} \tag{29}$$

for $y>0$. Here $k=(1,2)$ and the (A_k,B_k) are given by

$$A_1(q)=\frac{q}{R}\left(TN-\frac{2qb}{\sqrt{\omega}}S\right),\quad B_1(q)=-\frac{b}{R}(\sqrt{\omega}TS+2qaN) \tag{30a}$$

$$A_2(q)=\frac{a}{R}(2qbS-\sqrt{\omega}TN),\quad B_2(q)=\frac{q}{R}\left(TS+\frac{2qa}{\sqrt{\omega}}N\right), \tag{30b}$$

where the definitions

$$\sqrt{3}a=\sqrt{1-q^2c_a^2},\quad b=\sqrt{1-q^2c_b^2} \tag{31a}$$

$$T=1-q^2\left(\omega+\frac{1}{\omega}\right),\quad T_o=1+q^2\left(\frac{1}{\omega}-\omega\right) \tag{31b}$$

$$R=\frac{4}{\omega}q^2ab+T^2=\frac{3}{2}(b-a)\left[T_o a+\left(b^2-\frac{5}{3\omega}q^2\right)b\right] \tag{31c}$$

hold, and the dimensionless constants

$$c_b=\sqrt{\omega},\quad c_a=\sqrt{\omega+\frac{2}{\omega}}>c_b \tag{32}$$

define the effective dilatational and rotational wave speeds (v_rc_a,v_rc_b), respectively, in the *x*-direction. To ensure boundedness of (29) in $y>0$ we require that $\mathrm{Re}(a,b)>0$ in the q-plane with, respectively, the branch cuts $\mathrm{Im}(q)=0,|\mathrm{Re}(q)|>1/c_a$ and $\mathrm{Im}(q)=0,|\mathrm{Re}(q)|>1/c_b$. The first expression in (31*c*) shows that R is a form of the classical [7] Rayleigh function. The second expression is less common, but demonstrates in view of (31*a*) and (32) that a Rayleigh function arises in a sense because of the existence of two body waves. In either form, it is analytic in the q-plane with branch cuts $\mathrm{Im}(q)=0,1/c_a<|\mathrm{Re}(q)|<1/c_b$, and exhibits the roots $q=\pm(1/c_o,1/c_R)$, where

$$c_a=\sqrt{\omega-\frac{1}{\omega}}, \tag{33a}$$

$$c_R=\sqrt{\omega+\frac{1}{\omega}\left(1-\frac{2}{\sqrt{3}}\right)}. \tag{33b}$$

It can be shown that it is the last factor in the second form for R in (31*c*) that exhibits the roots associated with (33*b*), while those associated with (33*a*) are roots of both the first factor and the term T_o. For $\sqrt{2/\sqrt{3}-1}<\omega<\sqrt{2}$ it can be shown that $0<c_R<c_b$, i.e., v_rc_R is the Rayleigh speed in κ^*. For $0<\omega<\sqrt{2/\sqrt{3}-1}$, however, c_R is imaginary, and so has no meaning as a wave speed. In view of (18), this implies that Rayleigh waves do not exist for prestresses

$$\sigma<-2\mu(\sqrt{3}-1)\sqrt{2/\sqrt{3}+1}. \tag{34}$$

For $0<\omega<1(\sigma<0)$ c_o is also imaginary, while for $1<\omega<\sqrt{2}(0<\sigma<\mu/\sqrt{2})$, we have $0<c_o<c_R$. In either case, when

$q=\pm 1/c_o$, $b=a$ and (29) exhibits the numerator factor $T_o\to 0$. That is, (29) itself has no pole associated with (33*a*), and that value plays no role in solution behavior.

Rayleigh waves in prestressed media have been considered ([10,11]), and the vanishing of the Rayleigh speed has been interpreted ([11]) in terms of the instability of the homogeneously deformed configuration. In what follows, therefore, prestress is restricted by (20) and (34), i.e., to the range between Rayleigh wave annihilation and the negative Poisson effect.

General Transform Inversions

Operation on (29) with (28*b*) gives the unilateral transform

$$\mu\hat{u}_k=\frac{1}{2\pi i}\int A_k(q)e^{p(qx-\sqrt{\omega}ay)}dq+\frac{1}{2\pi i}\int B_k(q)e^{p(qx-\sqrt{\omega}by)}dq \tag{35}$$

where $k=(1,2)$, and the analyticity of both integrands allows the entire Im(q)-axis serve as the Bromwich contour. For the A_k-term, integrand decay is exponential for $|q|\to\infty$ in the left and right-hand halves if the q-plane for, respectively, $x>0$ and $x<0$. Cauchy theory can then be used to switch integration to the Cagniard [12] contour

$$r_a^2q_\pm=\tau x\pm iyc_a\sqrt{\frac{\omega}{3}}\sqrt{\tau^2-s_a^2},\quad s_a=\frac{r_a}{c_a},$$

$$r_a=\sqrt{x^2+\frac{1}{3}\omega c_a^2y^2} \tag{36}$$

parameterized by the positive real variable $\tau\geqslant s_a$. The schematic in Fig. 3(*a*) for $x>0$ shows that $q_\pm$ define the branches of hyperbolas with, in light of (32), asymptote slopes $\pm yc_ac_b/\sqrt{3}x$ and the intercept $q=-x/c_ar_a$ which always lies between the origin and the branch points $q=-(1/c_a,1/c_b)$ of (a, b). The A_k-integration can now be written as

$$\frac{1}{\pi}\mathrm{Re}\int_0^\infty a(q_+)A_k(q_+)\frac{\sqrt{3}H(\tau-s_a)}{c_a\sqrt{\tau^2-s_a^2}}e^{-p\tau}d\tau \tag{37}$$

where $H(\,)$ is the Heaviside function. In view of (27*a*), (37) is precisely the unilateral transform operation on the integrand, so that its inverse follows by inspection. A similar approach can be used for the B_k-term in (35), with the result

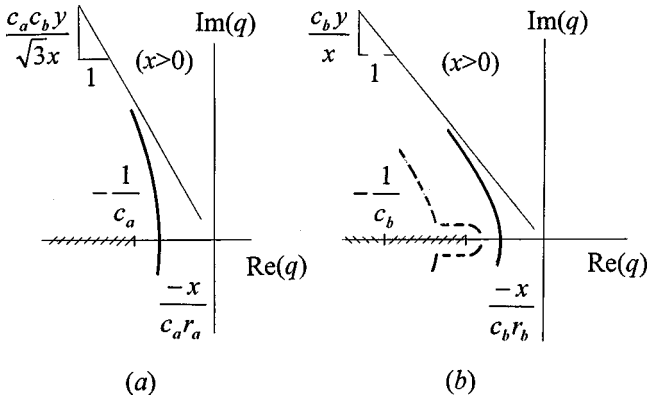


Fig. 3 (*a*) Cagniard contour for A_k-terms in displacement, (*b*) Cagniard contours for B_k-terms in displacement

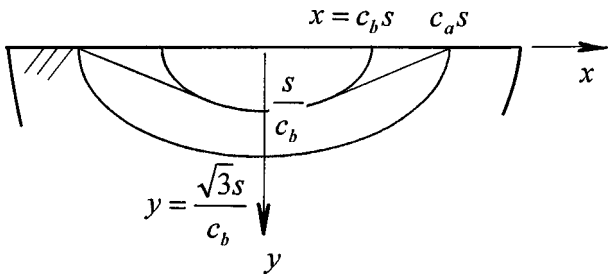


Fig. 4 Wave pattern

$$\pi\mu u_k=\mathrm{Re}[a(q_a)A_k(q_a)]\frac{\sqrt{3}H(s-s_a)}{c_a\sqrt{s^2-s_a^2}}+\mathrm{Re}[b(q_b)B_k(q_b)]\frac{H(s-s_b)}{c_b\sqrt{s^2-s_b^2}}+\mathrm{Im}[b(q_h)B_k(q_h)]\frac{H(s_b-s)}{c_b\sqrt{s_b^2-s^2}}H(s-s_h)\left(|x|>\frac{\omega^2y}{\sqrt{2}}\right) \tag{38}$$

for $y>0$, where $k=(1,2)$. In (38)

$$q_a=q_+(s) \tag{39a}$$

$$r_b^2q_b=-sx+iyc_b\sqrt{\omega}\sqrt{s^2-s_b^2},\quad s_b=\frac{r_b}{c_b},\quad r_b=\sqrt{x^2+\omega c_b^2y^2} \tag{39b}$$

$$r_b^2q_h=-sx+yc_b\sqrt{\omega}\sqrt{s_b^2-s^2},\quad s_h=\frac{1}{c_a}(x+\sqrt{2}y). \tag{39c}$$

As indicated by the curves in Fig. 3(*b*), q_b for $x>0$ defines the upper half of the branch of a hyperbola in the q-plane with asymptote slope c_by/x and intercept $q=-x/c_hr_b$. The intercept always lies between the origin and the branch point $q=-1/c_b$ of b, but for a given location (x, y) may (broken line in Fig. 3(*b*)) or may not (solid line in Fig. 3(*b*)) lie on the branch cut of a. In the former instance, the additional integration path parameterized by q_h is required. This explains the two contributions from B_k to (38).

In light of (36), (38), and (39*a*), the A_k-term is dilatational in nature, and occurs in the outer expanding semi-elliptical region in Fig. 4, while (38) and (39*b*) show that the $B_k(q_b)$-term is a rotational signal, and occurs in the inner expanding semi-elliptical region in Fig. 4. These shapes are manifestations of the anisotropy of the superposed deformation. The disturbances due to the $B_k(q_h)$-term in (38) are the head (bow) wave contributions, and occur in the wedge-shaped regions in Fig. 4.

Surface Behavior

Application of Cauchy theory when $y>0$ does not involve the roots of R because they lie on the branch cuts of (a, b). As $y\to 0$, however, the Cagniard contours in Fig. 3(a, b) collapse onto these cuts, thereby requiring that they be augmented with semicircular paths of vanishingly small radius that are centered at the root locations associated with (33*b*). The results for $y=0$ when $x>0$ then follow as

$$\mu u_1 = -\frac{S}{\pi x}\frac{4\alpha b^2 q^2}{\sqrt{\omega}D}H\left(s-\frac{x}{c_a}\right)+\frac{S}{\pi x}\frac{\sqrt{\omega}\beta T^2}{D}H\left(s-\frac{x}{c_b}\right)$$
$$+\frac{N}{\pi x}\frac{2\alpha bqT}{D}H\left(s-\frac{x}{c_a}\right)H\left(\frac{x}{c_b}-s\right)$$
$$-\frac{N}{4}\frac{3\omega}{2-\sqrt{3}(\omega^2+3)}\delta\left(s-\frac{x}{c_R}\right) \tag{40a}$$

$$\mu u_2 = \frac{N}{\pi x}\frac{\sqrt{\omega}\alpha T^2}{D}H\left(s-\frac{x}{c_a}\right)-\frac{N}{\pi x}\frac{4\alpha^2\beta q^2}{\sqrt{\omega}D}H\left(s-\frac{x}{c_b}\right)$$
$$-\frac{S}{\pi x}\frac{2\alpha bqT}{D}H\left(s-\frac{x}{c_a}\right)H\left(\frac{x}{c_b}-s\right)$$
$$+\frac{S}{4}\frac{3\omega}{2-\sqrt{3}(\omega^2+3)}\delta\left(s-\frac{x}{c_R}\right). \tag{40b}$$

In (40), (b, T) are defined in (31), but now

$$q=-\frac{s}{x},\quad \sqrt{3}\alpha=\sqrt{q^2c_a^2-1},\quad \beta=\sqrt{q^2c_b^2-1} \tag{41a}$$

$$D=\left(T-\frac{2q^2}{\omega}\right)\left(T^2-\frac{4q^2}{3\omega^2}\right). \tag{41b}$$

The third terms in both (40a,b) represent in view of Fig. 4 head wave contributions, while the first and second terms are, respectively, dilatational and rotational in nature. The last terms represent propagating Rayleigh wave spikes. Rayleigh signals are also manifest in the other S (shear force)-terms for u_1 and N (normal force)-terms for u_2 as propagating singularities. Such behavior is analogous to that for a linear elastic solid ([7]).

Translating Surface Load Problem

In this instance, the superposed infinitesimal deformation is triggered by shear and normal surface line loads that are applied at $(x_1,x_2)=0$ but translate over the surface in the positive x_1-direction with a constant speed v. The process is illustrated schematically in Fig. 2(b). Other than replacing the boundary conditions (25) with

$$T'_{12}-\sigma\frac{\partial u_2}{\partial x}=S\delta\left(s-\frac{x}{c}\right),\quad T'_{22}=N\delta\left(s-\frac{x}{c}\right) \tag{42}$$

the analysis of this problem is essentially the same. In (42) the dimensionless sliding speed

$$c=\frac{v}{v_r} \tag{43}$$

is introduced. In view of (27a,b) it can be shown that (29) need be modified only by dividing its right-hand side by $p(q+1/c)$. So long as $c<c_b$, therefore, the inversion process proceeds as before, although the factor $1/p$ implies integration with respect to s. The result is that for $y>0$

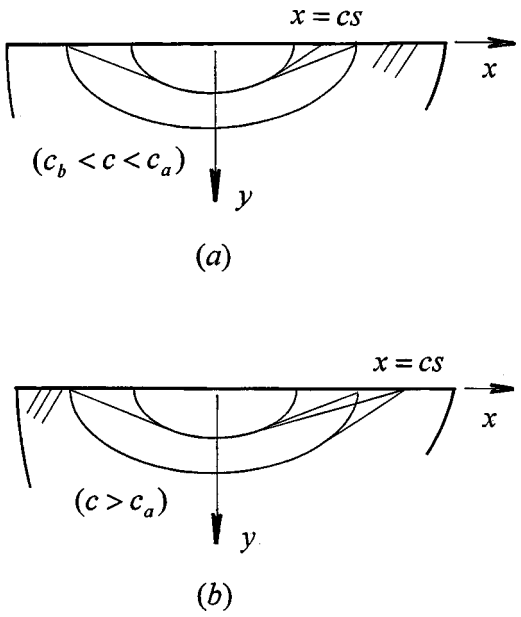


Fig. 5 (*a*) Wave pattern for transonic load speed, (*b*) wave pattern for supersonic load speed

$$\pi\mu u_k=\mathrm{Re}\int_{s_a}^{s}\frac{a(q_a)A_k(q_a)}{q_a+\frac{1}{c}}\frac{\sqrt{3}d\tau}{c_a\sqrt{\tau^2-s_a^2}}$$
$$+\mathrm{Re}\int_{s_b}^{s}\frac{b(q_b)B_k(q_b)}{q_b+\frac{1}{c}}\frac{d\tau}{c_b\sqrt{\tau^2-s_b^2}}$$
$$+\mathrm{Im}\int_{s_h}^{s_b}\frac{b(q_h)B_k(q_h)}{q_h+\frac{1}{c}}\frac{H(s-\tau)d\tau}{c_b\sqrt{s_b^2-\tau^2}}\left(|x|>\frac{\omega^2 y}{\sqrt{2}}\right) \tag{44}$$

where it is understood that integrals vanish unless the upper limit exceeds the lower. Here $k=(1,2)$, (A_k,B_k) are defined by (30) and (31) and the integration variable τ replaces s in (39).

The disturbances associated with (44) also occur in the expanding regions depicted in Fig. 4. When $c>c_b$, however, an integrand pole at $q=-1/c$ lies between the branch point $q=-1/c_b$ and the origin. Depending upon where the location (x, y) places the intercept of the Cagniard contour in Fig. 3(b), application of Cauchy theory may for $x>0$ now have to include the effect of a residue. Thus, for $(c>c_b, cx>c_b r_b)$ the term

$$B_k\left(-\frac{1}{c}\right)H\left(s-\frac{x}{c}-\sqrt{\omega}y\sqrt{1-\frac{c_b^2}{c^2}}\right) \tag{45}$$

must be added to (44). Similarly, for $(c>c_a, cx>c_a r_a)$ the term

$$A_k\left(-\frac{1}{c}\right)H\left(s-\frac{x}{c}-\sqrt{\frac{\omega}{3}}y\sqrt{1-\frac{c_a^2}{c^2}}\right) \tag{46}$$

must be included. As indicated by Fig. 5(a, b) the contributions (45) and (46) represent, respectively, rotational and dilatational disturbances generated in expanding wedge-like regions when the surface loads translate with trans and supersonic speeds. These disturbances are discontinuous along their fronts, and their evaluation is relatively simple because, as seen in (31), the quantities (q, a, b, T, R) are purely real.

Surface Behavior for Translating Loads

As in the stationary load case, the Cagniard contours collapse onto the branch cuts of (a, b) when $y=0$, and adjustments for the singularities at $q=\pm 1/c_R$ must be made. In this instance, the results for $y=0$, $x>0$ are

$$\mu u_1 = -\frac{S}{\pi x}\int_{x/c_a}^{s}\frac{4\alpha b^2 q^2}{\sqrt{\omega}D\left(q+\frac{1}{c}\right)}d\tau+\frac{S}{\pi x}\int_{x/c_b}^{s}\frac{\sqrt{\omega}\beta T^2}{D\left(q+\frac{1}{c}\right)}d\tau + \frac{N}{\pi x}\int_{x/x_a}^{x/c_b}\frac{2\alpha bqT}{D\left(q+\frac{1}{c}\right)}H(s-\tau)d\tau - \frac{N}{4}\frac{3\omega}{2-\sqrt{3}(\omega^2+3)}\frac{c_R c}{c_R-c}H\left(s-\frac{x}{c_R}\right)+A_1\left(-\frac{1}{c}\right) \times H\left(s-\frac{x}{c}\right)(c>c_a)+B_1\left(-\frac{1}{c}\right)H\left(s-\frac{x}{c}\right)(c>c_b) \tag{47a}$$

$$\mu u_2 = \frac{N}{\pi x}\int_{x/c_a}^{s}\frac{\sqrt{\omega}\alpha T^2}{D\left(q+\frac{1}{c}\right)}d\tau-\frac{N}{\pi x}\int_{x/c_b}^{s}\frac{4\alpha^2\beta q^2}{\sqrt{\omega}D\left(q+\frac{1}{c}\right)}d\tau - \frac{S}{\pi x}\int_{x/c_a}^{x/c_b}\frac{2\alpha bqT}{D\left(q+\frac{1}{c}\right)}H(s-\tau)d\tau + \frac{S}{4}\frac{3\omega}{2-\sqrt{3}(\omega^2+3)}\frac{c_R c}{c_R-c}H\left(s-\frac{x}{c_R}\right)+A_2\left(-\frac{1}{c}\right) \times H\left(s-\frac{x}{c}\right)(c>c_a)+B_2\left(-\frac{1}{c}\right)H\left(s-\frac{x}{c}\right)(c>c_b) \tag{47b}$$

where (41) and (31) hold, but the integration variable τ takes the place of s. In contrast to (40) for the stationary load case, (47) exhibit not only signals due to trans and supersonic speeds, but Rayleigh signals that are finite and discontinuous at their fronts.

Effects of Prestress

Equations (18*b*), (20) and (32)–(34) lead to the plot in Fig. 6, which shows that, for a compressive prestress, the dilatational wave speed $v_r c_a$ in κ^* increases with stress level, while the rotational and Rayleigh wave speeds $(v_r c_b, v_r c_R)$ decrease; the Rayleigh wave ceases to exist, of course, when (34) holds. For the tensile prestress, all three speeds would tend to the same limit at very high stress levels.

In terms of specific solution behavior, the discontinuities that travel over the surface at the Rayleigh wave speed in the moving-load case, *i.e.*, the fourth terms in (47*a,b*), exhibit an unbounded resonance when the load speed reaches the Rayleigh value ($c=c_R$). Behavior near resonance is also of note: For all $\sqrt{2\sqrt{3}-1}<\omega<\sqrt{2}$, the ω-dependent terms in the coefficients of the fourth terms are negative. Thus, in view of Fig. 2, the Rayleigh discontinuities change sign as v passes through the value $v_r c_R$. In the case of a classical isotropic linear elastic solid [13] this type of sign change has implications for the existence of solutions for sliding contact at subsonic speeds.

Indeed, the resonant and near-resonant response are readily extracted from the classical wave propagation results [7]; the difference here is that the phenomena depend both on the properties of the compressible neo-Hookean material and the prestress in the material.

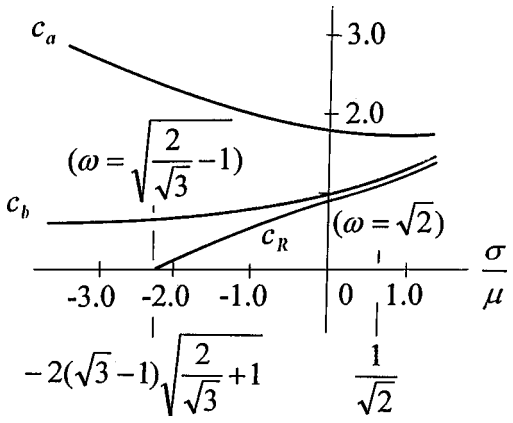


Fig. 6 Nondimensionalized wave speeds versus prestress

Some Comments

This article considered a transient plane-strain Green's function problem of shear and normal line loads applied to the surface of a prestressed highly elastic half-space. The loadings were either stationary and applied for an instant, or translated over the surface at an arbitrary constant speed. The prestress was aligned with the half-space surface, and could be either tensile or compressive, while an isotropic compressible neo-Hookean material served as the half-space.

Following [3], the problem was viewed as the superposition of infinitesimal deformations triggered by the surface loading upon (perhaps) finite deformations engendered by the prestress. Results by Beatty and Usmani [4] were used to formulate the problem, and exact solutions obtained for both deformations. In particular, integral transform and standard [12] inversion methods were used to derive the infinitesimal deformations. Complete expressions for the associated displacements both within and upon the half-space were given, and the separate contributions due to dilatational, rotational and head (bow) waves were identified, and the associated wave patterns described. In the case of surface loads moving at trans and supersonic speeds, the additional displacement contributions were also noted.

The results and the accompanying analysis showed that the isotropic compressible neo-Hookean solid behaved for small strains like a linear elastic material with Poisson's ratio 1/4. The prestress induced by the typical [3] *de facto* anisotropic behavior in the infinitesimal deformation was itself bounded above in tension by the value $\mu/\sqrt{2}$, where μ is the shear modulus. Tensions above this limit produced a negative Poisson effect in the infinitesimal response. Compressive prestress also exhibited the critical value $-2\mu(\sqrt{3}-1)\sqrt{2\sqrt{3}+1}$, below which the Rayleigh wave ceased to exist. Both critical values are of $O(\mu)$, but are relevant within the context of highly elastic response. Rayleigh waves in prestressed media have been discussed in more general terms [10,11] the latter work, indeed, showed that Rayleigh wave suppression can be associated with instability of the homogeneous deformation. This view is consistent with static analysis [4], which found instability criteria in terms of the roots of a polynomial that resembles a rationalization of a function of the Rayleigh type.

The unbounded resonance and sign change of surface displacements induced by the translation of surface loads at the Rayleigh speed that can be found in classical wave propagation [7] results was also seen in the present study. However, the resonance was now due both to material properties and to prestress levels. Such resonance in the classical case is known to be important in sliding contact [13]. Indeed, for sliding contact with friction at sub, trans, and supersonic speeds on a linear coupled thermoelastic material free of prestress, the unilateral Signorini conditions for contact cannot be satisfied without artifice for sliding speeds between the

Rayleigh and rotational wave values [14]. The same result holds for sliding contact on the prestressed isotropic compressible neo-Hookean material considered here [15]. Indeed, for compressive prestresses that exceed the critical value identified here, the Signorini conditions cannot be satisfied for subsonic sliding speeds. Subsonic sliding on a solid whose strain energy function reduces for infinitesimal strains to the classical isotropic form has also been treated [16], but the prestress deformations considered were not finite.

A plane-strain situation was treated here, and it is clear [3] that the effective material constants for the infinitesimal deformations are sensitive to the nature of the prestress as well as its level and to the properties of the highly elastic body. At present, therefore, efforts to study transient response both to surface and buried loads without the motion constraints imposed by plane strain are underway. Moreover, prestresses which exceed levels critical for Rayleigh wave annihilation are under consideration.

In closing, it should be noted that some basic equations used here, *e.g.*, (23), can be extracted from more general work [10,11]. The present results are intended as an illustration of a particular case of a highly elastic material in terms of exact and tractable full-field and surface solutions of the Green's function type. For example, the Cagniard functions (39) are similar in form to classical results [7] for isotropic elasticity.

References

[1] Achenbach, J. D., Pao, Y.-H., and Tiersten, H. F., 1976, ''Applications of Elastic Waves in Electrical Devices, Non-Destructive Testing, and Seismology,'' NSF Workshop Report, Northwestern University.

[2] Pao, Y.-H., 1987, ''Theory of Acoustoelasticity and Acoustoplasticity,'' *Solid Mechanics Research for Quantitative Non-Destructive Evaluation*, J. D. Achenbach and Y. Rajapakse, eds., Martinus Nijhoff, Dordrecht, pp. 257–273.

[3] Green, A. E., and Zerna, W., 1968, *Theoretical Elasticity*, 2nd ed., Oxford University, Oxford, UK.

[4] Beatty, M. F., and Usmani, S., 1975, ''On the Indentation of a Highly Elastic Half-Space,'' Q. J. Mech. Appl. Math., **28**, pp. 47–62.

[5] Truesdell, C., and Noll, W., 1965, ''The Non-Linear Field Theories of Mechanics,'' *Handbuch der Physik*, Vol. III/3, W. Flugge, ed., Springer-Verlag, Berlin.

[6] Hibbeler, R. C., 1997, *Mechanics of Materials*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ.

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

[8] Sneddon, I. N., 1972, *The Use of Integral Transforms*, McGraw-Hill, New York.

[9] van der Pol, B., and Bremmer, H., 1950, *Operational Calculus Based on the Two-Sided Laplace Integral*, Cambridge University Press, Cambridge, UK.

[10] Hayes, M. A., and Rivlin, R. S., 1961, ''Surface Waves in Deformed Elastic Material,'' Arch. Ration. Mech. Anal., **8**, pp. 358–380.

[11] Dowaikh, M. A., and Ogden, R. W., 1991, ''On Surface Waves in a Compressible Elastic Half-Space,'' Stab. Appl. Analy. Continuous Media, **1**, pp. 27–44.

[12] deHoop, A. T., 1960, ''A Modification of Cagniard's Method for Seismic Pulse Problems,'' Appl. Sci. Res., **B8**, pp. 349–356.

[13] Georgiadis, H. G., and Barber, J. R., 1993, ''On the Super-Rayleigh/Subseismic Elastodynamic Indentation Problem,'' J. Elast., **31**, pp. 141–161.

[14] Brock, L. M., and Georgiadis, H. G., 2000, ''Sliding Contact With Friction on a Thermoelastic Solid at Subsonic and Supersonic Speeds,'' J. Therm. Stresses, **23**, pp. 629–633.

[15] Brock, L. M., 1999, ''Sliding Contact With Friction at Arbitrary Constant Speeds on a Pre-Stressed Highly Elastic Half-Space,'' J. Elast., **57**, pp. 105–132.

[16] Brock, L. M., 1999, ''Rapid Sliding Indentation With Friction of a Pre-Stressed Thermoelastic Material,'' J. Elast., **53**, pp. 161–188.

Y. Huang
Department of Mechanical and Industrial Engineering, University of Illinois, Urbana, IL 61801

H. Gao
Division of Mechanics and Computation, Stanford University, Stanford, CA 94305

Intersonic Crack Propagation—Part I: The Fundamental Solution

Recent experiments of Rosakis et al. have clearly shown that the crack-tip velocity can exceed the shear wave speed for a crack tip propagating between two weakly bonded, identical and isotropic solids under shear-dominated loading. This has motivated recent theoretical and numerical studies on intersonic crack propagation. We have obtained analytically the fundamental solution for mode-II intersonic crack propagation in this paper. This fundamental solution can provide the general solutions for intersonic crack propagation under arbitrarily initial equilibrium fields. We have also developed a cohesive zone model to determine the crack-tip energy release for an intersonic shear crack.
[DOI: 10.1115/1.1357871]

1 Introduction

The present study is motivated by recent experiments on intersonic crack propagation by Rosakis et al. [1] who investigated shear dominated crack growth along weak planes in a brittle polyester resin under far-field asymmetrical loading. They observed crack propagation as fast as the longitudinal wave speed, c_l. This experiment is interesting because it has been widely believed that a brittle crack cannot propagate faster than the Rayleigh wave speed, c_R. The origin for this belief stems from the predictions of continuum mechanics with regard to the dynamic elastic solutions of the near-tip stress fields and energy release rates for various velocity regimes and different types of external loading. Freund [2] and Broberg [3] have elegantly summarized the solutions to dynamic crack propagation for all velocity regimes. For a mode-I crack, the physically admissible stress singularity and the energy release rate vanish for all crack velocities in excess of the Rayleigh wave speed, which implies that it is impossible for a mode-I crack to propagate at a velocity greater than the Rayleigh wave speed. In fact, Washabaugh and Knauss [4] demonstrated that the velocity of a mode-I crack tip propagating along a fabricated weak plane may asymptotically approach the Rayleigh wave speed. For a mode-II crack, however, the order of stress singularity becomes positive once the crack-tip velocity exceeds the shear wave speed, c_s, with a maximum value of 1/2 (square-root singularity) at a special speed $\sqrt{2}c_s$ at which the crack stops radiating ([5–7]). This singularity may lead to a positive crack-tip energy release rate for mode-II intersonic crack propagation once a cohesive view of fracture is adopted ([3,6]), i.e., the crack tip is viewed not as a point singularity but as a finite cohesive zone.

Evidence of shear crack propagation in excess of the shear wave speed has also been provided from observations of shallow crustal earthquakes ([8–10]). Among recent earthquakes where super-shear rupture velocity has been reported is the M7.4 earthquake at Kocaeli, Turkey on Aug. 17, 1999 ([11]). The late arrival of laboratory experiments on intersonic fracture ([1,12]) is due, in part, to the fact that a crack in elastic homogeneous and isotropic solids always kinks or branches out, deviating from the initial crack plane and having a zigzag crack path, once the crack-tip velocity exceeds only $0.3\sim0.4c_s$ ([2,13,14]). A wavy crack instability occurs at low crack velocities and prevents an exploration of the full range of possible velocities. In fact, the only possibility of attaining intersonic crack propagation is to introduce a weak path (a layer of lower toughness) so that crack growth is confined to this path ([1,12]).

There are also analytical and numerical studies on shear intersonic crack propagation. Freund [5], Burridge et al. [15], and Simonov [16] investigated various aspects of intersonic crack growth such as the stability regime as well as stress singularity at crack tip and at the radiating wave fronts. Broberg [17] solved the problem of self-similar propagation of an intersonic shear crack symmetrically expanded at constant speed from zero initial length. Yu and Suo [18] used the cohesive model to study the permissible velocity zones for intersonic crack growth along a bimaterial interface. Andrews [19] used a slip weakening model to investigate shear crack growth along a weak interface and found that the shear crack approaching the Rayleigh wave speed induces a microcrack that moves at speeds exceeding the shear wave speed. Similar mechanisms were also reported in the cohesive finite element simulations ([20,21]) and atomic simulations ([22]) of intersonic shear fracture. Gao et al. [23] compared the atomic simulations with the continuum analysis of intersonic crack propagation and established that, without any parameter fitting, they agree very well.

In this paper we obtained analytically a fundamental solution for intersonic shear crack propagation. A semi-infinite crack in an infinite solid is subjected to a pair of suddenly applied concentrate shear forces on the crack faces. The crack tip starts to propagate at a velocity between the shear and longitudinal wave speeds. The method of analytic continuation ([24]) for subsonic crack growth is extended for intersonic crack propagation. The analytical expressions are obtained for the shear stress ahead of the propagating crack tip and the sliding displacement on the crack faces, which is useful for the Part II of this study for a suddenly stopping intersonic shear crack. A cohesive zone model is also developed to determine the crack-tip energy release rate in intersonic crack propagation.

2 The Fundamental Solution

An infinite solid containing a semi-infinite crack on the negative x-axis is subjected to plane-strain deformation. The solid is linear elastic and isotropic, with the shear modulus μ and Poisson's ratio ν. For time $t\leqslant 0$, the solid is stress free and at rest everywhere, and the crack tip is at the origin (0,0) of the stationary coordinate system (x,y). The positive x-axis represents the preexisting weak plane for the crack to propagate. At time $t=0$, the crack tip begins to move intersonically at a constant velocity v in the positive x-direction, where v is between the shear wave speed c_s and longitudinal wave speed c_l. As the tip moves away, a pair of concentrated shear forces (in the x-direction) of constant magnitude τ^* (per unit length in the z-direction) is left at the

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 20, 2000; final revision, Nov. 9, 2000. Associate Editor: L. T. Wheeler. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

origin (the initial crack-tip location). Similar to subsonic crack growth ([2]), this is the fundamental solution for an intersonic shear crack subjected to an arbitrary initial equilibrium field.

The displacement potentials ϕ and ψ are introduced such that the in-plane displacements can be written as ([2])

$$u_x=\frac{\partial\phi}{\partial x}+\frac{\partial\psi}{\partial y},\quad u_y=\frac{\partial\phi}{\partial y}-\frac{\partial\psi}{\partial x}. \tag{1}$$

The in-plane stress components can be written in terms of ϕ and ψ as

$$\begin{aligned}\sigma_{xx}&=\mu\left(\frac{c_l^2}{c_s^2}\nabla^2\phi-2\frac{\partial^2\phi}{\partial y^2}+2\frac{\partial^2\psi}{\partial x\partial y}\right),\\ \sigma_{yy}&=\mu\left(\frac{c_l^2}{c_s^2}\nabla^2\phi-2\frac{\partial^2\phi}{\partial x^2}-2\frac{\partial^2\psi}{\partial x\partial y}\right),\\ \sigma_{xy}&=\mu\left(2\frac{\partial^2\phi}{\partial x\partial y}+\frac{\partial^2\psi}{\partial y^2}-\frac{\partial^2\psi}{\partial x^2}\right),\end{aligned} \tag{2}$$

where ∇^2 is the Laplace operator, the shear and longitudinal wave speeds $c_s=\sqrt{\mu/\rho}$ and $c_l=\sqrt{(\kappa+1)/(\kappa-1)}c_s$, ρ the mass density, and $\kappa=3-4\nu$ for plane-strain deformation. A coordinate system (ξ,y) moving with the crack tip is introduced as

$$\xi=x-vt, \tag{3}$$

where v is the crack-tip velocity. The equation of motion in the moving coordinate system becomes ([2])

$$\begin{aligned}\alpha_l^2\frac{\partial^2\phi}{\partial\xi^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{2v}{c_l^2}\frac{\partial^2\phi}{\partial\xi\partial t}-\frac{1}{c_l^2}\frac{\partial^2\phi}{\partial t^2}=0,\\ -\hat{\alpha}_s^2\frac{\partial^2\psi}{\partial\xi^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{2v}{c_s^2}\frac{\partial^2\psi}{\partial\xi\partial t}-\frac{1}{c_s^2}\frac{\partial^2\psi}{\partial t^2}=0,\end{aligned} \tag{4}$$

where

$$\alpha_l=\sqrt{1-\frac{v^2}{c_l^2}},\quad \hat{\alpha}_s=\sqrt{\frac{v^2}{c_s^2}-1}. \tag{5}$$

The coefficient of $\partial^2\psi/\partial\xi^2$ has become negative in intersonic crack propagation, leading to shock waves associated with the intersonically moving crack tip.

Only the upper half-plane $(y\geqslant 0)$ is analyzed due to symmetry. The boundary conditions in the moving coordinate system can be written as

$$\begin{aligned}\sigma_{yy}(\xi,y=0,t)&=0,\\ \sigma_{xy}(\xi,y=0,t)&=\tau_+(\xi,t)-\tau^*\delta(\xi+vt)H(t)H(-\xi),\\ u_x(\xi,y=0,t)&=u_-(\xi,t)\end{aligned} \tag{6}$$

for $-\infty<\xi<\infty$, where τ_+ is the unknown shear stress ahead of the moving crack tip ($\xi>0$, $y=0$), which can be taken as zero for $\xi<0$; δ and H are the Dirac delta function and the unit step function, respectively; and u_- is the unknown sliding displacement on the crack face ($\xi<0$, $y=0$), which vanishes for $\xi>0$.

The Laplace transform and two-sided Laplace transform are applied with respect to time t and moving coordinate ξ, respectively, i.e.,

$$\begin{aligned}\hat{\phi}(\xi,y,s)&=\int_0^\infty\phi(\xi,y,t)e^{-st}dt,\\ \Phi(\zeta,y,s)&=\int_{-\infty}^\infty\hat{\phi}(\xi,y,s)e^{-s\zeta\xi}d\xi.\end{aligned} \tag{7}$$

For the displacement potential ϕ, the two-sided Laplace transform exists only in the strip $-1/(c_l-v)<\mathrm{Re}(\zeta)<1/(c_l+v)$, which is identical with that for subsonic crack propagation ([2]), where Re stands for the real part of a complex number. The Laplace transform for ψ is defined in the same way as in (7), but it converges only in the strip $0<\mathrm{Re}(\zeta)<1/(c_s+v)$ because the crack tip propagates faster than the shear wave speed. Accordingly, the common of these two strips, $0<\mathrm{Re}(\zeta)<1/(c_s+v)$, is the domain within which the Laplace transform converges.

The governing Eqs. (4) become the ordinary differential equations for Φ and Ψ after the Laplace transform, and they have the general solution in the upper half-plane $(y\geqslant 0)$ as

$$\Phi=\frac{P(\zeta)}{s^3}e^{-\alpha(\zeta)sy},\quad \Psi=\frac{Q(\zeta)}{s^3}e^{-\beta(\zeta)sy}, \tag{8}$$

where the functions P and Q are to be determined by boundary conditions, and α and β are given by

$$\begin{aligned}\alpha&=-i\alpha_l\sqrt{\zeta+\frac{1}{c_l-v}}\sqrt{\zeta-\frac{1}{c_l+v}},\\ \beta&=-\hat{\alpha}_s\sqrt{\zeta-\frac{1}{v-c_s}}\sqrt{\zeta-\frac{1}{v+c_s}}.\end{aligned} \tag{9}$$

Here, α has a branch cut on each side of the strip $0<\mathrm{Re}(\zeta)<1/(c_s+v)$, i.e., from $-1/(c_l-v)$ to $-\infty$ and from $1/(c_l+v)$ to ∞. The branch cuts for β, however, are both to the right of the strip, i.e., from $1/(v+c_s)$ and $1/(v-c_s)$ to ∞, which are different from that for subsonic crack growth ([24]). These branch cuts ensure $\mathrm{Re}(\alpha)\geqslant 0$ and $\mathrm{Re}(\beta)\geqslant 0$ for ζ in the strip.

The Laplace transform of the boundary conditions (6) gives

$$\begin{aligned}\left[\frac{c_l^2}{c_s^2}(\alpha^2+\zeta^2)-2\zeta^2\right]P+2\beta\zeta Q&=0,\\ -2\alpha\zeta P+(\beta^2-\zeta^2)Q&=\frac{T_+(\zeta)}{\mu}+\frac{\tau^*}{\mu v}\frac{1}{\zeta-\frac{1}{v}},\\ \zeta P-\beta Q&=U_-,\end{aligned} \tag{10}$$

where

$$\begin{aligned}T_+(\zeta)&=s\int_0^\infty\hat{\tau}_+(\xi,s)e^{-s\zeta\xi}d\xi,\\ U_-(\zeta)&=s^2\int_{-\infty}^0\hat{u}_-(\xi,s)e^{-s\zeta\xi}d\xi\end{aligned} \tag{11}$$

are analytic for $\mathrm{Re}(\zeta)\geqslant 0$ and $\mathrm{Re}(\zeta)\leqslant 1/(c_l+v)$, respectively, on the ζ-plane, $\hat{\tau}_+(\xi,s)$ is the transform of the shear stress τ_+ ahead of the moving crack tip with respect to time, and $\hat{u}_-(\xi,s)$ is the transformed sliding displacement u_- on the crack face. Elimination of P and Q from three equations in (10) yields

$$T_+(\zeta)+\frac{\tau^*}{v}\frac{1}{\zeta-\frac{1}{v}}=-\mu\frac{c_s^2}{v^2}\frac{R(\zeta)}{\left(\zeta-\frac{1}{v}\right)^2\beta}U_-, \tag{12}$$

where

$$R(\zeta)=4\alpha\beta\zeta^2+\left[2\zeta^2-\frac{v^2}{c_s^2}\left(\zeta-\frac{1}{v}\right)^2\right]^2, \tag{13}$$

and α and β are given in (9). The function $R(\zeta)$ has four roots ([24]), namely $\zeta=1/(v-c_R)$, $1/(v+c_R)$, and a double root at $1/v$, where c_R is the Rayleigh wave speed. These roots are all to the right of the strip of convergence, $0<\mathrm{Re}(\zeta)<1/(c_s+v)$. Unlike subsonic crack growth, the function $R(\zeta)$ is no longer analytic for large ζ in intersonic crack propagation, and has to be decomposed differently from that in subsonic crack growth ([2]). We define a new function s by

$$s(\zeta)=\frac{1}{4i\alpha_l\hat{\alpha}_s}\left(\frac{\zeta-\dfrac{1}{c_l+v}}{\zeta+\dfrac{1}{c_l-v}}\right)^{1/2}\times\frac{R(\zeta)}{\left(\zeta-\dfrac{1}{v-c_R}\right)\left(\zeta-\dfrac{1}{v+c_R}\right)\left(\zeta-\dfrac{1}{v}\right)^2},\tag{14}$$

which has no roots on the ζ-plane, but has branch cuts from $-1/(c_l-v)$ to $-\infty$, and from $1/(c_l+v)$, $1/(v+c_s)$, and $1/(v-c_s)$ to ∞. For large ζ,

$$s\sim 1+[(2-v^2/c_s^2)^2/(4i\alpha_l\alpha_s)]((\zeta-1/(c_l+v))/(\zeta+(1/c_l-v)))^{1/2},$$

and is therefore not analytic in the entire range of intersonic crack propagation ($c_s<v<c_l$) except when the crack-tip velocity v is $\sqrt{2}$ times the shear wave speed c_s. In fact, it is well known ([2]) that only at this particular velocity of $\sqrt{2}c_s$ stresses near an intersonic shear crack tip have the conventional square-root singularity, and the crack-tip energy release rate remains finite and nonzero.

In order to solve T_+ and U_- from (12), we decompose $s(\zeta)$ as

$$s(\zeta)=s_+(\zeta)s_-(\zeta),\tag{15}$$

where s_+ is analytic for $\mathrm{Re}(\zeta)\geqslant 0$, while s_- is analytic for $\mathrm{Re}(\zeta)\leqslant 1/(c_l+v)$. For any ζ in the strip $0\leqslant\mathrm{Re}(\zeta)\leqslant 1/(c_l+v)$, the Cauchy's integral formula gives

$$\log s(\zeta)=\frac{1}{2\pi i}\oint_\Gamma\frac{\log s(z)}{z-\zeta}dz,\tag{16}$$

where the closed counterclockwise contour of integration Γ is the boundary of the strip, i.e., the two vertical straight lines $\Gamma_1(\mathrm{Re}(\zeta)=0)$ and $\Gamma_2(\mathrm{Re}(\zeta)=1/(c_l+v))$ in the complex ζ-plane. For subsonic crack propagation, $s(\zeta)$ is decomposed by taking $\log s_+(\zeta)$ and $\log s_-(\zeta)$ as the integrations on these two straight lines Γ_1 and Γ_2, respectively. This standard method, however, does not work for intersonic crack propagation because $s(\zeta)$ is no longer analytic for large ζ such that the integration on each straight line would diverge. In the following, we adopt a different method to decompose $s(\zeta)$ for intersonic crack propagation. The derivative of (16) with respect to ζ is

$$\frac{s'(\zeta)}{s(\zeta)}=\frac{1}{2\pi i}\oint_\Gamma\frac{\log s(z)}{(z-\zeta)^2}dz.\tag{17}$$

According to the decomposition in (15), the left-hand side of (17) can be written as $s'_+/s_++s'_-/s_-$. Therefore, we may take

$$\frac{s'_+(\zeta)}{s_+(\zeta)}=\frac{1}{2\pi i}\int_{\Gamma_1}\frac{\log s(z)}{(z-\zeta)^2}dz,\quad \frac{s'_-(\zeta)}{s_-(\zeta)}=\frac{1}{2\pi i}\int_{\Gamma_2}\frac{\log s(z)}{(z-\zeta)^2}dz.\tag{18}$$

A similar decomposition was used by Noble [25] in the study of functions that are neither bounded nor single-valued at infinity. Based on Cauchy's theorem, the integration path $\Gamma_1(\mathrm{Re}(\zeta)=0)$ for s'_+/s_+ can be augmented to a path on both sides of the branch cut between $-1/(c_l-v)$ and $-\infty$. Further integration with respect to ζ gives

$$\frac{s_+(\zeta)}{s_+(0)}=\exp\left[-\frac{\zeta}{\pi}\int_{1/(c_l-v)}^{\infty}\arctan V_+(r)\frac{dr}{r(r+\zeta)}\right],\tag{19}$$

where

$$V_\pm(r)=\frac{\left(2r^2-\dfrac{v^2}{c_s^2}\left(r\pm\dfrac{1}{v}\right)^2\right)^2}{4\alpha_l\hat{\alpha}_s r^2\sqrt{r\mp\dfrac{1}{c_l-v}}\sqrt{r\pm\dfrac{1}{c_l+v}}\sqrt{\left|r\pm\dfrac{1}{v-c_s}\right|}\sqrt{\left|r\pm\dfrac{1}{v+c_s}\right|}}.\tag{20}$$

As $\zeta\to+\infty$, $s_+(\zeta)/s_+(0)$ is on the order of $\zeta^{q-1/2}$, where

$$q=\frac{1}{\pi}\arctan\frac{4\alpha_l\hat{\alpha}_s}{\left(2-\dfrac{v^2}{c_s^2}\right)^2}\tag{21}$$

is the order of stress singularity near an intersonic shear crack tip ([2]), and it is always less than 1/2 except at a single crack-tip velocity $v=\sqrt{2}c_s$. Similarly, the integration path $\Gamma_2(\mathrm{Re}(\zeta)=1/(c_l+v))$ for s'_-/s_- in (18) can be augmented to both sides of the branch cut between $1/(c_l+v)$ and ∞. This leads to

$$\frac{s_-(\zeta)}{s_-(0)}=\exp\left[-\frac{\zeta}{\pi}\int_{1/(c_l+v)}^{\infty}\times\left[\frac{\pi}{2}+\left(\frac{\pi}{2}-\arctan V_-(r)\right)H_1(r)\right]\frac{dr}{r(r-\zeta)}\right],\tag{22}$$

where the function V_- is given in (20), and $H_1(r)=H(1/(v+c_s)-r)-H(r-1/(v-c_s))$. As $\zeta\to-\infty$, $s_-(\zeta)/s_-(0)$ is on the order of $|\zeta|^{1/2-q}$.

The function $s(\zeta)$ in (14) can then be decomposed as

$$s(\zeta)=s(0)\frac{s_+(\zeta)}{s_+(0)}\frac{s_-(\zeta)}{s_-(0)}=\frac{v^2(v^2-c_R^2)}{4c_s^4\hat{\alpha}_s\left(1+\dfrac{v}{c_l}\right)}\frac{s_+(\zeta)}{s_+(0)}\frac{s_-(\zeta)}{s_-(0)},\tag{23}$$

where $s(0)$ is evaluated from (14). Finally, Eq. (12) governing the unknowns T_+ and U_- can be decomposed as

$$\frac{s_+(0)}{s_+(\zeta)}\frac{T_+}{\sqrt{(c_l-v)\zeta+1}}+\left[\frac{s_+(0)}{s_+(\zeta)}\frac{1}{\sqrt{(c_l-v)\zeta+1}}-\frac{s_+(0)}{s_+\left(\dfrac{1}{v}\right)}\sqrt{\frac{v}{c_l}}\right]\frac{\tau^*}{v\zeta-1}=\frac{i\mu}{c_s}\frac{s_-(\zeta)}{s_-(0)}\frac{[(v-c_R)\zeta-1][(v+c_R)\zeta-1]U_-}{\sqrt{(c_l+v)\zeta-1}\sqrt{(v-c_s)\zeta-1}\sqrt{(v+c_s)\zeta-1}}-\frac{s_+(0)}{s_+\left(\dfrac{1}{v}\right)}\sqrt{\frac{v}{c_l}}\frac{\tau^*}{v\zeta-1}.\tag{24}$$

Its left-hand side is analytic to the right ($\mathrm{Re}(\zeta)>0$) of the strip since $\zeta=1/v$ is a removable singularity, while its right-hand side is analytic to the left of the strip ($\mathrm{Re}(\zeta)<1/(c_l+v)$). Therefore, (24) defines an entire function. The Laplace transform of the stress ($\tau\sim r^{-q}$) and displacement ($u\sim r^{1-q}$) near an intersonic shear crack tip ($r\to 0$) gives the asymptotic behavior of the transformed stress and displacement for $|\zeta|\to\infty$ as $T_+\sim\zeta^{q-1}$ and $U_-\sim\zeta^{q-2}$. It is then straightforward to show that both sides of (24) approach zero as $|\zeta|\to\infty$. Based on the Liouville's theorem, both sides of (24) must vanish for all ζ. This gives

$$T_+=\left[\frac{s_+(\zeta)}{s_+\left(1/v\right)}\sqrt{\frac{v}{c_l}}\sqrt{(c_l-v)\zeta+1}-1\right]\frac{\tau^*}{v\zeta-1}, \tag{25}$$

$$U_-=-\frac{ic_s}{\mu}\frac{s_+(0)}{s_+\left(1/v\right)}\frac{s_-(0)}{s_-(\zeta)}\times\frac{\sqrt{(c_l+v)\zeta-1}\sqrt{(v-c_s)\zeta-1}\sqrt{(v+c_s)\zeta-1}}{[(v-c_R)\zeta-1][(v+c_R)\zeta-1]}\sqrt{\frac{v}{c_l}}\frac{\tau^*}{v\zeta-1}. \tag{26}$$

The functions P and Q can be obtained by substituting T_+ in (25) and U_- in (26) into (10), which in turn give solutions for the transformed displacement potentials Φ and Ψ in (8). The means of Cagniard-de Hoop method is then used to invert the double transforms for ϕ and ψ. In the following we focus on the shear stress ahead of the moving crack tip, $\tau_+(\xi>0,t)$. Inversion of the two-sided Laplace transform is considered first,

$$\hat{\tau}_+(\xi,s)=\frac{1}{2\pi i}\int_{\zeta_0-i\infty}^{\zeta_0+i\infty}T_+(\zeta)e^{s\zeta\xi}d\zeta, \tag{27}$$

where ζ_0 is a real number in the strip, i.e., $0<\zeta_0<1/(c_l+v)$. Based on Cauchy's theorem, the integration path can be augmented to a path on both sides of the branch cut between $-1/(c_l-v)$ and $-\infty$, and $\hat{\tau}_+$ becomes

$$\hat{\tau}_+(\xi,s)=-\frac{1}{\pi}\int_{1/(c_l-v)}^{+\infty}\mathrm{Im}[T_+(-\eta)]e^{-s\eta\xi}d\eta, \tag{28}$$

where Im stands for the imaginary part of a complex number, and the integrand is understood as the limiting value just above the branch cut. The inverse Laplace transform of (28) with respect to s can be obtained by observation since the inversion of $e^{-s\eta\xi}$ is simply the Dirac delta function $\delta(t-\xi\eta)$. In conjunction with the analytical expression of T_+ in (25), the shear stress ahead of the propagating crack tip is found as

$$\tau_+(\xi>0,t)=\frac{4c_s^3}{\pi c_l^{3/2}}\tau^*\sqrt{\frac{v}{\xi}}\frac{s_+(0)}{s_+\left(\frac{1}{v}\right)}\frac{s_-(0)}{s_-\left(-\frac{t}{\xi}\right)}t^2\times[(c_l+v)t+\xi]H[(c_l-v)t-\xi]*\frac{\sqrt{(c_l-v)t-\xi}\sqrt{(v-c_s)t+\xi}\sqrt{(v+c_s)t+\xi}}{[(v-c_R)t+\xi][(v+c_R)t+\xi](vt+\xi)^3}. \tag{29}$$

It has the asymptotic form near the crack tip ($\xi\to 0+$)

$$\tau_+(\xi\to 0+,t)=\frac{4\alpha_l\hat{\alpha}_s c_s^4 f(v)}{\pi\sqrt{c_l v^3}(v^2-c_R^2)}\frac{\tau^*}{vt}\frac{s_+(0)}{s_+\left(\frac{1}{v}\right)}\left[\frac{(v^2-c_s^2)t}{(c_l+v)\xi}\right]^q, \tag{30}$$

where α_l and $\hat{\alpha}_s$ are defined by (5), the power of stress singularity q is given in (21), and the function f is related to the crack-tip velocity v by

$$f(v)=\exp\left[\int_{1/(c_l+v)}^{1/(v+c_s)}\frac{f_0(r)}{\pi\gamma}dr-\int_{1/(v-c_s)}^{+\infty}\frac{f_0(r)}{\pi r}dr\right]. \tag{31}$$

Here the function f_0 is given by

$$f_0(r)=\tan^{-1}\left[4\alpha_l\hat{\alpha}_s\frac{f_1(r)-\left(2-\frac{v^2}{c_s^2}\right)^2}{\left(2-\frac{v^2}{c_s^2}\right)^2 f_1(r)+16\alpha_l^2\hat{\alpha}_s^2}\right], \tag{32}$$

and

$$f_1(r)=\frac{\left[2r^2-\frac{v^2}{c_s^2}\left(r-\frac{1}{v}\right)^2\right]^2}{r^2\sqrt{\left(r+\frac{1}{c_l-v}\right)\left(r-\frac{1}{c_l+v}\right)\left|r-\frac{1}{v-c_s}\right|\left|r-\frac{1}{v+c_s}\right|}}. \tag{33}$$

It is observed that the near-tip stress decreases with time via a power of $1-q$, i.e., $\tau\sim 1/t^{1-q}$.

The sliding displacement on the crack face, $u_-(\xi<0,t)$, is important to the study of a suddenly stopping crack in Part II of this paper. The inversion of U_- in (26), together with the Cagniard-de Hoop method, gives

$$u_-(\xi<0,t)=\frac{1}{\pi}PV\int_{1/(c_l+v)}^{-t/\xi}\mathrm{Im}[U_-(\eta)]d\eta H[(c_l+v)t+\xi], \tag{34}$$

where PV stands for the Cauchy principal value integral; the integrand is understood as the limiting value just above the branch cut from $1/(c_l+v)$ to ∞, and $\mathrm{Im}[U_-(\eta)]$ is obtained from (26) as

$$\mathrm{Im}[U_-(\eta)]=\frac{\sqrt{c_l v}}{c_s^2}\frac{\tau^*}{\mu}\frac{s_+(\eta)}{s_+\left(\frac{1}{v}\right)}(1-v\eta)\sqrt{(c_l-v)\eta+1}\frac{F_n(\eta)}{F_d(\eta)}, \tag{35}$$

where

$$\begin{aligned}F_n(\eta)=&4\eta^2\sqrt{(c_l-v)\eta+1}\sqrt{(c_l+v)\eta-1}[(v-c_s)\eta-1]\\&\times[(v+c_s)\eta-1]\\&-c_l c_s\left[2\eta^2-\frac{1}{c_s^2}(v\eta-1)^2\right]^2\\&\times\sqrt{1-(v-c_s)\eta}\sqrt{(v+c_s)\eta-1}*H[1-(v-c_s)\eta]\\&\times H[(v+c_s)\eta-1],\end{aligned} \tag{36}$$

$$\begin{aligned}F_d(\eta)=&16\eta^4[(c_l-v)\eta+1][(c_l+v)\eta-1][(v-c_s)\eta-1]\\&\times[(v+c_s)\eta-1]+c_l^2c_s^2\left[2\eta^2-\frac{1}{c_s^2}(v\eta-1)^2\right]^4.\end{aligned} \tag{37}$$

It is noted that the denominator $F_d(\eta)$ has simple poles at $1/(v-c_R)$ and $1/(v+c_R)$, as well as a double pole at $1/v$.

3 Discussion

The fundamental solution in the previous section can provide the general solution for an intersonic shear crack under an arbitrary initial equilibrium field. This is because that the process of crack propagation is essentially the negation of the equilibrium traction distribution during dynamic crack growth (see [2] for details).

The stress singularity around an intersonic shear crack tip is always weaker than the conventional square-root singularity except at a single crack-tip velocity of $v=\sqrt{2}c_s$. Only when the crack-tip velocity becomes $\sqrt{2}$ times the shear wave speed the square-root singularity is preserved and the crack-tip energy release rate remains finite and nonzero. In fact, this is consistent with the first experimental observation on the shear-dominated

intersonic crack propagation ([1]), in which the crack-tip velocity was observed to approach a steady-state value of $\sqrt{2}c_s$. The molecular dynamics simulation of the shear-dominated intersonic crack propagation ([22]) has also shown that the crack-tip velocity tends to approach $\sqrt{2}c_s$. At this particular crack-tip velocity $v=\sqrt{2}c_s$, the stress and displacement fields near the propagating crack tip are

$$\sigma_{xx}=\frac{1+\alpha_l^2}{\alpha_l}\frac{K_{II}}{\sqrt{2\pi}}\mathrm{Im}[(\xi+i\alpha_l y)^{-1/2}],\quad \sigma_{yy}=0,$$

$$\sigma_{xy}=\frac{K_{II}}{\sqrt{2\pi}}\mathrm{Re}[(\xi+i\alpha_l y)^{-1/2}], \tag{38}$$

$$u_x=\frac{1}{\alpha_l}\frac{K_{II}}{\mu\sqrt{2\pi}}\mathrm{Im}[(\xi+i\alpha_l y)^{1/2}],\quad u_y=\frac{K_{II}}{\mu\sqrt{2\pi}}\mathrm{Re}[(\xi+i\alpha_l y)^{1/2}], \tag{39}$$

where α_l, defined in (5), becomes $\alpha_l=\sqrt{1-2c_s^2/c_l^2}$; K_{II} is the crack-tip stress intensity factor for $v=\sqrt{2}c_s$,

$$K_{II}(v=\sqrt{2}c_s)=\lim_{\xi\to 0+}\tau_+(\xi,t)\sqrt{2\pi\xi}. \tag{40}$$

It is clearly observed from (38) that stresses are singular only at the crack tip for this particular crack-tip velocity of $\sqrt{2}c_s$. This indicates that $\sqrt{2}c_s$ is the radiation-free crack-tip velocity, i.e., there is no shock wave emanating from the crack tip at this velocity.

From the asymptotic expression (30), the stress intensity factor for $v=\sqrt{2}c_s$ is found as

$$K_{II}(v=\sqrt{2}c_s)=\frac{2c_s^3}{c_l(2c_s^2-c_R^2)}\sqrt{1-\frac{\sqrt{2}c_s}{c_l}}\frac{s_+(0)}{s_+\left(\dfrac{1}{\sqrt{2}c_s}\right)}f(\sqrt{2}c_s)K_0, \tag{41}$$

where the function f is defined in (31), and $K_0=K_{II}(v\to 0)=\tau^*\sqrt{2/(\pi l)}$ is the equilibrium stress intensity factor for a stationary crack subjected to a pair of shear forces τ^* at the same distance of $l=\sqrt{2}c_s t$ behind the crack tip. For Poisson's ratio $\nu=1/3$, the factor before K_0 is 0.464, i.e., $K_{II}(v=\sqrt{2}c_s)=0.464K_{II}(v\to 0)$, such that the near-tip shear stress is slightly less than one half of its counterpart for a stationary crack. This also indicates that the stress intensity factor reaches its ''steady-state'' limit ($0.464K_0$) instantaneously.

The crack-tip energy release rate for $v=\sqrt{2}c_s$ can be obtained via the energy flux integral ([2]) and the asymptotic crack-tip field in (38) and (39) as ([7])

$$G(v=\sqrt{2}c_s)=\frac{K_{II}^2(v=\sqrt{2}c_s)}{4\mu\sqrt{1-\dfrac{2c_s^2}{c_l^2}}}=\frac{1-\nu^2}{E}\frac{K_{II}^2(v=\sqrt{2}c_s)}{2\sqrt{\nu(1-\nu)}}. \tag{42}$$

This relation between the crack-tip energy release rate and the stress intensity factor at $v=\sqrt{2}c_s$ is different from its counterpart for a stationary crack tip ($v\to 0$) by the factor $1/2\sqrt{\nu(1-\nu)}$, which is 1.06 for Poisson's ratio $\nu=1/3$. Therefore, $G(v=\sqrt{2}c_s)=0.228G_0$, where $G_0=((1-\nu^2)/E)K_0^2$ is the energy release rate for a stationary crack tip subjected to a pair of shear forces τ^* at the same distance of $l=\sqrt{2}c_s t$ behind the crack tip.

For the crack-tip velocity other than $v=\sqrt{2}c_s$, the stress singularity around an intersonic shear crack tip is weaker than the conventional square-root singularity, which leads to a vanishing crack-tip energy release rate. Broberg [6] has introduced a process zone model of Dugdale-Barenblatt type to remedy the pathology of zero crack-tip energy release rate, when the crack-tip velocities are different from $\sqrt{2}c_s$. This provides an energy absorption mechanism near the crack tip. We have used such a model of shear cohesive process zone, with details of the model given in the Appendix. The shear cohesive zone propagates with the moving crack tip. The constant shear cohesive strength is denoted by τ_c. The energy release rate, G, is evaluated from the energy flux ([2]) into the cohesive zone. Let $G_0=((1-\nu^2)/E)K_0^2$ denote the crack-tip energy release rate for a stationary crack tip subjected to a pair of shear forces τ^* at the same distance of vt behind the crack tip, and $K_0=\tau^*\sqrt{2/(\pi vt)}$ be the corresponding stress intensity factor. The normalized energy release rate, G/G_0, is given by

$$\frac{G}{G_0}=\left[q\frac{c_s^4}{\sqrt{c_l v^3(v^2-c_R^2)}}\frac{s_+(0)}{s_+(1/v)}f(v)\right]^{1/q}\times\left[\sqrt{(1-\hat{\alpha}_s^2)^4+16\alpha_l^2\hat{\alpha}_s^2}\frac{\tau^*}{\tau_c vt}\right]^{1/q-2}*\times\frac{16\alpha_l\hat{\alpha}_s^2(\alpha_l^2+\hat{\alpha}_s^2)}{1-q}\frac{v^2-c_s^2}{(c_l+v)v}. \tag{43}$$

At the radiation-free crack-tip velocity $v=\sqrt{2}c_s$, the above expression degenerates to that in (42). The normalized energy release rate does not depend on time, indicating the ''steady-state'' limit is reached instantaneously. At other velocities, however, the normalized energy release rate becomes time-dependent. Figure 1 shows the normalized energy release rate versus the intersonic crack-tip velocity for time $\tau_c vt/\tau^*=1$ and 10, and Poisson's ratio $\nu=1/3$. The normalized energy release rate decreases as time increases, which means the ''steady-state'' limit is not reached instantaneously, and the energy release rate decreases faster than t^{-1} represented in G_0. This does not occur in subsonic crack growth and is unique for intersonic crack propagation. In fact, as time approaches infinity, the normalized energy release rate approaches zero in the entire range of intersonic crack propagation except at the radiation-free velocity $v=\sqrt{2}c_s$.

The length of the cohesive zone, L, is obtained from the requirement to eliminate the crack-tip singularity

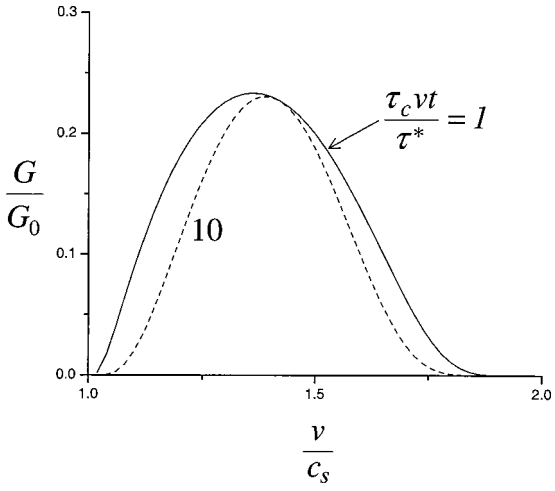


Fig. 1 The energy release rate, G, is shown versus the crack-tip velocity v at time $\tau_c vt/\tau^*=1$ and 10; G_0 is the energy release rate for a stationary crack tip subjected to a pair of shear forces τ^* at the same distance of vt behind the crack tip, τ_c is the cohesive strength, c_s the shear wave speed, and Poisson's ratio $\nu=1/3$

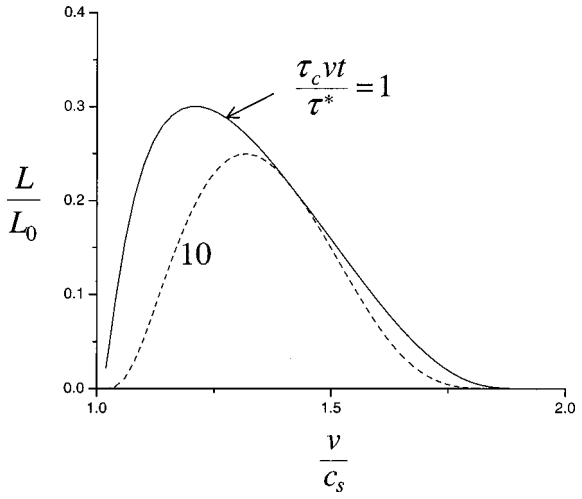


Fig. 2 The length of cohesive zone, L, is shown versus the crack-tip velocity v at time $\tau_c vt/\tau^*=1$ and 10; L_0 is the cohesive zone length for a stationary crack tip subjected to a pair of shear forces τ^* at the same distance of vt behind the crack tip, τ_c is the cohesive strength, c_s the shear wave speed, and Poisson's ratio $\nu=1/3$

$$L=\left[q\sqrt{(1-\hat{\alpha}_s^2)^4+16\alpha_l^2\hat{\alpha}_s^2}\frac{c_s^4}{\sqrt{c_l v^3}(v^2-c_R^2)}\frac{\tau^*}{\tau_c vt}\frac{s_+(0)}{s_+\left(\frac{1}{v}\right)}f(v)\right]^{1/q} \times\frac{v^2-c_s^2}{c_l+v}t. \tag{44}$$

Let $L_0=\tau^{*2}/(4\tau_c^2 vt)$ denote the corresponding cohesive zone length for a stationary crack tip subjected a pair of shear forces τ^* at the same distance of vt behind the crack tip. The length of cohesive zone in (44), normalized by L_0, is shown in Fig. 2 for time $\tau_c vt/\tau^*=1$ and 10, and Poisson's ratio $\nu=1/3$. Similar to Fig. 1, the normalized cohesive zone length decreases with increasing time, except at the radiation-free velocity $v=\sqrt{2}c_s$. The normalized cohesive zone length remains unchanged at the $v=\sqrt{2}c_s$, because the square-root crack-tip singularity is preserved.

4 Concluding Remarks

We have conducted a fully transient analysis to obtain analytically the fundamental solution for intersonic shear crack propagation. A semi-infinite crack in an infinite solid is subjected to a pair of suddenly applied concentrated shear forces on the crack faces. The crack tip starts to propagate at a velocity between the shear and longitudinal wave speeds. This fundamental solution can provide the general solutions for intersonic crack propagation under arbitrary initial equilibrium fields. There exists a single crack-tip velocity, $v=\sqrt{2}c_s$, at which the crack tip has the conventional square-root singularity, where c_s is the shear wave speed. The crack-tip stress intensity factor at this crack-tip velocity is slightly less than one half of its counterpart for a stationary crack. We have also developed a cohesive zone model to determine the crack-tip energy release rate for an intersonic shear crack.

Acknowledgments

Y. H. is grateful to Prof. G. F. Carrier of Harvard University for an insightful discussion. The work of Y. H. was supported by the ONR through Grant N00014-01-1-0205 (Dr. Y. D. S. Rajapakse, project monitor) and the NSF through Grant CMS9983779. The work of H. G. was supported by the NSF through grant CMS9820988.

Appendix

Cohesive Zone Model. We develop a cohesive zone model in order to obtain the nonvanishing crack-tip energy release rate for this fundamental solution in intersonic crack propagation. A cohesive zone is imposed behind the propagating crack tip to eliminate the crack-tip singularity in (30). The stress field is composed of two parts that lead to positive and negative shear stresses ahead of the crack tip, respectively; (i) the fundamental solution in Section 2; and (ii) the cohesive-zone solution, which has a a constant shear stress traction and vanishing normal stress traction on the crack face:

$$\sigma_{xy}(\xi<0,y=0)=\tau_c H(\xi+L),\quad \sigma_{yy}(\xi<0,y=0)=0, \tag{45}$$

where τ_c is the shear cohesive strength, L the length of the cohesive zone, and H the unit step function.

Instead of conducting a fully transient analysis as in Section 2 to obtain the exact cohesive-zone solution, we use a steady-state solution to approximate this transient one such that the time derivative is related to the spatial derivative with respect to the moving coordinate ξ by $\partial/\partial t=-v\partial/\partial\xi$. The resulted cohesive-zone solution is independent of time, except the cohesive zone length L, which is to be determined by eliminating the time-dependent singular crack-tip field in (30). The energy release rate obtained from this approximate solution will be validated against the exact energy release rate in (42) at the radiation-free crack-tip velocity $v=\sqrt{2}c_s$.

The equation of motion (4) becomes

$$\alpha_l^2\frac{\partial^2\phi}{\partial\xi^2}+\frac{\partial^2\phi}{\partial y^2}=0,\quad -\hat{\alpha}_s^2\frac{\partial^2\psi}{\partial\xi^2}+\frac{\partial^2\psi}{\partial y^2}=0. \tag{46}$$

Its solution in the upper half-plane can be generally written as ([2])

$$\phi=\mathrm{Re}[F(\xi+i\alpha_l y)],\quad \psi=\psi(\xi+\hat{\alpha}_s y), \tag{47}$$

where $F(z)$ is an analytic function of the complex variable z, $i=\sqrt{-1}$, and ψ depends on a single real variable $\xi+\hat{\alpha}_s y$. The symmetry conditions $\sigma_{yy}=u_1=0$ ahead of a mode-II crack tip ($\xi>0$, $y=0$) lead to $\mathrm{Re}[F''(\xi>0)]=0$ and $\psi(\xi>0)=0$. An analytic function can then be defined by

$$\theta(z)=F''(z)\quad \text{if } \mathrm{Im}(z)\geq 0,$$
$$=-\overline{F''}(z)\quad \text{if } \mathrm{Im}(z)\leq 0, \tag{48}$$

where $\overline{F}(z)=\overline{F(\bar{z})}$ is analytic in the lower half plane, and θ is analytic on the entire plane except on the crack face.

The traction boundary condition (45) on the crack face is expressed in terms of F and ψ as

$$2\alpha_l\,\mathrm{Im}\,F''(\xi<0)+(1-\hat{\alpha}_s^2)\psi''(\xi<0)=-\frac{\tau_c}{\mu}H(\xi+L),$$
$$(1-\hat{\alpha}_s^2)\mathrm{Re}\,F''(\xi<0)+2\alpha_s\psi''(\xi<0)=0. \tag{49}$$

Elimination of $\psi''(\xi<0)$ and substitution of θ in (48) into (49) yields

$$[(1-\hat{\alpha}_s^2)^2+4i\alpha_l\hat{\alpha}_s]\theta^+(\xi<0)-[(1-\hat{\alpha}_s^2)^2-4i\alpha_l\hat{\alpha}_s] \times\theta^-(\xi<0)=4\hat{\alpha}_s\frac{\tau_c}{\mu}H(\xi+L)H(-\xi), \tag{50}$$

where $\theta^{\pm}=\theta(y\to 0\pm)$. The above equation constitutes a standard Riemann-Hilbert problem whose solution is given by

$$\theta(z)=\frac{z^{-q}}{2\pi i}\frac{4\hat{\alpha}_s}{\sqrt{(1-\hat{\alpha}_s^2)^4+16\alpha_l^2\hat{\alpha}_s^2}}\frac{\tau_c}{\mu}\int_{-L}^{0}\frac{|\xi'|^q}{\xi'-z}d\xi'. \tag{51}$$

As z approaches the crack face ($\mathrm{Im}(z)\to 0$), (51) has the limits

$$\theta_{\pm}(\xi<0)=\frac{4\hat{\alpha}_s|\xi|^{-q}e^{\mp iq\pi}}{\sqrt{(1-\hat{\alpha}_s^2)^4+16\alpha_l^2\hat{\alpha}_s^2}}\frac{\tau_c}{\mu}\times\left[\frac{1}{2\pi i}PV\int_{-L}^{0}\frac{|\xi'|^q}{\xi'-\xi}d\xi'\pm\frac{1}{2}|\xi|^qH(\xi+L)\right],\tag{52}$$

where PV stands for the Cauchy principal value integral. The stress field in the cohesive-zone solution can be obtained by substituting the potentials into (2).

The cohesive-zone solution in (51) and the fundamental solution in (30) have the same order of stress singularity, r^{-q}, around the crack tip. By requiring their net coefficient of r^{-q} to vanish so as to cancel the stress singularity, we have determined the cohesive zone length as in (44). The net crack-tip energy release rate is obtained from the energy flux into the cohesive zone as ([2])

$$G=\tau_c\delta_c=\frac{\tau_c^2L}{2\pi(1-q)\mu}\frac{16\alpha_l\hat{\alpha}_s^2(1+\hat{\alpha}_s^2)}{(1-\hat{\alpha}_s^2)^4+16\alpha_l^2\hat{\alpha}_s^2},\tag{53}$$

where $\delta_c=u_1(\xi=-L,y\rightarrow 0+)-u_1(\xi=-L,y\rightarrow 0-)$ is the net sliding displacement at the end of the cohesive zone, inclusive of both the fundamental solution and the cohesive zone solution. This, in conjunction with (44), leads to the net energy release rate in (43). At the radiation-free crack-tip velocity $v=\sqrt{2}c_s$, (43) degenerates to the exact crack-tip energy release rate in (42).

References

[1] Rosakis, A. J., Samudrala, O., and Coker, D., 1999, ''Cracks Faster Than the Shear Wave Speed,'' Science, **284**, pp. 1337–1340.

[2] Freund, L. B., 1990, *Dynamic Fracture Mechanics*, Cambridge University Press, Cambridge, UK.

[3] Broberg, K. B., 1999, *Cracks and Fracture*, Academic Press, San Diego, CA.

[4] Washabaugh, P. D., and Knauss, W. G., 1994, ''A Reconciliation of Dynamic Crack Velocity and Rayleigh-Wave Speed in Isotropic Brittle Solids,'' Int. J. Fract., **65**, pp. 97–114.

[5] Freund, L. B., 1979, ''The Mechanics of Dynamic Shear Crack Propagation,'' J. Geophys. Res., **84**, pp. 2199–2209.

[6] Broberg, K. B., 1989, ''The Near-Tip Field at High Crack Velocities,'' Int. J. Fract., **39**, pp. 1–13.

[7] Gao, H., Huang, Y., Gumbsch, P., and Rosakis, A. J., 1999, ''On Radiation-Free Transonic Motion of Cracks and Dislocations,'' J. Mech. Phys. Solids, **47**, pp. 1941–1961.

[8] Archuleta, R. J., 1982, ''Analysis of Near Source Static and Dynamic Measurements From the 1979 Imperial Valley Earthquake,'' Bull. Seismol. Soc. Am., **72**, pp. 1927–1956.

[9] Beroza, G. C., and Spudich, P., 1988, ''Linearized Inversion for Fault Rupture Behavior—Application to the 1984 Morgan-Hill, California, Earthquake,'' J. Geophys. Res., **93**, pp. 6275–6296.

[10] Wald, D. J., and Heaton, T. H., 1994, ''Spatial and Temporal Distribution of Slip for the 1992 Landers, California, Earthquake,'' Bull. Seismol. Soc. Am., **84**, pp. 668–691.

[11] Ellsworth, W. L., and Celebi, M., 1999, ''Near Field Displacement Time Histories of the M 7.4 Kocaeli (Izimit), Turkey, Earthquake of August 17, 1999,'' EOS Trans. Am. Geophys. Union, **80**, No. 46, F648.

[12] Rosakis, A. J., Samudrala, O., and Coker, D., 2000, ''Intersonic Shear Crack Growth Along Weak Planes,'' Mater. Res. Innovations, **3**, pp. 236–243.

[13] Gao, H., 1993, ''Surface Roughening and Branching Instabilities in Dynamic Fracture,'' J. Mech. Phys. Solids, **41**, pp. 457–486.

[14] Abraham, F. F., Brodbeck, D., Rafey, R. A., and Rudge, W. E., 1994, ''Instability Dynamics of Fracture: A Computer Simulation Investigation,'' Phys. Rev. Lett., **73**, pp. 272–275.

[15] Burridge, R., Conn, G., and Freund, L. B., 1979, ''The Stability of a Plane Strain Shear Crack With Finite Cohesive Force Running at Intersonic Speeds,'' J. Geophys. Res., **84**, pp. 2210–2222.

[16] Simonov, I. V., 1983, ''Behavior of Solutions of Dynamic Problems in the Neighborhood of the Edge of a Cut Moving at Transonic Speed in an Elastic Medium,'' Mech. Solids, **18**, pp. 100–106.

[17] Broberg, K. B., 1994, ''Intersonic Bilateral Slip,'' Geophys. J. Int., **119**, pp. 706–714.

[18] Yu, H. H., and Suo, Z., 2000, ''Intersonic Crack Growth on an Interface,'' Proc. R. Soc. London, Ser. A, **A456**, pp. 223–246.

[19] Andrews, D. J., 1976, ''Rupture Velocity of Plane Strain Shear Cracks,'' J. Geophys. Res., **81**, pp. 5679–5687.

[20] Needleman, A., 1999, ''An Analysis of Intersonic Crack Growth Under Shear Loading,'' ASME J. Appl. Mech., **66**, pp. 847–857.

[21] Geubelle, P. H., and Kubair, D., 2001, ''Intersonic Crack Propagation in Homogeneous Media Under Shear Dominated Loading: Numerical Analysis,'' J. Mech. Phys. Solids, **49**, pp. 571–587.

[22] Abraham, F. F., and Gao, H., 2000, ''How Fast Can Cracks Propagate?'' Phys. Rev. Lett., **84**, pp. 3113–3116.

[23] Gao, H., Huang, Y., and Abraham, F. F., 2001, ''Continuum and Atomistic Studies of Intersonic Crack Propagation,'' J. Mech. Phys. Solids (in press).

[24] Freund, L. B., 1972, ''Crack Propagation in an Elastic Solid Subjected to General Loading. I. Constant Rate of Extension,'' J. Mech. Phys. Solids, **20**, pp. 129–140.

[25] Noble, B., 1958, *Methods Based on the Wiener-Hopf Technique*, Pergamon, New York.

S. Candan
Laboratoire de Recherches et Applications en Mécanique Avançée,
Institute Français de Mécanique Avançee,
Aubière F-63175, France

I. Elishakoff[1]
Mem. ASME,
Department of Mechanical Engineering,
Florida Atlantic University,
Boca Raton, FL 33431
e-mail: ielishak@me.fau.edu

Apparently First Closed-Form Solution for Frequencies of Deterministically and/or Stochastically Inhomogeneous Simply Supported Beams

An infinite number of closed-form solutions is reported for a deterministically or stochastically nonhomogeneous beam, for both natural frequencies and reliabilities, for specialized cases. These solutions may prove useful as benchmark solutions. Numerical examples are evaluated. [DOI: 10.1115/1.1355034]

1 Introduction

The aim of this study is to find some closed-form solutions to the dynamic equation of a beam in which the Young's modulus and the density both are polynomial functions, with both the deterministic and stochastic inhomogeneities included. The exact mode shape is searched also as a polynomial function, with attendant closed-form expression for the natural frequencies. The considered case is the beam simply supported at both ends. For the bibliography of investigations on vibration and buckling of nonhomogeneous beams, one may consult with the papers by Eisenberger [1], and Rollot and Elishakoff [2].

The importance of the found solutions lies in the possibility of their use as benchmark solutions against which the efficacy of various approximate methods could be ascertained. Additionally, presently there is a considerable literature on so-called stochastic finite element method (SFEM), that deals with inhomogeneous structures involving random fields. The latter random functions can be represented as mean functions superimposed with deviation functions. Solution of the problem with properly chosen mean functions often constitutes an important part of the analysis (see, e.g., [3,4]). Thus, the closed-form solutions, both in deterministic and stochastic settings possess attractive analytical advantages over approximate solutions where inherent approximations of various natures are needed. For alternative formulations of random eigenvalue problem, the reader may consult with papers by Shinozuka and Astill [5] and Zhu and Wu [6].

2 Formulation of the Problem

The dynamic behavior of a beam is described by the following equation:

$$\frac{d^2}{dx^2}\left[\mathcal{D}(x)\frac{d^2w(x)}{dx^2}\right]-\mathcal{R}(x)\omega^2 w(x)=0 \tag{1}$$

where $\mathcal{D}=EI$ is the flexural stiffness, $E=$Young's modulus, $\rho=$density, $I=$moment of inertia of the cross section, $\mathcal{R}(x)=\rho A$ is the inertial coefficient, $A=$area of the cross section, $w(x)=$displacement, and ω is the natural frequency.

In this study, it is assumed that the cross-sectional area is constant, but both $\mathcal{D}$ and $\mathcal{R}$ are specified as polynomial functions, given by

$$\mathcal{R}(\xi)=\sum_{i=0}^{m} a_i\xi^i \tag{2}$$

$$\mathcal{D}(\xi)=\sum_{i=0}^{n} b_i\xi^i \tag{3}$$

where $\xi=x/L$ is a nondimensional axial coordinate.

We assume that $w(\xi)$ is also polynomial

$$w(\xi)=\sum_{i=0}^{p} w_i\xi^i \tag{4}$$

where w_i are sought coefficients. In these expressions, m, n, and p are, respectively, the degree of the polynomials for $\mathcal{R}(\xi)$, $\mathcal{D}(\xi)$, and $w(\xi)$.

Equation (1) can be rewritten as

$$\frac{d^2}{d\xi^2}\left[\mathcal{D}(\xi)\frac{d^2w(\xi)}{d\xi^2}\right]-kL^4\mathcal{R}(\xi)w(\xi)=0 \tag{5}$$

where

$$k=\omega^2. \tag{6}$$

As the involved functions are assumed to be polynomial ones, the degrees of each polynomial function must be linked, namely

$$n+(p-2)-2=m+p \tag{7}$$

or, simply

$$n-m=4. \tag{8}$$

We observe that Eq. (8) does not depend on the degree p of the displacement $w(\xi)$. We arrive at the seemingly unexpected conclusion that any polynomial function for the displacement may be used in Eq. (5) if it also satisfies the boundary conditions. This fact will be used at a later stage. In view of Eq. (8) the expression for $\mathcal{D}(\xi)$ can be written as follows:

$$\mathcal{D}(\xi)=\sum_{i=0}^{m+4} b_i\xi^i. \tag{9}$$

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 25, 1998; final revision, Sept. 7, 2000. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

3 Boundary Conditions

The case of the simply supported beam is associated with the following boundary conditions:

$$w(0)=0 \tag{10}$$

$$\mathcal{D}(0)w''(0)=0 \tag{11}$$

$$w(1)=0 \tag{12}$$

$$\mathcal{D}(1)w''(1)=0. \tag{13}$$

Solution to Eq. (11) can be found with either $\mathcal{D}(0)=0$ or $w''(0)=0$. However, Young's modulus, which is zero on one point, has no physical sense, thus Eq. (11) is equivalent to $w''(0)=0$. Hence we postulate $b_0>0$. The same reasoning can be applied to Eq. (13). So the displacement has to satisfy the following conditions:

$$w(0)=0 \tag{14}$$

$$w''(0)=0 \tag{15}$$

$$w(1)=0 \tag{16}$$

$$w''(1)=0. \tag{17}$$

Satisfaction of the boundary conditions (14)–(17) requires that the degree of the displacement polynomial must at least be 4. Assuming that $w(\xi)$ is a fourth-order polynomial

$$w(\xi)=w_0+w_1\xi+w_2\xi^2+w_3\xi^3+w_4\xi^4. \tag{18}$$

The satisfaction of the boundary conditions yields

$$w(\xi)=w_1(\xi-2\xi^3+\xi^4). \tag{19}$$

4 Expansion of the Differential Equation

By substituting the different expressions of $\mathcal{D}(\xi)$, $\mathcal{R}(\xi)$, $w(\xi)$ in Eq. (5), we obtain

$$w_1\Bigg[\sum_{i=2}^{m+4} i(i-1)b_i\xi^{i-2}(-12\xi+12\xi^2) + \sum_{i=0}^{m+4} 24b_i\xi^i+2\sum_{i=1}^{m+4} ib_i\xi^{i-1}(-12+24\xi) - kL^4\sum_{i=0}^{m} a_i\xi^i(\xi-2\xi^3+\xi^4)\Bigg]=0. \tag{20}$$

The latter expression can be rewritten as follows:

$$-12\sum_{i=1}^{m+3} i(i+1)b_{i+1}\xi^i+12\sum_{i=2}^{m+4} i(i-1)b_i\xi^i +24\sum_{i=0}^{m+4} b_i\xi^i-24\sum_{i=0}^{m+3}(i+1)b_{i+1}\xi^i+48\sum_{i=1}^{m+4} ib_i\xi^i -kL^4\sum_{i=1}^{m+1} a_{i-1}\xi^i+2kL^4\sum_{i=3}^{m+3} a_{i-3}\xi^i-kL^4\sum_{i=4}^{m+4} a_{i-4}\xi^i=0. \tag{21}$$

The Eq. (21) has to be satisfied for any ξ. This requirement yields the following relations:

$$-24(b_1-b_0)=0, \quad \text{for } i=0 \tag{22}$$

$$-kL^4a_0+72(b_1-b_2)=0, \quad \text{for } i=1 \tag{23}$$

$$-kL^4a_1+144(b_2-b_3)=0, \quad \text{for } i=2 \tag{24}$$

$$L^4(2ka_0-ka_2)+240(b_3-b_4)=0, \quad \text{for } i=3 \tag{25}$$

$$\cdots$$

$$L^4(2ka_{i-3}-ka_{i-4}-ka_{i-1})+12(i+1)(i+2)(b_i-b_{i+1})=0, \quad \text{for } 4\leqslant i\leqslant m+1 \tag{26}$$

$$\cdots$$

$$L^4(2ka_{m-1}-ka_{m-2})+12(m+3)(m+4)(b_{m+2}-b_{m+3})=0, \quad \text{for } i=m+2 \tag{27}$$

$$L^4(2ka_m-ka_{m-1})+12(m+4)(m+5)(b_{m+3}-b_{m+4})=0, \quad \text{for } i=m+3 \tag{28}$$

$$-kL^4a_m+12(m^2+11m+30)b_{m+4}=0, \quad \text{for } i=m+4. \tag{29}$$

Note that the Eqs. (22)–(29) are valid only if $m\geqslant 3$. For cases that satisfy the inequality $m<3$, the reader is referred to the Appendix A. Note also that the Eqs. (22)–(29) have a recursive form.

The sole unknown in Eqs. (22)–(29) is the natural frequency coefficient k, yet we observe that we have $m+5$ equations. We conclude that the parameters b_i and a_i have to satisfy some auxiliary conditions so that Eqs. (22)–(29) are compatible.

5 Compatibility Conditions

A first compatibility condition is given by the Eq. (22), leading to $b_0=b_1$. From the other equations, several expressions for k can be found. Its values determined from Eqs. (22)–(29), respectively, are listed below:

$$k=72(b_1-b_2)/L^4a_0 \tag{30}$$

$$k=144(b_2-b_3)/L^4a_1 \tag{31}$$

$$k=240(a_2-2a_0)^{-1}(b_3-b_4)/L^4 \tag{32}$$

$$\cdots$$

$$k=12(i+1)(i+2)(a_{i-1}+a_{i-4}-2a_{i-3})^{-1}[b_i-b_{i+1}]/L^4, \quad \text{for } 4\leqslant i\leqslant m+1 \tag{33}$$

$$\cdots$$

$$k=12(m+3)(m+4)(a_{m-2}-2a_{m-1})^{-1}[b_{m+2}-b_{m+3}]/L^4 \tag{34}$$

$$k=12(m+4)(m+5)(a_{m-1}-2a_m)^{-1}[b_{m+3}-b_{m+4}]/L^4 \tag{35}$$

$$k=12(m^2+11m+30)b_{m+4}/L^4a_m. \tag{36}$$

To check the compatibility of these expressions, all expressions for k have to be equal to each other. We consider two separate problems: (i) material density coefficients a_i are specified; find coefficients b_i so that closed-form solution holds; (ii) elastic modulus coefficients b_i are specified; find coefficients a_i so that closed-form solution is obtainable.

6 Specified Inertial Coefficient Function

Let us assume that the function $\mathcal{R}(\xi)$, of the inertial coefficient, and hence all a_i $(i=0,2\ldots m)$ are given. Let us observe that if b_{m+4} is specified then the expression given in Eq. (36) is the final formula for the natural frequency coefficient k. Then Eqs. (30)–(36) allow an evaluation of remaining parameters b_i. Note that b_{m+4} and a_m have to have the same sign due to the positivity of k.

From Eq. (35) we get

$$b_{m+3}=\left\{\left[\frac{m^2+11m+30}{(m+4)(m+5)}\right]\left(\frac{a_{m-1}}{a_m}-1\right)+1\right\}b_{m+4}. \tag{37}$$

Equation (34) yields

Table 1

m	$\mathcal{D}(\xi)$	k
0	$b\left(3+3\xi-2\xi^2-2\xi^3+\xi^4\right)$	$360\frac{b}{L^4}$
1	$(b/10)\left(59+59\xi-11\xi^2-46\xi^3-4\xi^4+10\xi^5\right)$	$504\frac{b}{L^4}$
2	$(b/15)\left(135+135\xi-5\xi^2-75\xi^3-33\xi^4-5\xi^5+15\xi^6\right)$	$672\frac{b}{L^4}$
3	$(b/35)\left(434+434\xi+14\xi^2-196\xi^3-70\xi^4-70\xi^5-10\xi^6+35\xi^7\right)$	$864\frac{b}{L^4}$
4	$(b/28)\left(452+452\xi+32\xi^2-178\xi^3-52\xi^4-52\xi^5-52\xi^6-7\xi^7+28\xi^8\right)$	$1080\frac{b}{L^4}$
5	$(b/36)\left(648+648\xi+69\xi^2-224\xi^3-56\xi^4-56\xi^5-56\xi^6-56\xi^7-8\xi^8+36\xi^9\right)$	$1320\frac{b}{L^4}$
6	$(b/15)\left(371+371\xi+41\xi^2-124\xi^3-25\xi^4-25\xi^5-25\xi^6-25\xi^7-25\xi^8-3\xi^9+15\xi^{10}\right)$	$1584\frac{b}{L^4}$
7	$(b/55)\left(1628+1628\xi+198\xi^2-517\xi^3-88\xi^4-88\xi^5-88\xi^6-88\xi^7-88\xi^8-88\xi^9-10\xi^{10}+55\xi^{11}\right)$	$1872\frac{b}{L^4}$
8	$(b/330)\left(5715+57151\xi+1492\xi^2-3513\xi^3-510\xi^4-510\xi^5-510\xi^6-510\xi^7-510\xi^8-510\xi^9-510\xi^{10}-55\xi^{11}+330\xi^{12}\right)$	$2184\frac{b}{L^4}$
9	$(b/26)\left(1053+1053\xi+143\xi^2-312\xi^3-39\xi^4-39\xi^5-39\xi^6-39\xi^7-39\xi^8-39\xi^9-39\xi^{10}-39\xi^{11}-4\xi^{12}+26\xi^{13}\right)$	$2520\frac{b}{L^4}$

$$b_{m+2}=\left[\left(\frac{m+5}{m+3}\right)\frac{a_{m-2}-2a_{m-1}}{a_{m-1}-2a_m}+1\right]b_{m+3}-\left(\frac{m+5}{m+3}\right)\frac{a_{m-2}-2a_{m-1}}{a_{m-1}-2a_m}b_{m+4}. \tag{38}$$

Equation (33) results in

$$b_i=\left[\left(\frac{i+3}{i+1}\right)\frac{a_{i-1}-a_{i-4}-2a_{i-3}}{a_i-a_{i-3}-2a_{i-2}}+1\right]b_{i+1}-\left(\frac{i+3}{i+1}\right)\frac{a_{i-1}-a_{i-4}-2a_{i-3}}{a_i-a_{i-3}-2a_{i-2}}b_{i+2} \tag{39}$$

where i belongs to the set $\{4,5,\ldots,m+1\}$.

From Eq. (32) we obtain

$$b_3=\left[\left(\frac{3}{2}\right)\frac{a_2-2a_0}{a_3-a_0-2a_1}+1\right]b_4-\left(\frac{3}{2}\right)\frac{a_2-2a_0}{a_3-a_0-2a_1}b_5. \tag{40}$$

Eq. (31) leads to

$$b_2=\left[\left(\frac{5}{3}\right)\frac{a_1}{a_2-2a_0}+1\right]b_3-\left(\frac{5}{3}\right)\frac{a_1}{a_2-2a_0}b_4. \tag{41}$$

From Eq. (29) we get

$$b_1=\left(2\frac{a_0}{a_1}+1\right)b_2-2\frac{a_0}{a_1}b_3. \tag{42}$$

And finally, Eq. (22) yields

$$b_0=b_1. \tag{43}$$

Thus, for specified coefficients a_0, a_1, $\ldots$ a_m and b_{m+4}, Eqs. (37)–(43) lead to the set of coefficients in elastic modulus such that the beam possesses mode shape given in Eq. (19). Note that if $a_i=a$, then coefficients b_i do not depend on the parameter a.

To sum up, if

$$\mathcal{R}(\xi)=\sum_{i=0}^{m}a_i\xi^i \quad \mathcal{D}(\xi)=\sum_{i=0}^{m+4}b_i\xi^i \tag{44}$$

where b_i are computed via Eqs. (37)–(43), the fundamental mode shape of a beam is

$$w(\xi)=w_1(\xi-2\xi^3+\xi^4) \tag{45}$$

and the fundamental natural frequency squared reads

$$\omega^2=12(m^2+11m+30)b_{m+4}/a_mL^4. \tag{46}$$

As we have seen, in order to obtain closed-form solution it is sufficient that (1) all a_i coefficients and (2) the coefficient b_{m+4}

Table 2

m	$\mathcal{D}(\xi)$	k
0	$b\left(3+3\xi-2\xi^2-2\xi^3+\xi^4\right)$	$360\frac{b}{L^4}$
1	$(b/10)\left(38+38\xi+3\xi^2-32\xi^3-11\xi^4+10\xi^5\right)$	$252\frac{b}{L^4}$
2	$(b/45)\left(203+203\xi+63\xi^2-77\xi^3-119\xi^4-35\xi^5+45\xi^6\right)$	$224\frac{b}{L^4}$
3	$(b/140)\left(725+725\xi+305\xi^2-115\xi^3-241\xi^4-325\xi^5-85\xi^6+140\xi^7\right)$	$216\frac{b}{L^4}$
4	$(b/140)\left(815+815\xi+395\xi^2-25\xi^3-151\xi^4-235\xi^5-295\xi^6-70\xi^7+140\xi^8\right)$	$216\frac{b}{L^4}$
5	$(b/1512)\left(9751+9751\xi+5131\xi^2+511\xi^3-875\xi^4-1799\xi^5-2549\xi^6-2954\xi^7-644\xi^8+1512\xi^9\right)$	$220\frac{b}{L^4}$
6	$(b/1470)\left(10388+10388\xi+5768\xi^2+1148\xi^3-238\xi^4-1162\xi^5-1822\xi^6-1617\xi^7-2702\xi^8-546\xi^9+1470\xi^{10}\right)$	$\frac{1584}{7}\frac{b}{L^4}$

be specified. Yet, the requirements are not necessary ones. Indeed one can assume that a_i coefficient are given and instead of b_{m+4} any coefficients $b_j (j \neq m+4)$ is specified. If this is the case, then from Eq. (33) one expresses b_{i+1} via b_i and k; substitution into subsequent equations allows us to express $b_{m+2}, b_{m+3}, b_{m+4}$ via b_i; analogously, substitution of b_i into Eqs. (30)–(32) yields sought exact solutions.

In Tables 1 and 2, some sample specified function $\mathcal{D}(x)$ and the attendant fundamental natural frequency coefficients are given. The polynomial functions $\mathcal{R}(x)$ were specified as

$$\mathcal{R}(\xi)=\sum_{i=0}^{m} \xi^i, \quad \mathcal{R}(\xi)=\sum_{i=0}^{m} (i+1)\xi^i, \tag{47}$$

respectively, in Tables 1 and 2.

7 Specified Flexural Stiffness Function

Consider now the case when the flexural stiffness function is specified, implying that all $b_i (i=0, \ldots, m+4)$ are given. The following question arises: Is it possible to determine the material density coefficients $a_i (i=0, \ldots, m)$, such that equations corresponding to Eqs. (22)–(29) are compatible? One immediately observes that there are $(m+5)$ Eqs. (22)–(29), while one has only $m+1$ unknowns, $a_0, a_1, \ldots, a_m$. In actuality, however, one has only m unknowns. In order for the process of determining of coefficient a_i to proceed, one of the a_j coefficients should be specified. The most convenient assumption is to fix either a_0 or a_l or a_m, since in these cases only one equation, respectively, Eq. (30) or Eq. (31) or Eq. (36) will be sufficient to determine the sought expression of the natural frequency coefficient. Let us assume that the coefficient a_0 is given thus, to check the compatibility of Eqs. (22)–(29), four b_i coefficients cannot be chosen arbitrarily.

Note that the natural frequency coefficient k has to be positive thus, the difference b_1-b_2 and the coefficient a_0 have to have the same sign. Moreover, as the coefficient a_0 is positive, the difference b_1-b_2 should be positive. So, for $b_1 > b_2$, one substitutes the value of k determined from Eq. (30) into Eq. (31); this allows to determine the coefficient a_1 so that the frequency coefficient k in Eq. (31) is positive, and so on.

First, Eq. (22) leads to

$$b_0 = b_1. \tag{48}$$

From Eq. (31) we get

$$a_1 = 2a_0 \frac{b_3 - b_2}{b_2 - b_1}. \tag{49}$$

Eq. (32) yields

$$a_2 = \frac{5a_1(b_4 - b_3) + 6a_0(b_3 - b_2)}{3(b_3 - b_2)}. \tag{50}$$

Eq. (33), where $i=4$, results in

$$a_3 = \frac{a_0(4b_4 - 6b_5 + 2b_3) + a_1(4b_4 - 4b_3) + a_2(3b_5 - 3b_4)}{2(b_4 - b_3)}. \tag{51}$$

From Eq. (33), where $4 \leqslant i \leqslant m$, we obtain

$$\begin{aligned} a_i = \frac{1}{(i+1)(b_{i+1} - b_i)} \{ & a_{i-1}(i+3)(b_{i+2} - b_{i+1}) + 2a_{i-2}(i+1) \\ & \times (b_{i+1} - b_i) + a_{i-3}[b_{i+1}(i+5) - 2b_{i+2}(i+3) + b_i(i+1)] \\ & + a_{i-4}(i+3)(b_{i+2} - b_{i+1}) \}. \end{aligned} \tag{52}$$

Then from Eq. (36) and Eq. (30), one can find an expression of b_{m+4}, so that the compatibility of Eqs. (22)–(29) is checked,

$$b_{m+4} = \frac{6a_m(b_1 - b_2)}{a_0(m^2 + 11m + 30)}. \tag{53}$$

From Eq. (35) and Eq. (36), a relation for b_{m+3} can be found,

$$b_{m+3} = \frac{a_0 b_{m+4}(m^2 + 9m + 20) + 6a_{m-1}(b_1 - b_2) + 12a_m(b_2 - b_1)}{a_0(m+4)(m+5)}. \tag{54}$$

Finally, Eq. (34) and Eq. (35) yield to an evaluation of b_{m+2},

$$b_{m+2} = \frac{a_0 b_{m+3}(m+3)(m+4) + 6a_{m-2}(b_1 - b_2) + 12a_{m-1}(b_2 - b_1)}{a_0(m+3)(m+4)}. \tag{55}$$

To sum up, while specifying the elastic modulus function, only $m+1$ coefficients b_i can be chosen arbitrarily; the other remaining four coefficients are connected with the arbitrary ones via Eq. (48) Eqs. (53)–(55).

Thus, if

$$\mathcal{R}(\xi)=\sum_{i=0}^{m} a_i \xi^i \quad \mathcal{D}(\xi)=\sum_{i=0}^{m+4} b_i \xi^i, \tag{56}$$

where a_i and four of b_i coefficients are computed via Eqs. (48)–(55), the fundamental mode shape of a beam is

$$w(\xi) = w_1(\xi - 2\xi^3 + \xi^4). \tag{57}$$

The fundamental natural frequency square reads

$$\omega^2 = 72(b_1 - b_2)/a_0 L^4. \tag{58}$$

The closed-form solutions could be utilized for comparison with approximate techniques. For example, utilization of the single term Boobnov-Galerkin method for the case

$$\mathcal{R}(\xi) = 1 + 2\xi + 3\xi^2 + 4\xi^3 + 5\xi^4 \tag{59}$$

$$\mathcal{D}(\xi) = b_8 L^8 \left(\frac{163}{28} + \frac{163}{28}\xi + \frac{79}{28}\xi^2 - \frac{5}{28}\xi^3 - \frac{151}{140}\xi^4 - \frac{47}{28}\xi^5 - \frac{59}{28}\xi^6 - \frac{1}{2}\xi^7 + \xi^8 \right) \tag{60}$$

yields, with $\sin(\pi\xi)$ taken as a comparison function, the following expression:

$$k = \frac{b_8 L^4}{22{,}050} \left(\frac{6{,}945{,}750 - 409{,}185\pi^4 + 391{,}612\pi^8 - 2{,}716{,}875\pi^2 + 17{,}356\pi^6}{10\pi^4 - 19\pi^2 + 15} \right) \tag{61}$$

or numerically $k=216.29697b_8L^4$, which differs from the exact solution $k=216b_8L^4$ by 0.13 percent.

8 Stochastic Analysis

The preceeding formulation allows one to perform a stochastic analysis to account for the possible randomness in the material density and elastic modulus.

8.1 Probabilistically Specified Inertial Coefficient Function. Assume that the coefficients a_i form a random vector with a joint probability density $f_A(a_1,a_2,\ldots,a_m)$, where $A^T=(A_1,A_2,\ldots,A_m)$ and capital letters denote a random variable A_i whose possible values are denoted by the lowercase notation a_i. As Eq. (36) suggests the natural frequency, squared Ω^2 is also a random variable denoted by

$$\Omega^2=\alpha^2B_{m+4}/A_m$$
$$\alpha^2=12(m^2+11m+30), \tag{62}$$

where the coefficient B_{m+4} constitutes either a deterministic or a random variable. Several cases allow closed-form evaluation of the reliability r, defined in the present circumstances as the probability that the natural frequency squared Ω^2 does not exceed a pre-selected deterministic value ω_0^2,

$$r=\text{Prob}(\Omega^2\leqslant\omega_0^2)=\text{Prob}(\alpha^2B_{m+4}\leqslant\omega_0^2A_m). \tag{63}$$

Let $B_{m+4}\equiv B$ be an exponentially distributed random variable with density

$$f_B(b)=\frac{1}{E(B)}\exp\left[-\frac{b}{E(B)}\right],\quad b\geqslant 0 \tag{64}$$

and zero otherwise, $E(B)$ being the mean value of B. Likewise, the coefficient $A_m\equiv A$ has an exponential density

$$f_A(a)=\frac{1}{E(A)}\exp\left[-\frac{a}{E(A)}\right],\quad a\geqslant 0 \tag{65}$$

and vanishes if $a<0$, with $E(A)$ indicating the mean value of A. Since A is exponentially distributed the random variable ω_0^2A is also exponentially distributed with mean $\omega_0^2E(A)$. Likewise α^2B is an exponential random variable with mean $\alpha^2E(B)$. The reliability is obtained as

$$r=\frac{\omega_0^2E(A)}{\omega_0^2E(A)+\alpha^2E(B)}. \tag{66}$$

It is remarkable that although all coefficients A_j $(j=1,\ldots m)$ are random, the probabilistic characterization of only a single coefficient A_m turns out to be needed, in addition to that of B_{m+4} for the reliability evaluation.

8.2 Specified Flexural Stiffness Function. Assume now that $m+1$ coefficients b_i form a random vector with a joint probability density $f_B(b_1,\ldots.,b_{m+1})$. The remaining four coefficients b_0, b_{m+2}, b_{m+3}, b_{m+4} are related with the above coefficients via Eqs. (48), (53)–(55). Due to the randomness of coefficients $B_1,\ldots.,B_{m+1}$, we conclude, that the natural frequency squared is itself a random variable

$$\Omega^2=\sigma^2(B_1-B_2) \tag{67}$$

where σ^2 is a coefficient

$$\sigma^2=72/L^4a_0. \tag{68}$$

The coefficient a_0 in Eq. (68) can be treated either as a deterministic or as a random variable. For the sake of illustration, a particular case will be considered hereinafter, namely, when a_0 is a deterministic variable. The mean natural frequency squared equals

$$E[\Omega^2]=\sigma^2[E(B_1)-E(B_2)] \tag{69}$$

whereas its variance reads

$$\text{Var}[\Omega^2]=\sigma^4[\text{Var}(B_1)+\text{Var}(B_2)]. \tag{70}$$

As in Eq. (63), reliability is defined as a probability that the natural frequency squared Ω^2 does not exceed a pre-selected value ω_0^2. Once the joint probability density of the coefficients b_1, b_2, a_0 are specified, the reliability function r can be derived directly. The reliability is cast as

$$r=\text{Prob}[\sigma^2(B_1-B_2)<\omega_0^2]. \tag{71}$$

A new random variable is introduced

$$Z=\sigma^2(B_1-B_2)-\omega_0^2. \tag{72}$$

The reliability is re-written as

$$r=\text{Prob}(Z<0)$$
$$=\int_{-\infty}^{\infty}db_2\left[\int_{-\infty}^{b_2+\omega_0^2/\sigma^2}f_{B_2B_1}(b_2,b_1)db_1-\int_{-\infty}^{b_2}f_{B_2B_1}(b_2,b_1)db_1\right]. \tag{73}$$

Let B_1 and B_2 be independent random variables, Eq. (73) becomes

$$r=\int_{-\infty}^{\infty}f_{B_2}(b_2)\left[F_{B_1}\left(b_2+\frac{\omega_0^2}{\sigma^2}\right)-F_{B_1}(b_2)\right]db_2 \tag{74}$$

$$F_{B_1}\left(b_2+\frac{\omega_0^2}{\sigma^2}\right)=\int_{-\infty}^{b_2+\omega_0^2/\sigma^2}f_{B_1}(b_1)db_1$$

$$F_{B_1}(b_2)=\int_{-\infty}^{b_2}f_{B_1}(b_1)db_1. \tag{75}$$

Let B_1 be an uniformly distributed random variable with density

$$f_{B_1}(b_1)=(\beta-\alpha)^{-1},\quad \text{if}\quad b_1\in[\alpha,\beta] \tag{76}$$

and zero otherwise; likewise, the coefficient B_2 has an uniform density

$$f_{B_2}(b_2)=(\delta-\gamma)^{-1},\quad \text{if}\quad b_2\in[\gamma,\delta]. \tag{77}$$

Let us assume that $\alpha>\delta$. Thus, the positivity of Ω^2 is always checked since $B_1>B_2$. We first calculate the expressions in Eq. (75)

$$F_{B_1}\left(b_2+\frac{\omega_0^2}{\sigma^2}\right)=\begin{cases}0, & \text{if}\quad b_2+\dfrac{\omega_0^2}{\sigma^2}\leqslant\alpha\\[2ex] \dfrac{b_2-\left(\alpha-\dfrac{\omega_0^2}{\sigma^2}\right)}{\beta-\alpha}, & \text{if}\quad \alpha<b_2+\dfrac{\omega_0^2}{\sigma^2}<\beta\\[2ex] 1, & \text{if}\quad b_2+\dfrac{\omega_0^2}{\sigma^2}\geqslant\beta\end{cases} \tag{78}$$

$$F_{B_1}(b_2)=0$$

From Eq. (74) it follows that for the reliability evaluation we need to find a region in which both f_{B_2} and F_{B_1} are nonzero. As Eq. (77) suggests, f_{B_2} is nonzero if $\gamma\leqslant b_2\leqslant\delta$. The function F_{B_1} differs from both zero and unity if $\alpha-\omega_0^2/\sigma^2<b_2<\beta-\omega_0^2/\sigma^2$; F_{B_1} equals unity if $b_2>\beta-\omega_0^2/\sigma^2$. Thus, in order for the product $f_{B_2}F_{B_1}$ to be nonzero it is necessary and sufficient that b_2 belongs to the two following intervals:

$$I_1=[\gamma,\delta],\quad I_2=[\alpha-\omega_0^2/\sigma^2,\infty]. \tag{79}$$

It is natural to inquire when these two intervals have no intersection. This takes place when the lower end of interval I_2 exceeds

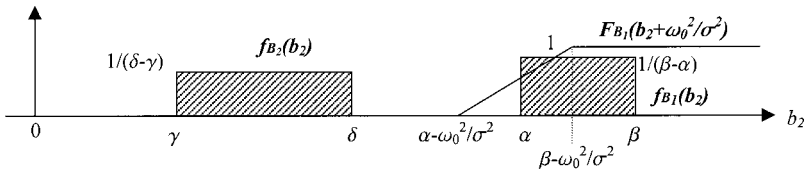


Fig. 1 Probability density functions of variables B_1 and B_2, and the probability distribution function F_{B1} for $\omega_0^2 \leq \sigma^2(\alpha-\delta)$, leading to zero reliability

the upper end of interval I_1, i.e., when $\alpha-\omega_0^2/\sigma^2 \geq \delta$, or in terms of ω_0^2, when $\omega_0^2 \leq \sigma^2(\alpha-\delta)$ (see Fig. 1). In these circumstances the integrand in Eq. (69), and hence the reliability r both vanish identically.

Assume now that the lower end of the interval I_2 belongs to the interval I_1, i.e., $\gamma \leq \alpha-\omega_0^2/\sigma^2 \leq \delta$ and, moreover, $\beta-\omega_0^2/\sigma^2 \geq \delta$ (Fig. 2). This implies that $\sigma^2(\alpha-\delta) < \omega_0^2 \leq \sigma^2(\beta-\delta)$. The sought region for b_2 is the interval $[\alpha-\omega_0^2/\sigma^2, \delta]$. The reliability is obtained as

$$r=\int_{\alpha-\omega_0^2/\sigma^2}^{\delta} \frac{b_2-\left(\alpha-\dfrac{\omega_0^2}{\sigma^2}\right)}{(\delta-\gamma)(\beta-\alpha)}\,db_2=\frac{[\sigma^2(\delta-\alpha)+\omega_0^2]^2}{2\sigma^4(\delta-\gamma)(\beta-\alpha)}. \quad (80)$$

Consider now the following case: $\alpha-\omega_0^2/\sigma^2 \geq \gamma$ and $\gamma \leq \beta-\omega_0^2/\sigma^2 \leq \delta$ (Fig. 3) implying $\sigma^2(\alpha-\gamma) < \omega_0^2 \leq \sigma^2(\beta-\gamma)$. The b_2 region is the interval $[\gamma,\delta]$. But to evaluate the reliability, this region has to be split into the union of two regions $[\gamma,\beta-\omega_0^2/\sigma^2]$, $[\beta-\omega_0^2/\sigma^2,\delta]$ since function $F_{B_1}(b_2+\omega_0^2/\sigma^2)$ takes value of unity at $b_2=\beta-\omega_0^2/\sigma^2$. Hence the reliability reads

$$r=\int_{r}^{\beta-\omega_0^2/\sigma^2} \frac{b_2-\left(\alpha-\dfrac{\omega_0^2}{\sigma^2}\right)}{(\delta-\gamma)(\beta-\alpha)}\,db_2+\int_{\beta-\omega_0^2/\sigma^2}^{\delta}\frac{1}{\delta-\gamma}\,db_2$$
$$=\frac{(\sigma^2\beta-\omega_0^2)+2\sigma^4\delta(\alpha-\beta)+\sigma^4\gamma(\sigma^2\gamma-2\alpha\sigma^2+2\omega_0^2)}{2\sigma^4(\delta-\gamma)(\alpha-\beta)}. \quad (81)$$

We consider now the case $\beta-\omega_0^2/\sigma^2 \leq \gamma$ (Fig. 4), meaning $\sigma^2(\beta-\gamma) < \omega_0^2$. The integration domain is the interval $[\gamma,\delta]$

$$r=\int_{r}^{\delta} 1\,\frac{1}{\delta-\gamma}\,db_2=1. \quad (82)$$

There are two intermediate situations between the ones depicted in the Figs. 2 and 3, depending on the lengths of the intervals $\beta-\alpha$ and $\delta-\gamma$. If the length $\beta-\alpha$ of b_1 interval is smaller than the length $\gamma-\delta$ of b_2 interval, then two quantities, $\alpha-\omega_0^2/\sigma^2$ and $\beta-\omega_0^2/\sigma^2$ belong to the b_2 interval as shown in Fig. 5. The reliability reads

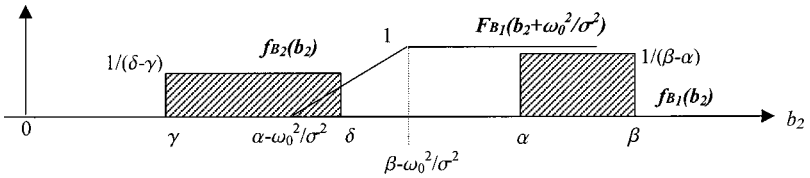


Fig. 2 Probability density functions of variables B_1 and B_2, and the probability distribution function F_{B1} for $\beta-\omega_0^2/\sigma^2 \geq \delta$; reliability is given by expression (80)

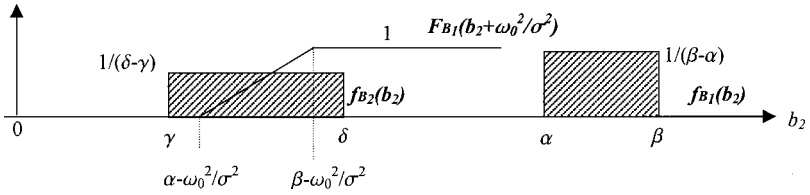


Fig. 3 Probability density functions of variables B_1 and B_2, and the probability distribution function F_{B1} for $\alpha-\omega_0^2/\sigma^2 \geq \gamma$ and $\gamma \leq \beta-\omega_0^2/\sigma^2 \leq \delta$; reliability is given in Eq. (81)

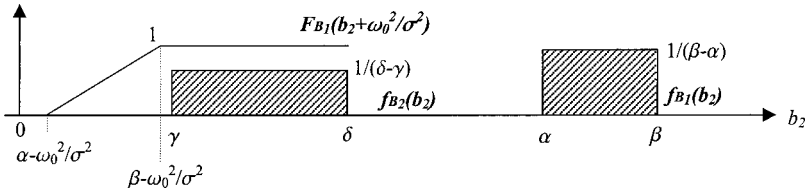


Fig. 4 Probability density functions of variables B_1 and B_2, and the probability distribution function F_{B1} for $\beta-\omega_0^2/\sigma^2 \leq \gamma$; leading to unity reliability

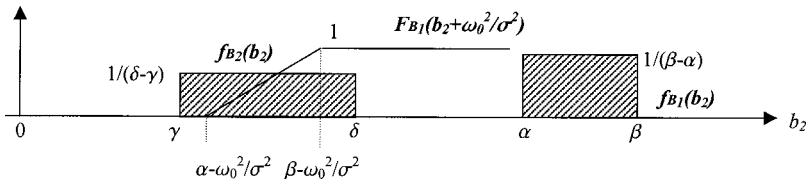


Fig. 5 Case when length $\beta-\alpha$ of B_1 interval is smaller than the length of B_2 interval; reliability is given in Eq. (83)

$$r=\int_{\alpha-\omega_0^2/\sigma^2}^{\beta-\omega_0^2/\sigma^2}\frac{b_2-\left(\alpha-\dfrac{\omega_0^2}{\sigma^2}\right)}{(\delta-\gamma)(\beta-\alpha)}db_2+\int_{\beta-\omega_0^2/\sigma^2}^{\delta}\frac{1}{\delta-\gamma}db_2$$
$$=\frac{\sigma^2(2\delta-\alpha-\beta)+2\omega_0^2}{2\sigma^2(\delta-\gamma)}. \tag{83}$$

If, however, the length $\beta-\alpha$ of b_1 interval is bigger than the length $\gamma-\delta$ of b_2 interval, the sought region is illustrated in Fig. 6, and the reliability is expressed as follows:

$$r=\int_r^{\delta}\frac{b_2-\left(\alpha-\dfrac{\omega_0^2}{\sigma^2}\right)}{(\delta-\gamma)(\beta-\alpha)}db_2=\frac{\delta^2-\gamma^2}{2(\delta-\gamma)(\beta-\alpha)}+\frac{\omega_0^2-\sigma^2\alpha}{\sigma^2(\beta-\alpha)}. \tag{84}$$

The coefficients of variations c_1 and c_2 of random variables B_1 and B_2, respectively,

$$c_1=\frac{\sqrt{\mathrm{Var}(B_1)}}{E(B_1)},\quad c_2=\frac{\sqrt{\mathrm{Var}(B_2)}}{E(B_2)} \tag{85}$$

are chosen to be equal to $c_1=c_2=c$. This implies that for a specified coefficient of variation c, the upper bounds of the interval are related to the lower bounds, as follows:

$$\beta=\alpha\frac{1+\sqrt{3}c}{1-\sqrt{3}c},\quad \delta=\gamma\frac{1+\sqrt{3}c}{1-\sqrt{3}c}. \tag{86}$$

α is fixed at 14, γ is fixed at 1. For the coefficient of variation 0.3, the unity reliability is manifested for $y\geqslant 43.286$. Values associated with transition to the unity reliability for the coefficients of variation of 0.4 and 0.5, are, respectively, $y\geqslant 76.151$ and $y\geqslant 193.994$.

It should be remarked that the transitional values of y from nonunity to unity reliability can be predicted, for the uniformly distributed B_1 and B_2, *without* resort to the reliability calculations. As the natural frequency squared is proportional to the difference of B_1 and B_2 in Eq. (67), the largest value of the natural frequency is obtained when B_1 takes on the value of the upper bound of the interval, β whereas B_2 takes on the value of the lower bound of the interval γ. Thus, if the maximum natural frequency squared $\omega_{\max}^2=\sigma^2(\beta-\gamma)$ is smaller than $y=\omega_0^2$, then the inequality in Eq. (71) is satisfied automatically; hence the reliability is identically unity. Likewise, the minimum natural frequency squared $\omega_{\min}^2=\sigma^2(\alpha-\delta)$. If this value exceeds $y=\omega_0^2$, then obviously for any pairs of values in intervals B_1 and B_2, the inequality (71) will be violated, with resulting vanishing reliability.

On the other hand for exponentially distributed variables, unity reliability is never achieved, as Eq. (66) suggests.

9 Nature of Imposed Restrictions

In this paper, in order to obtain the closed-form solutions for natural frequencies deterministically and/or stochastically inhomogeneous simply supported beam, the flexural stiffness and the inertial coefficient were assumed to be polynomial functions whose powers differ by four. One should stress that the a and b coefficients in Eqs. (2) and (3) cannot be specified independently in order for a closed-form solution to exist. It is quite interesting to comment on the physical meaning of this restriction. Does it signify that a and b coefficients and therefore the inertial coefficient and flexural stiffness must depend on each other? To reply to this question consider a classical case of the closed-form solution, reported for nonlinear stochastic dynamics. Nigam [7] studies the following set of equations:

$$\ddot{Y}_j+\beta_j\dot{Y}_j+\partial V/\partial Y_j=Q_j(t),\quad j=1,2,\ldots,n \tag{87}$$

where Y_j = generalized forces, β_j = damping coefficients, V = potential function, $Q_j(t)$ = generalized forces, n = number of degree-of-freedom. Then the Fokker-Planck equation is constructed that is not reproduced here. Nigam [7] notes

''Assume that

$$\beta_j/\Phi_j=\gamma \text{ for every } j, \tag{88}$$

and define

$$H=\frac{1}{2}\sum_{j=1}^{n}z_{j+n}^2+V(z_1,z_2,\ldots,z_n), \tag{89}$$

then the solution can be expressed as

$$p(z_1,z_2,\ldots,z_{n+1},\ldots,z_{2n})=C\exp[-(\gamma/\pi)H]'' \tag{90}$$

where p is the probability density function and Φ_j is the spectral density of $Q_j(t)$. As is seen the closed-form solution (90) is obtainable when the ratios between the *inner* characteristics—damping coefficients of the system on one hand and the spectral densities Φ_j of *external* excitation on the other, satisfy the condi-

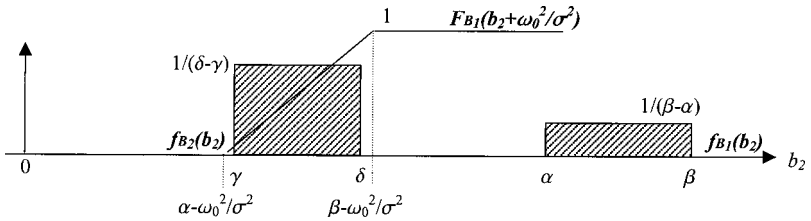


Fig. 6 Case when length $\beta-\alpha$ of B_1 interval is bigger than the length of B_2 interval; reliability is given in Eq. (84)

tion given in Eq. (88). This indicates that when these characteristics have a common numerical parameter then the closed-form solution is derivable. The condition (88) is necessary for the solution to be given by Eq. (90) which was the first derived by Ariaratnam [8] for systems with two-degrees-of-freedom and was extended to multi-degree-of-freedom systems by Caughey [9]. For other closed-form solutions of this and other kinds, the reader may consult for example, with works by Dimentberg [10], Soize [11], Scheurkogel and Elishakoff [12], and others.

One should stress that the importance of the derived closed-form solution is not diminished by the fact that certain conditions must be met. The appearance of conditions is natural too. Indeed, it can be expected that the solution of the *inverse* problem would depend upon the part or entire given data. Thus, if the inhomogeneous beam has a polynomial inertial coefficient with given coefficients, it must be no surprise that the sought flexural stiffness of the beam possesses the pre-selected mode shape, that is directly related to the specified inertial coefficients, in order to derive the closed-form solutions.

The following question arises: Is there any resemblance in the previous literature to the type of thinking adopted in this paper? The connection with the previous work was found via Saint-Venant's *semi-inverse* method. As Timoshenko [13] writes:

> ''In 1853, Saint-Venant presented his epochmaking memoir on torsion to the French Academy. The committee, composed of Cauchy, Poncelet, Piobert, and Lamé, were very impressed by the work and recommended its publication . . . In the introduction Saint-Venant states that the stresses at any point of an elastic body can be readily calculated if the functions representing the components u, v, and w of the displacements are known . . . Saint-Venant then proposes the *semi-inverse method* by which he assumes only some features of the displacements and the forces and determines the remaining features of those quantities so as he ask by all the equations of elasticity. He remarks that an engineer, guided by the approximate solution of the elementary strength of materials, can obtain rigorous solutions of practical importance in this way.''

Indeed, Saint-Venant in 1853 postulated the prior knowledge of the two displacement functions $u=\theta zy$ and $v=\theta zx$, and then determined the function $w=\theta\varphi(x,y)$ where $\varphi(x,y)$ is some function of x and y determined from basic equations. In the present paper we essentially utilized a semi-inverse method: We assumed the knowledge of the mode shape and derived the stiffness from the basic equations. It appears that the clarifying comments in this section further enhance the usefulness of this study.

10 Conclusion

The described class of deterministic and stochastic solutions contains infinite number of closed-form solutions. Indeed, the degree m of polynomial in the expression of the inertial coeficient in Eq. (44) can be chosen arbitrarily. Likewise, the coefficients a_i can be prescribed at will subject to a condition of positivity of both $\mathcal{R}(\xi)$ and $\mathcal{D}(\xi)$.

It should be noted that there is a connection between the present work with the subject of ''inverse problems'' of vibration [14,15]. Indeed, whereas mathematical ''direct problems'' consist of finding solutions to equations with known *input* parameters, mathematical ''inverse problem'' deals with the reconstruction of the parameters of the governing equations when the *output* quantities are known. According to Gladwell [15], ''inverse problems are concerned with the construction of a model of a given type; e.g., a mass-spring system, a string, etc; which has given eigenvalues and/or eigenvectors or eigenfunctions; i.e., given *spectral* data. In general, if some such spectral data is given, there can be no system, a unique system, or many systems, having these properties.'' It is remarkable as the present study demonstrates, that there exist infinite beams, corresponding to $m=0,1,2,\ldots$, that possess the fundamental mode given in Eq. (19).

The natural question arises: *Is it possible to formulate the problem so as to obtain a unique solution*? The reply to this question is affirmative. Indeed, one can pre-select not only the fundamental mode shape, but also the fundamental natural frequency denoted by ω_1. Then the Eq. (46) yields the coefficient b_{m+4} that accomplishes this goal

$$b_{m+4}=\omega_1^2 a_m L^4/12(m^2+11m+30). \tag{91}$$

The polynomial expressions have been used prior to this study in deterministic analyses; yet, to the best of authors' knowledge, this is the first collection of closed-form results in either deterministic or probabilistic setting for the natural frequencies and associated reliabilities of inhomogeneous beams.

It is also notable that whereas in usual finite element method in stochastic setting, only small coefficients of variation can be allowed, the present formulation is not bound to small coefficients of variation. Therefore, the deterministic and probabilistic closed-form solutions that were uncovered in this study can be utilized as benchmark solutions.

Acknowledgment

This study was performed when Mr. Suleyman Candan was a Visiting Research Scholar at the Florida Atlantic University. The authors appreciate the anonymous reviewers whose inquiries led to writing Section 9. The helpful discussion with Dr. Joseph Neuringer and Mr. Denis Meyer is gratefully recorded.

Appendix A

Case 1: $m=0$. In this subcase, the expressions of $\mathcal{R}(\xi)$ and $\mathcal{D}(\xi)$ read (*A*1)

$$\mathcal{R}(\xi)=a_0,\quad \mathcal{D}(\xi)=\sum_{i=0}^{4} b_i\xi^i. \tag{A1}$$

By substituting the latter expressions in Eq. (5), we obtain (*A*2)

$$\begin{aligned}&-12\sum_{i=1}^{3} i(i+1)b_{i+1}\xi^i+12\sum_{i=2}^{4} i(i-1)b_i\xi^i\\&+24\sum_{i=0}^{4} b_i\xi^i-24\sum_{i=0}^{3} i(i+1)b_{i+1}\xi^i\\&+48\sum_{i=1}^{4} ib_i\xi^i-kL^4a_0(\xi-2\xi^3+\xi^4)=0.\end{aligned} \tag{A2}$$

The Eq. (*A*2) has to be satisfied for any ξ. This requirement yields (*A*3)–(*A*7)

$$-24b_1+24b_0=0 \tag{A3}$$

$$-72b_2+72b_1-ka_0=0 \tag{A4}$$

$$-144b_3+144b_2=0 \tag{A5}$$

$$-240b_4+240b_3+2kL^4a_0=0 \tag{A6}$$

$$360b_4-kL^4a_0=0. \tag{A7}$$

To satisfy the compatibility equations, b_i, where $i=\{0,1,2,3\}$, has to be

$$b_1=3b_4 \tag{A8}$$

$$b_0=3b_4 \tag{A9}$$

$$b_3=-2b_4 \tag{A10}$$

$$b_2=-2b_4. \tag{A11}$$

To sum up, if conditions (*A*1) are satisfied, where b_i are given by Eqs. (*A*8)–(*A*11), then the fundamental mode shape is expressed by Eq. (44), where the fundamental natural frequency reads (*A*12)

$$\omega^2 = 360 b_4 / a_0 L^4. \tag{A12}$$

Case 2: $m=1$. In this subcase, the expressions of $E(\xi)$ and $\rho(\xi)$ read, respectively (A13),

$$\mathcal{R}(\xi) = a_0 + a_1 \xi, \quad D(\xi) = \sum_{i=0}^{5} b_i \xi^i. \tag{A13}$$

By substituting the latter expressions in Eq. (5), we obtain

$$\begin{aligned} &-12 \sum_{i=1}^{4} i(i+1) b_{i+1} \xi^i + 12 \sum_{i=2}^{5} i(i-1) b_i \xi^i \\ &+24 \sum_{i=0}^{5} b_i \xi^i - 24 \sum_{i=0}^{4} (i+1) b_{i+1} \xi^i \\ &+48 \sum_{i=1}^{5} i b_i \xi^i - kL^4 a_0 (\xi - 2\xi^3 + \xi^4) \\ &- kL^4 a_1 \xi (\xi - 2\xi^3 + \xi^4) = 0. \end{aligned} \tag{A14}$$

The Eq. (A14) has to be satisfied for any ξ. This requirement yields (A15)–(A20)

$$-24 b_i + 24 b_0 = 0 \tag{A15}$$

$$-72 b_2 + 72 b_1 - kL^4 a_0 = 0 \tag{A16}$$

$$-144 b_3 + 144 b_2 - kL^4 a_1 = 0 \tag{A17}$$

$$-240 b_4 + 240 b_3 + 2kL^4 a_0 = 0 \tag{A18}$$

$$-360 b_5 + 360 b_4 + 2kL^4 a_1 - kL^4 a_0 = 0 \tag{A19}$$

$$-504 b_5 + kL^4 a_0 = 0. \tag{A20}$$

To satisfy the compatibility equations, b_i, $i=\{0,1,2,3,4\}$, has to be

$$b_4 = -\frac{9a_1 - 7a_0}{5a_1} b_5 \tag{A21}$$

$$b_3 = -\frac{9a_1 + 14a_0}{5a_1} b_5 \tag{A22}$$

$$b_2 = \frac{17a_1 + 28a_0}{10a_1} b_5 \tag{A23}$$

$$b_1 = \frac{17a_1 + 42a_0}{10a_1} b_5 \tag{A24}$$

$$b_0 = \frac{17a_1 + 42a_0}{10a_1} b_5 \tag{A25}$$

To sum up, if conditions (A13) are satisfied, where b_i are given by Eqs. (A21)–(A25), then the fundamental mode shape is expressed by Eq. (45), where the fundamental natural frequency reads (A26)

$$\omega^2 = 504 b_5 / a_1 L^4. \tag{A26}$$

Case 3: $m=2$. In this subcase, the expressions of $\mathcal{R}(\xi)$ and $\mathcal{D}(\xi)$ reads (A27)

$$\mathcal{R}(\xi) = a_0 + a_1 \xi + a_2 \xi^2, \quad \mathcal{D}(\xi) = \sum_{i=0}^{6} b_i \xi^i. \tag{A27}$$

By substituting the latter expressions in Eq. (5), we obtain (A28)

$$\begin{aligned} &-12 \sum_{i=1}^{5} i(i+1) b_{i+1} \xi^i + 12 \sum_{i=2}^{6} i(i-1) b_i \xi^i \\ &+24 \sum_{i=0}^{6} b_i \xi^i - 24 \sum_{i=0}^{5} (i+1) b_{i+1} \xi^i + 48 \sum_{i=1}^{6} i b_i \xi^i \\ &- kL^4 (a_0 + a_1 \xi + a_2 \xi^2)(\xi - 2\xi^3 + \xi^4) = 0. \end{aligned} \tag{A28}$$

The Eq. (A14) has to be satisfied for any ξ. This requirement yields (A29)–(A35)

$$-24 b_1 + 24 b_0 = 0 \tag{A29}$$

$$-72 b_2 + 72 b_1 - kL^4 a_0 = 0 \tag{A30}$$

$$-144 b_3 + 144 b_2 - kL^4 a_1 = 0 \tag{A31}$$

$$-240 b_4 + 240 b_3 + 2kL^4 a_0 - kL^4 a_2 = 0 \tag{A32}$$

$$-360 b_5 + 360 b_4 + 2kL^4 a_1 - kL^4 a_0 = 0 \tag{A33}$$

$$-504 b_6 + 504 b_5 + 2kL^4 a_2 - kL^4 a_1 = 0 \tag{A34}$$

$$672 b_6 - kL^4 a_2 = 0. \tag{A35}$$

To satisfy the compatibility equations, b_i, $i=\{0,1,2,3,4,5\}$ has to be

$$b_5 = -\frac{5a_2 - 4a_1}{3a_2} b_6 \tag{A36}$$

$$b_4 = -\frac{25a_2 + 36a_1 - 28a_0}{15a_2} b_6 \tag{A37}$$

$$b_3 = \frac{17a_2 - 36a_1 - 56a_0}{15a_2} b_6 \tag{A38}$$

$$b_2 = \frac{17a_2 + 34a_1 - 56a_0}{15a_2} b_6 \tag{A39}$$

$$b_1 = \frac{17a_2 - 36a_1 + 84a_0}{15a_2} b_6 \tag{A40}$$

$$b_0 = \frac{17a_2 - 36a_1 + 84a_0}{15a_2} b_6. \tag{A41}$$

To sum up, if conditions (A27) are satisfied, where b_i are given by Eqs. (A36)–(A41), then the fundamental mode shape is expressed by Eq. (45), where the fundamental natural frequency reads

$$\omega^2 = 672 b_6 / a_2 L^4. \tag{A42}$$

References

[1] Eisenberger, M., 1997, ''Dynamic Stiffness Vibration Analysis of Non-Uniform Members,'' *International Symposium on Vibrations on Continuous Systems* (A. W. Leissa, organizer), Estes Park, CO, Aug. 11–15, pp. 13–15.

[2] Rollot, O., and Elishakoff, I., 1999, ''A Note on New Closed-Form Solutions for Buckling of a Variable Stiffness Column by Mathematica®,'' J. Sound Vib., **224**, No. 1, pp. 172–182.

[3] Köylüoğlu, H., Cakmak, A. S., and Nielsen, S. A. K., 1994, ''Response of Stochastically Loaded Bernoulli-Euler Beams with Randomly Varying Bending Stiffness,'' *Structural Safety and Reliability*, G. I. Schuëller, M. Shinozuka, and J. T. P. Yao, eds., pp. 267–274.

[4] Elishakoff, I., Ren, Y. J., and Shinozuka, M., 1995, ''Some Exact Solutions for Bending of Beams With Spatially Stochastic Stiffness,'' Int. J. Solids Struct., **32**, pp. 2315–2327 (Corrigendum: **33**, p. 3491, 1996).

[5] Shinozuka, M., and Astill, C. J., 1972, ''Random Eigenvalue Problems in Structural Analysis,'' AIAA J., **10**, pp. 456–462.

[6] Zhu, W. Q., and Wu, W. Q., 1991, ''A Stochastic Finite Element Method for Real Eigenvalue Problem,'' *Stochastic Structural Dynamics*, Vol. 2, I. Elishakoff and Y. K. Lin, eds., Springer-Verlag, Berlin, pp. 337–351.

[7] Nigam, N. C., 1983, *Introduction to Random Vibrations*, MIT Press, Cambridge, MA, pp. 258–259.

[8] Ariaratnam, S. T., 1960, ''Random Vibration of Nonlinear Suspensions,'' J. Mech. Eng. Sci., **2**, No. 3, pp. 195–201.

[9] Caughey, T. K., 1963, ''Derivation and Application of the Fokker-Planck

Equation to Discrete Nonlinear Dynamic Systems Subjected to White Random Excitation,'' J. Acoust. Soc. Am., **35**, No. 11, pp. 1683–1692.
[10] Dimentberg, 1982,
[11] Soize, C., 1988, ''Steady-State Solution of Fokker-Planck Equation in High Dimension,'' Probabilistic Engineering Mechanics, **3**, pp. 196–206.
[12] Scheurkogel, A., and Elishakoff, I., 1988, ''Nonlinear Random Vibration of a Two Degree-of-Freedom Systems,'' *Non-Linear Stochastic Engineering Systems*, F. Ziegler and G. I. Schuëller, eds., Springer, Berlin, pp. 285–299.
[13] Timoshenko, S. P., 1953, *History of Strength of Materials*, McGraw-Hill, New York.
[14] Gladwell, G. M. L., 1986, *Inverse Problems in Vibration*, Martinus Nijhoff Publishers, Dordrecht.
[15] Gladwell, G. M. L., 1996, ''Inverse Problem in Vibration II,'' Appl. Mech. Rev., **49**, No. 10, Part 2, pp. 13–15.

Exact Solutions for Out-of-Plane Vibration of Curved Nonuniform Beams

S. Y. Lee[1]
Professor

J. C. Chao
Graduate Student

Mechanical Engineering Department,
National Cheng Kung University,
Tainan, Taiwan 701, R. O. C.

The governing differential equations for the out-of-plane vibrations of curved nonuniform beams of constant radius are derived. Two physical parameters are introduced to simplify the analysis, and the explicit relations between the torsional displacement, its derivative and the flexural displacement are derived. With these explicit relations, the two coupled governing characteristic differential equations can be decoupled and reduced to one sixth-order ordinary differential equation with variable coefficients in the out-of-plane flexural displacement. It is shown that if the material and geometric properties of the beam are in arbitrary polynomial forms, then the exact solutions for the out-of-plane vibrations of the beam can be obtained. The derived explicit relations can also be used to reduce the difficulty in experimental measurement. Finally, two limiting cases are considered and the influence of taper ratio, center angle, and arc length on the first two natural frequencies of the beams are illustrated. [DOI: 10.1115/1.1346679]

1 Introduction

Curved beam structures have been used in many civil, mechanical, and aerospace engineering applications such as spring design, curved wires in missile-guidance floated gyroscopes, curved girder bridges, brake shoes within drum brakes, tire dynamics, stiffeners in aircraft structures, and turbomachinery blades. It can also be used as a simplified model of a shell structure. Research in this area can be traced back to the 19th century ([1,2]). An interesting review can be found in the papers by Markus and Nanasi [3], Laura and Maurizi [4], Chidamparam and Leissa [5], and Auciello and De Rosa [6].

In general, the out-of-plane and the in-plane vibrations of curved beams are coupled. However, based on the Bernoulli-Euler hypothesis, if the cross section of the curved beam is uniform and doubly symmetric, then the out-of-plane and the in-plane vibrations are independent ([7]).

Out-of-plane vibrations of curved beams have been studied by many investigators ([5]). The associated governing differential equations are two coupled differential equations in terms of the out-of-plane flexural displacement and the torsional displacement. It is known that if the beam is uniform, then the coefficients of the differential equations are constants. After some simple arithmetic operations, the two coupled differential equations can be reduced to a simple sixth-order ordinary differential equation with constant coefficients in the out-of-plane flexural displacement ([2,7]). Hence the problem can be solved by different analytical methods and the exact solutions can be obtained ([1,2,7,8]). However, it is not the case for the nonuniform beams.

Due to the complexity in the coefficients of the governing differential equations, the two coupled differential equations never have been reduced into one sixth-order ordinary differential equation. Exact solutions for the curved nonuniform beam problem are only found in the work by Suzuki, Kosawada, and Takahashi [9], who gave an exact series solution to the beams with the same boundary conditions at both ends. Nevertheless, their method has difficulty in handing the problems with other kinds of boundary conditions. Hence, curved nonuniform beam problems have been studied mainly by approximate methods such as the Rayleigh-Ritz method ([10]), the lumped mass approach ([11]), the transfer matrix method ([12]), the finite element method, and the discrete Green function method ([13]).

In this paper, the governing differential equations for the out-of-plane vibrations of a curved nonuniform beam of constant radius are derived via Hamilton's principle. By introducing two physical parameters, the analysis is simplified and it is found that the torsional displacement and its derivative can be explicitly expressed in terms of the flexural displacement. With these explicit relations, the two coupled governing characteristic differential equations are decoupled and reduced to one sixth-order ordinary differential equation with variable coefficients in the out-of-plane flexural displacement. It can be shown that if the material and geometric properties of the beam are in arbitrary polynomial forms, exact solutions for the out-of-plane vibrations of nonuniform curved beams can be obtained.

It is worth mentioning that by employing the explicit relations, one only has to measure one variable instead of measuring two variables simultaneously in the experimental study of the curved beam. Hence, it greatly reduces the difficulty in experimental measurement.

When the radius of a curved beam becomes infinite, the curved beam reduces to be a straight beam. Consequently, by setting the radius to be infinite, the sixth-order differential equation in terms of the flexural displacement should reduce to a fourth-order differential equation. However, it is not possible to perform this limiting process from the reduced sixth-order ordinary differential equation for the beam ([2,7]) and the limiting study had never been successfully explored before. In this paper, by employing the explicit relation, the limiting study for a curved nonuniform beam is successfully revealed. Finally, the influence of taper ratio, center angle, and arc length on the first two natural frequencies of the beams are studied.

2 Basic Analysis

Consider the out-of-plane and in-plane vibrations of a nonuniform curved beam of radius R, as shown in Fig. 1. Based on the Bernoulli-Euler hypothesis, if the thickness of the beam is small in comparison with the radius of the beam, the displacement fields of the curved beam in cylindrical coordinates are ([5])

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 20, 1999; final revision, May 16, 2000. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$u_r(r,s,z,t)=v(s,t)+z\phi(s,t),$$

$$u_\theta(r,s,z,t)=u(s,t)-r\frac{\partial v(s,t)}{\partial s}-z\frac{\partial w(s,t)}{\partial s}, \tag{1}$$

$$u_z(r,s,z,t)=w(s,t)-r\phi(s,t),$$

where u_r, u_θ, and u_z denote the displacements of the beam in the r, θ, and z directions, respectively. $s=R\theta$ and ϕ is the torsion angle. u, v, and w are the neutral axis displacements of the beam in the r, θ, and z-directions, respectively. Substituting Eq. (1) into the strain-displacement relations in cylindrical coordinates ([14]) yields

$$\epsilon_{\theta\theta}=-r\frac{\partial^2 v}{\partial s^2}+z\left(\frac{\phi}{R}-\frac{\partial^2 w}{\partial s^2}\right)+\left(\frac{\partial u}{\partial s}+\frac{v}{R}\right),$$

$$\gamma_{r\theta}=z\left(\frac{\partial\phi}{\partial s}+\frac{1}{R}\frac{\partial w}{\partial s}\right), \tag{2}$$

$$\gamma_{\theta z}=-r\left(\frac{\partial\phi}{\partial s}+\frac{1}{R}\frac{\partial w}{\partial s}\right),$$

$$\epsilon_{rr}=\epsilon_{zz}=\epsilon_{rz}=0.$$

Here, $\epsilon_{rr}=\epsilon_{zz}=0$ is consistent with the standard assumption in the thin beam theory that the normal stresses are negligible.

The potential energy and the kinetic energy of the beam are

$$\begin{aligned}V&=\frac{1}{2}\int_0^L\int_A(E\epsilon_{\theta\theta}^2+G\gamma_{r\theta}^2+G\gamma_{\theta z}^2)dAds\\&=\int_0^L E\left[-\alpha_1\frac{\partial^2 v}{\partial s^2}\left(\frac{\phi}{R}-\frac{\partial^2 w}{\partial s^2}\right)+\alpha_2\left(\frac{\partial u}{\partial s}+\frac{v}{R}\right)\left(\frac{\phi}{R}-\frac{\partial^2 w}{\partial s^2}\right)\right.\\&\quad\left.-\delta_1\frac{\partial^2 v}{\partial s^2}\left(\frac{\partial u}{\partial s}+\frac{v}{R}\right)\right]ds+\frac{1}{2}\int_0^L\left[EI_z\left(\frac{\partial^2 v}{\partial s^2}\right)^2\right.\\&\quad\left.+EI_r\left(\frac{\phi}{R}-\frac{\partial^2 w}{\partial s^2}\right)^2+EA\cdot\left(\frac{\partial u}{\partial s}+\frac{v}{R}\right)^2\right]ds\\&\quad+\frac{1}{2}\int_0^L GJ\left(\frac{\partial\phi}{\partial s}+\frac{1}{R}\frac{\partial w}{\partial s}\right)^2 ds,\end{aligned} \tag{3}$$

and

$$T=\frac{1}{2}\int_0^L\left\{\rho A\left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial v}{\partial t}\right)^2+\left(\frac{\partial w}{\partial t}\right)^2\right]+\rho J\left(\frac{\partial\phi}{\partial t}\right)^2\right\}ds, \tag{4}$$

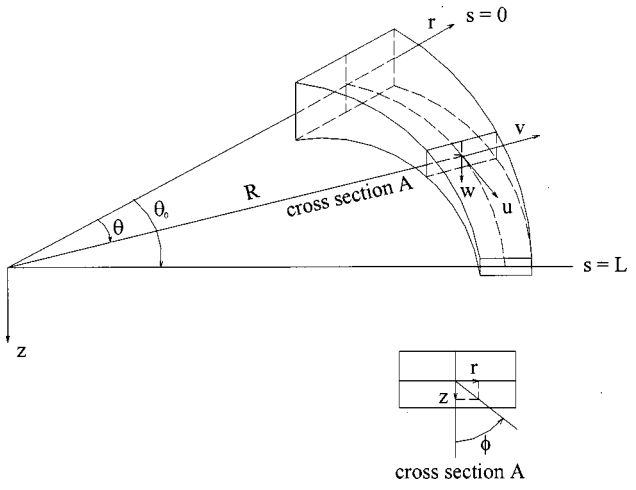


Fig. 1 Geometry and coordinate system of a curved nonuniform beam of constant radius

respectively, where $\alpha_1\sim\alpha_2$ and δ_1 are given as

$$\alpha_1=\int_A rzdA,$$

$$\alpha_2=\int_A zdA, \tag{5}$$

$$\delta_1=\int_A rdA.$$

$E(s)$ is the Young's modulus, $G(s)$ is the shear modulus, $A(s)$ is the cross-sectional area, L is the length of the neutral axis, and $\rho(s)$ is the mass per unit volume of the beam. $I_r(s)$ and $I_z(s)$ denote the area moments of inertia of the beam section about the r and z-axes, respectively. $G(s)J(s)$ is the torsional rigidity. When the beam cross section is circular, $J(s)$ is the polar moment of inertia about the θ-axis ([15]).

Via Hamilton's principle, the governing differential equations and the boundary conditions of the system can be derived. It can be shown that if the cross section of the beam is doubly symmetric about the r and z-axes, then the terms $\alpha_1\sim\alpha_2$ will vanish. As a result, the out-of-plane and in-plane vibrations of curved nonuniform beams are independent. This conclusion is consistent with that for the curved uniform beam theory ([7]).

The governing differential equations for the out-of-plane vibrations are two coupled differential equations expressed in the out-of-plane flexural and the torsional displacements

$$\begin{aligned}&\left[EI_r\left(\frac{1}{R}\phi-w''\right)\right]''+\left[GJ\left(\frac{1}{R}\phi'+\frac{1}{R^2}w'\right)\right]'=\rho A\ddot{w},\\&\left[GJ\left(\phi+\frac{1}{R}w'\right)\right]'-EI_r\left(\frac{1}{R^2}\phi-\frac{1}{R}w''\right)=\rho J\ddot{\phi},\end{aligned} \tag{6}$$

where the primes denote differentiation with respect to the s variable. The governing differential equations for the in-plane vibrations are

$$\begin{aligned}&\left[EI_z\left(\frac{1}{R^2}u'-\frac{1}{R}v''\right)\right]'+\left[EA\left(u'+\frac{1}{R}v\right)\right]'=\rho A\ddot{u},\\&\left[EI_z\left(\frac{1}{R}u'-v''\right)\right]''-EA\left(\frac{1}{R}u'+\frac{1}{R^2}v\right)=\rho A\ddot{v}.\end{aligned} \tag{7}$$

They are two coupled differential equations in terms of the in-plane flexural and the longitudinal displacements.

3 Out-of-Plane Vibrations

For time-harmonic out-of-plane vibrations of curve beams with angular frequency Ω, one assumes

$$\begin{aligned}\phi(s,t)&=\Phi(s)e^{i\Omega t},\\w(s,t)&=W(s)e^{i\Omega t}.\end{aligned} \tag{8}$$

The coupled governing characteristic differential equations for the out-of-plane vibrations of a curved nonuniform beam are

$$\left[EI_r\left(\frac{1}{R}\Phi-W''\right)\right]''+\left[GJ\left(\frac{1}{R}\Phi'+\frac{1}{R^2}W'\right)\right]'+\rho A\Omega^2 W=0, \tag{9}$$

$$\left[GJ\left(\Phi'+\frac{1}{R}W'\right)\right]'-EI_r\left(\frac{1}{R^2}\Phi-\frac{1}{R}W''\right)+\rho J\Omega^2\Phi=0. \tag{10}$$

The associated boundary conditions are at $s=0$ and L:

$$\left[EI_r\left(\frac{1}{R}\Phi-W''\right)\right]'+GJ\left(\frac{1}{R}\Phi'+\frac{1}{R^2}W'\right)=0\ \text{ or }\ W=0, \tag{11}$$

$$EI_r\left(\frac{1}{R}\Phi - W''\right) = 0 \quad \text{or} \quad W' = 0, \tag{12}$$

$$GJ\left(\Phi' + \frac{1}{R}W'\right) = 0 \quad \text{or} \quad \Phi = 0. \tag{13}$$

If the beam is clamped at the boundary, then the boundary conditions are $W=0$, $W'=0$ and $\Phi=0$. If the beam is free at the boundary, then the other identities in Eqs. (11)–(13) are specified.

3.1 Governing Differential Equation in Terms of the Flexural Displacement Parameter *W*

3.1.1 Nonuniform Curved Beams. To simplify the analysis, one defines two physical parameters

$$T_w = \left(GJ\frac{W'}{R}\right)' + EI_r\frac{W''}{R}, \tag{14}$$

$$F_{zw} = (EI_r W'')'' - \left(GJ\frac{W'}{R^2}\right)' - \rho A\Omega^2 W, \tag{15}$$

where T_w and F_{zw} are the torque per unit arc length and the force per unit arc length in the *z*-direction, caused by the flexural deflection parameter *W*, respectively. In terms of T_w and F_{zw}, Eqs. (9)–(10) can be rewritten as

$$\frac{1}{R}\Phi = \frac{a_1}{R}\Phi' + \frac{F_{zw}}{g_1 p_w} + \frac{T_w}{g_1 RGJ}, \tag{16}$$

$$\frac{1}{R}\Phi'' = \frac{a_2}{R}\Phi' + \frac{a_3}{g_1 p_w}F_{zw} - \frac{(EI_r)''}{g_1 p_w RGJ}T_w, \tag{17}$$

where p_w, g_1, a_1, a_2, and a_3 are referred in the Appendix. Substituting Eq. (17) into the equation resulted from differentiating Eq. (16), the relation between Φ'' and *W* can be obtained. Substituting the result into Eqs. (16)–(17), one can explicitly express Φ and Φ' in terms of the flexural displacement *W* parameter

$$\frac{1}{R}\Phi' = \frac{1}{a_4}q_w + \frac{a_1}{p_w g_1(1-a_1')}\left[a_3 F_{zw} - \frac{(EI_r)''}{RGJ}T_w\right], \tag{18}$$

$$\frac{1}{R}\Phi = \frac{a_1}{a_4}q_w + \frac{1}{p_w g_1(1-a_1')}\left\{(1-a_1'+a_1^2 a_3)F_{zw} + \frac{1}{RGJ}[p_w(1-a_1') - a_1^2(EI_r)'']T_w\right\} \tag{19}$$

where

$$q_w = \left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' \tag{20}$$

and a_4 is given in the Appendix. Substituting Eqs. (18)–(19) into Eqs. (9)–(10), one obtains two uncoupled sixth-order ordinary differential equations. After making some arithmetic operations with the two sixth-order equations and Eqs. (18)–(19), it can be found that they are dependent. Thus, the uncoupled governing equation in the lateral displacement parameter *W* is

$$\left\{GJ\left\{\frac{1}{a_4}\left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' + \frac{a_1}{p_w g_1(1-a_1')}\left[a_3 F_{zw} - \frac{(EI_r)''}{RGJ}T_w\right] + \frac{1}{R^2}W'\right\}\right\}'$$
$$- EI_r\left\{\frac{1}{R^2}\left\{\frac{a_1}{a_4}\left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' + \frac{1}{p_w g_1(1-a_1')}\right.\right.$$
$$\left.\left.\times\left\{(1-a_1'+a_1^2 a_3)F_{zw} + \frac{1}{RGJ}[p_w(1-a_1') - a_1^2(EI_r)'']T_w\right\}\right\}'$$
$$- \frac{1}{R^2}W''\right\} + \rho J\Omega^2\left\{\frac{a_1}{a_4}\left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' + \frac{1}{p_w g_1(1-a_1')}\right.$$
$$\left.\times\left\{(1-a_1'+a_1^2 a_3)F_{zw} + \frac{1}{RGJ}[p_w(1-a_1') - a_1^2(EI_r)'']T_w\right\}\right\} = 0. \tag{21}$$

The boundary conditions (11)–(13) in terms of *W* are

$$\left\{EI_r\left\{\frac{a_1}{a_4}\left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' + \frac{1}{p_w g_1(1-a_1')}\left\{(1-a_1'+a_1^2 a_3)F_{zw} + \frac{1}{RGJ}\cdot[p_w(1-a_1') - a_1^2(EI_r)'']T_w\right\} - W''\right\}\right\}'$$
$$+ GJ\left\{\frac{1}{a_4}\left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' + \frac{a_1}{p_w g_1(1-a_1')}\left[a_3 F_{zw} - \frac{(EI_r)''}{RGJ}T_w\right] + \frac{1}{R^2}W'\right\} = 0 \tag{22}$$

or

$$W = 0, \tag{23}$$

$$\frac{a_1}{a_4}\left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' + \frac{1}{p_w g_1(1-a_1')}\times\left\{(1-a_1'+a_1^2 a_3)F_{zw} + \frac{1}{RGJ}\cdot[p_w(1-a_1') - a_1^2(EI_r)'']T_w\right\} - W'' = 0 \tag{24}$$

or

$$W' = 0, \tag{25}$$

and

$$\frac{1}{a_4}\left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' + \frac{a_1}{p_w g_1(1-a_1')}\left[a_3 F_{zw} - \frac{(EI_r)''}{RGJ}T_w\right] + \frac{1}{R^2}W' = 0 \tag{26}$$

or

$$\frac{a_1}{a_4}\left[\frac{1}{g_1}\left(\frac{T_w}{RGJ} + \frac{F_{zw}}{p_w}\right)\right]' + \frac{1}{p_w g_1(1-a_1')}\times\left\{(1-a_1'+a_1^2 a_3)F_{zw} + \frac{1}{RGJ}\cdot[p_w(1-a_1') - a_1^2(EI_r)'']T_w\right\} = 0. \tag{27}$$

3.1.2 Curved Uniform Beams. For curved uniform beams, $a_1 = a_2 = 0$, $a_4 = 0$ and

$$a_3 = g_1 = \frac{EI_r}{R^2 GJ} - \frac{\rho\Omega^2}{G}, \tag{28}$$

T_w and F_{zw} now are rewritten as

$$T_w = GJ\frac{W''}{R} + EI_r\frac{W''}{R}, \tag{29}$$

$$F_{zw} = EI_r W'''' - GJ\frac{W''}{R^2} - \rho A\Omega^2 W. \tag{30}$$

Consequently, the governing Eq. (21) reduces to

$$W^{(6)}+\left(\frac{2}{R^2}+\frac{\rho\Omega^2}{G}\right)W''''+\left[\frac{1}{R^4}-\frac{\rho\Omega^2}{E}\left(\frac{A}{I_r}+\frac{1}{R^2}\cdot\frac{J}{I_r}\right)\right]W''$$
$$+\frac{\rho A\Omega^2}{GJ}\left(\frac{1}{R^2}-\frac{J}{I_r}\cdot\frac{\rho\Omega^2}{E}\right)W=0, \tag{31}$$

where $W^{(6)}$ indicates the sixth derivative of W with respect to s.

If the extension effect is not considered, i.e., E is infinite but EI_r is finite, then in terms of $s^*=s/R\theta_0=s/L$ and $W^*=W/R\theta_0=W/L$, the governing differential Eq. (31) is reduced to

$$\frac{d^6W^*}{ds^{*6}}+\left(2+\frac{\rho R^2\Omega^2}{G}\right)\frac{d^4W^*}{ds^{*4}}+\left[1-\frac{\rho R^4\Omega^2}{EI_r}\left(A+\frac{J}{R^2}\right)\right]\frac{d^2W^*}{ds^{*2}}$$
$$+\frac{\rho A\Omega^2R^4}{GJ}\left(1-\frac{\rho JR^2\Omega^2}{EI_r}\right)W^*=0 \tag{32}$$

where θ_0 is the center angle and is the total arc angle of the curve beam. This equation is exactly the same as the one given by Lee [16].

3.1.3 Straight Nonuniform Beams. When the radius of a curved beam becomes infinite, the curved beam becomes straight. Consequently, by setting the radius being infinite, the reduced six-order differential Eqs. (21), (31)–(32) should reduce to fourth-order differential equations. However, it can be found that it is not possible to take such a direct limiting study. In this section, by employing the explicit relation (19), the limiting study is revealed in the following:

By letting R being infinite in Eqs. (14)–(15), T_w and F_{zw} yield to

$$T_w=0, \tag{33}$$
$$F_{zw}=(EI_rW'')''-\rho A\Omega^2W. \tag{34}$$

Equation (21) becomes

$$\left\{GJ\left[\frac{1}{a_4}\left(\frac{F_{zw}}{g_1p_w}\right)'+\frac{a_1a_3}{p_wg_1(1-a_1')}F_{zw}\right]\right\}'$$
$$+\rho J\Omega^2\left[\frac{a_1}{a_4}\left(\frac{F_{zw}}{g_1p_w}\right)'+\frac{1-a_1'+a_1^2a_3}{p_wg_1(1-a_1')}F_{zw}\right]=0. \tag{35}$$

Equation (19) now is

$$\frac{1}{a_4}\left(\frac{F_{zw}}{g_1p_w}\right)'=\frac{-a_1a_3}{p_wg_1(1-a_1')}F_{zw}. \tag{36}$$

Substituting Eq. (36) with Eq. (35), the sixth-order governing differential Eq. (21) can be reduced to

$$F_{zw}=(EI_rW'')''-\rho A\Omega^2W=0. \tag{37}$$

By setting R being infinite, the boundary conditions Eqs. (11)–(12) become

$$(EI_rW'')'=0 \quad \text{or} \quad W=0, \tag{38}$$
$$W''=0 \quad \text{or} \quad W'=0. \tag{39}$$

Equation (37) and Eqs. (38)–(39) are the governing differential equation and the associated boundary conditions for the flexural vibration of a straight nonuniform beam ([17]).

It should be mentioned that for a straight beam, the flexural and the torsional vibrations are independent. Upon letting R being infinite, Eq. (37) can also be reduced from Eq. (9) directly.

3.2 Exact Fundamental Solutions and Frequency Equation. The uncoupled governing characteristic differential equation, in terms of W, for the out-of-plane vibration of curved nonuniform beams can be expressed as a sixth-order differential equation with variable coefficients in the form of

$$e_0(x)\frac{d^6Y(x)}{dx^6}+e_1(x)\frac{d^5Y(x)}{dx^5}+\cdots+e_5(x)\frac{dY(x)}{dx}$$
$$+e_6(x)Y(x)=0.$$
$$x\in(0,1). \tag{40}$$

Assume $e_0(x)$, $e_6(x)$ do not vanish on the closed interval [0, 1] and are in the polynomial forms, i.e.,

$$e_p(x)=\sum_{j=0}^{m_p}a_{p,j}(x-x_o)^j, \quad p=0\sim6 \tag{41}$$

where x_o is a constant to be selected for the fastest convergence and $0<x_o<1$. m_p, $p=0\sim6$ are integers representing the number of terms in the series. Then one can assume the six fundamental solutions of the differential equation to be in the form of ([18])

$$Y_i(x)=\frac{1}{i!}(x-x_o)^i+\sum_{q=n}^{\infty}A_{q,i}(x-x_o)^q, \quad i=0\sim5, \quad n=6. \tag{42}$$

Substituting Eqs. (41)–(42) into Eq. (40), collecting the coefficients of like powers, the following recurrence formula can be obtained:

$$A_{q,i}=-\frac{(q-n)!}{(q!)a_{0,0}}\left\{\sum_{k=1}^{q-n}\frac{(q-k)!}{(q-k-n)!}a_{0,k}A_{q-k,i}\right.$$
$$+\sum_{m=1}^{n}\left[\frac{1}{(i-n+m)!}a_{m,q-m-i}\right.$$
$$\left.\left.+\sum_{k=m}^{q-n}\frac{(q-k)!}{(q-k-n+m)!}a_{m,k-m}A_{q-k,i}\right]\right\}, \quad q\geqslant n, \quad n=6. \tag{43}$$

With this recurrence formula, one can generate the six exact fundamental solutions of the governing characteristic differential equation. After substituting these fundamental solutions into the associated boundary conditions, the frequency equation and natural frequencies of the beams are obtained, consequently.

4 Numerical Results

To illustrate the previous analysis, the out-of-plane vibrations of a curved nonuniform beam of circular cross section are studied. The following dimensionless parameters will be used in the following numerical analysis:

$$R_\theta^*=\frac{R}{\gamma_\theta(0)},$$
$$m(s^*)=\frac{\rho(s)A(s)}{\rho(0)A(0)},$$
$$b_\theta(s^*)=\frac{G(s)J(s)}{G(0)J(0)},$$
$$b_{r\theta}(s^*)=\frac{E(s)I_r(s)}{G(0)J(0)}, \tag{44}$$
$$L_\theta=R_\theta^*\theta_0,$$
$$\Lambda_\theta=\left[\Omega\Big/\sqrt{\frac{G(0)}{\rho(0)L^2}}\right]^2,$$

where γ_θ is the radius of gyration about the θ-axis.

In Table 1, the first out-of-plane natural frequencies of clamped-clamped curved beams are compared with those in the existing literature. It shows the results are very consistent.

Table 1 The first out-of-plane natural frequencies of curved uniform beams with clamped-clamped ends

θ_0 (degree)	R	$b_{r\theta}(0)=0.615$		$b_{r\theta}(0)=1.0$		$b_{r\theta}(0)=2.0$		
		#	&	#	&	#	&	*
0	∞	⋯	22.373	⋯	22.373	⋯	22.373	22.373
90	50	⋯	20.840	⋯	20.694	⋯	20.363	⋯
180	50	18.379	18.361	18.132	18.128	17.564	17.560	⋯
270	50	17.767	17.765	16.877	16.875	15.343	15.342	⋯

*: exact frequency parameter p in Rao [19].
#: $\alpha^2\sqrt{k\Lambda_n}$ in Volterra and Morell [20].
&: $c\sqrt{\Lambda_\theta}$ shown in the present paper, where $c=L_\theta/\sqrt{b_{r\theta}(0)}$

In the following, the natural frequencies of linearly tapered curved beams with clamped-free ends are studied. The material and geometric properties of the beams with taper ratio η are $m(s^*)=(1-\eta s^*)^2$, $b_\theta(s^*)=(1-\eta s^*)^2$ and $b_{r\theta}(s^*)=b_{r\theta}(0)\times(1-\eta s^*)^4$, respectively.

In Figs. 2 and 3, the influence of taper ratio and curvature (center angle) on the first and the second dimensionless natural frequencies $\sqrt{\Lambda_\theta}$ of the beams with constant dimensionless arc length $L_\theta=30$ and various center angle θ_0 is shown. For the beam with center angle θ_0 being zero, the radius of the beam is infinite. It represents a straight beam. The ones in the figures with cross mark are the dimensionless natural frequencies of a straight uniform beam and are consistent with the exact solutions given by Meirovitch [17]. From Fig. 2, it can be observed that the first natural frequencies of the beams with the same center angle increase as the taper ratio is increased. Those of the beams with the same taper ratio increase as the center angle is increased. Figure 3 shows that the influences of taper ratio and curvature on the second natural frequencies are quite different from that on the first natural frequencies. The second natural frequencies of the beams with the same center angle decrease as the taper ratio is increased and those of the beams with the same taper ratio decreased as the center angle is increased.

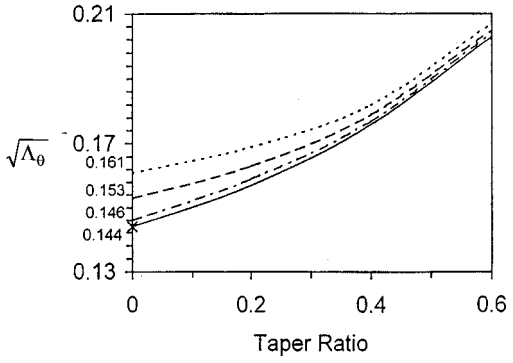


Fig. 2 The influence of taper ratio on the first dimensionless natural frequencies $\sqrt{\Lambda_\theta}$ of clamped-free beams with various center angle θ_0 ($L_\theta=30$, $b_{r\theta}(0)=1.5$; ——: $\theta_0=0$ deg; —·—·—·—: $\theta_0=20$ deg; — — —: $\theta_0=40$ deg; · · · ·: $\theta_0=60$ deg)

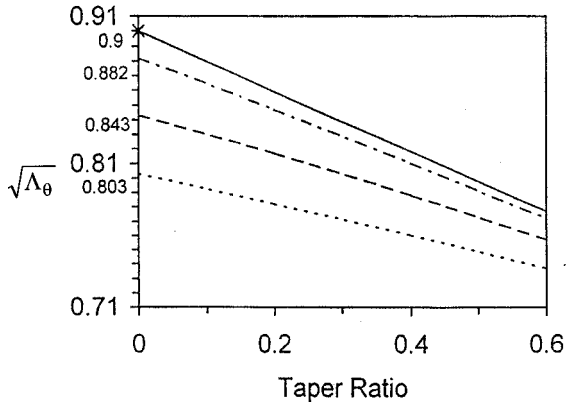


Fig. 3 The influence of taper ratio on the second dimensionless natural frequencies $\sqrt{\Lambda_\theta}$ of clamped-free beams with various center angle θ_0 ($L_\theta=30$, $b_{r\theta}(0)=1.5$; ——: $\theta_0=0$ deg; —·—·—·—: $\theta_0=20$ deg; — — —: $\theta_0=40$ deg; · · · ·: $\theta_0=60$ deg)

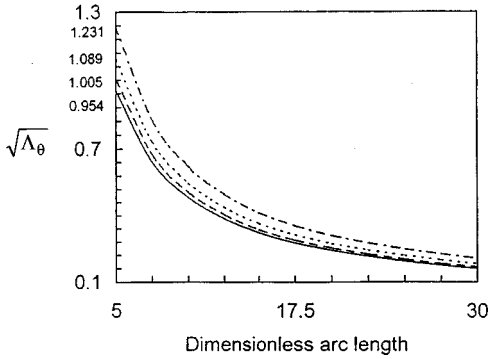


Fig. 4 The influence of dimensionless arc length on the first dimensionless natural frequencies $\sqrt{\Lambda_\theta}$ of clamped-free beams with various taper ratio ($b_{r\theta}(0)=1.5$, $\theta_0=60$ deg; ——: $\eta=0$; — — —: $\eta=0.2$; · · · ·: $\eta=0.4$; —·—·—·—: $\eta=0.6$)

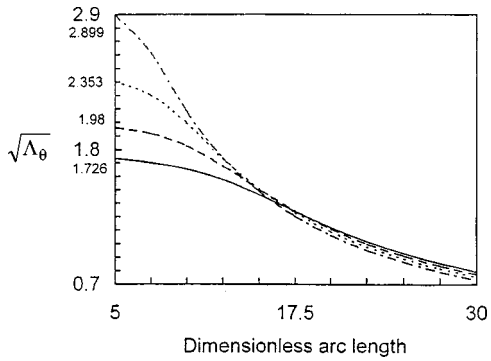


Fig. 5 The influence of dimensionless arc length on the second dimensionless natural frequencies of clamped-free beams with various taper ratio ($b_{r\theta}(0)=1.5$, $\theta_0=60$ deg; ——: $\eta=0$; — — —: $\eta=0.2$; · · · ·: $\eta=0.4$; —·—·—·—: $\eta=0.6$)

Figures 4 and 5 show the influence of dimensionless arc length L_θ and taper ratio on the first two natural frequencies of the beams with constant center angle ($\theta_0=60$ deg). From Fig. 4 it can be found that the first natural frequencies of the beams with the same taper ratio decrease as the dimensionless arc length is increased and those with the same dimensionless arc length increase as the taper ratio is increased. In Fig. 5 it shows that the second natural frequencies of the beams with the same taper ratio also decrease as the dimensionless arc length is increased. However, the second natural frequencies of the beams with the same dimensionless arc length will increase as the taper ratio is increased only when the dimensionless arc length is approximately less than 13. Those of the beams will decrease as the taper ratio is increased when the dimensionless arc length is approximately greater than 17.5.

5 Conclusions

In this paper, two physical parameters are introduced to simplify the analysis of out-of-plane vibrations of curved beams. The explicit relations between the torsional displacement, its derivative, and the flexural displacement are derived. With these explicit

relations, the two-coupled governing characteristic differential equations are reduced to one sixth-order ordinary differential equation with variable coefficients in the out-of-plane flexural displacement. It is shown that if the coefficients of the reduced differential equation are in arbitrary polynomial forms, then the exact solutions for the out-of-plane vibrations of nonuniform curved beams can be obtained. The method can be easily applied to the problems with general elastically restrained boundary conditions. By employing the explicit relations, one only has to measure one variable instead of measuring two variables simultaneously in the experimental study of curved beams. Hence, the difficulty in experimental measurement can be greatly reduced. In addition, a successful limiting study from the curved beam theory to the straight beam theory is revealed. Numerical analysis shows the influence of taper ratio, center angle, and dimensionless arc length on the first two natural frequencies are quite different.

Acknowledgment

This research work was sponsored by the National Science Council of Taiwan, R.O.C. under Grant NSC85-2212-E006-042 and is gratefully acknowledged.

Appendix

The forms of p_w and $a_1 \sim a_8$:

$$p_w = EI_r + GJ,$$

$$g_1 = \frac{(EI_r)''}{p_w} + \frac{EI_r}{R^2 GJ} - \frac{\rho\Omega^2}{G},$$

$$a_1 = -\frac{GJ}{p_w g_1 EI_r}\left(\frac{E^2 I_r^2}{GJ}\right)',$$

$$a_2 = \frac{1}{p_w g_1}\left(\frac{E^2 I_r^2}{GJ}\right)'\left(\frac{\rho J\Omega^2}{EI_r} - \frac{1}{R^2}\right) - \frac{(GJ)'}{GJ},$$

$$a_3 = \frac{EI_r}{R^2 GJ} - \frac{\rho\Omega^2}{G},$$

$$a_4 = 1 - \left[\frac{(GJ)'}{GJ} - \frac{(p_w + EI_r)'}{p_w}\right]\cdot\left[\left(\frac{1}{g_1}\right)' + \frac{1}{p_w g_1^2}\left(\frac{E^2 I_r^2}{GJ}\right)\times\left(\frac{\rho J\Omega^2}{EI_r} - \frac{1}{R^2}\right)\right] - \frac{(GJ)'}{g_1 GJ} - \frac{1}{g_1}\left[\frac{(GJ)'}{GJ} - \frac{(p_w + EI_r)'}{p_w}\right]',$$

$$a_5 = \left[\frac{(GJ)'}{p_w}\right]^2\cdot\left[\frac{(GJ)'}{p_w}\right]',$$

$$a_6 = -\frac{(GJ)'}{p_w^3}\left\{\left\{1 + \left[\frac{(GJ)'}{p_w}\right]'\right\}(GJ)'(GJ + p_w)' - z_w(GJ + p_w)''\right\}\frac{(GJ)'''}{p_w},$$

$$a_7 = \left\{-\left[\frac{EI_r(GJ)'}{p_w}\right]'' - \frac{(GJ)'}{R^2}\left(1 - \frac{GJ}{p_w}\right) + \frac{(GJ)'}{p_w}\cdot\left\{2\left[\frac{EI_r(GJ)'}{p_w}\right]' + EI_r\cdot\left\{\left[\frac{(GJ)'}{p_w}\right]' - \frac{(GJ)'}{p_w}\right\}\right\}\right\}\cdot\frac{1}{\rho A\Omega^2},$$

$$a_8 = \frac{1}{\rho A\Omega^2 p_w}\left\{(EI_r)'' - \frac{GJ}{R^2} + EI_r\left\{\left[\frac{(GJ)'}{p_w}\right]^2 - \left[\frac{(GJ)'}{p_w}\right]'\right\} - \frac{2(EI_r)'(GJ)'}{p_w}\right\},$$

$$a_9 = \left\{\left[\frac{EI_r(GJ)'}{p_w}\right]'' - \frac{(GJ)'}{R^2}\left(1 - \frac{GJ}{p_w}\right) - \frac{(GJ)'}{p_w}\cdot\left\{2\left[\frac{EI_r(GJ)'}{p_w}\right]' + EI_r\cdot\left\{\left[\frac{(GJ)'}{p_w}\right]' - \frac{(GJ)'}{p_w}\right\}\right\}\right\}\cdot\left[\left(\frac{1}{\rho A\Omega^2}\right)' + \frac{(GJ)'}{p_w\rho A\Omega^2}\right] + 1 + \frac{1}{\rho A\Omega^2 R}\cdot\left\{-\frac{(GJ)'}{p_w}\left\{2\left[\frac{EI_r(GJ)'}{p_w}\right]' + EI_r\left\{\left[\frac{(GJ)'}{p_w}\right]' - \frac{(GJ)'}{p_w}\right\}\right\} + \left[\frac{EI_r(GJ)'}{p_w}\right]'' - \frac{(GJ)'}{R^2}\cdot\left(1 - \frac{GJ}{p_w}\right)\right\}',$$

$$a_{10} = \frac{GJ}{R^2} + \left[EI_r\frac{(GJ)'}{p_w}\right]' - EI_r\left[\frac{(GJ)'}{p_w}\right]^2.$$

References

[1] Michell, J. H., 1890, ''The Small Deformation of Curves and Surfaces With Application to the Vibrations of a Helix and a Circular Ring,'' Messenger Math., **19**, pp. 68–76.

[2] Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, 4th Ed., Dover, New York.

[3] Markus, S., and Nanasi, T., 1981, ''Vibrations of Curved Beams,'' Shock Vib. Digest, **7**, pp. 3–14.

[4] Laura, P. A. A., and Maurizi, M. J., 1987, ''Recent Research on Vibrations of Arch-Type Structures,'' Shock Vib. Digest, **7**, pp. 3–14.

[5] Childamparam, P., and Leissa, A. W., 1993, ''Vibrations of Planar Curved Beams, Rings and Arches,'' Appl. Mech. Rev., **46**, pp. 467–483.

[6] Auciello, N. M., and De Rosa, M. A., 1994, ''Free Vibrations of Circular Arches: A Review,'' J. Sound Vib., **174**, No. 4, pp. 433–458.

[7] Ojalvo, I. U., 1962, ''Coupled Twist-Bending Vibrations of Incomplete Elastic Rings,'' Int. J. Mech. Sci., **4**, pp. 53–72.

[8] Nelson, F. C., 1963, ''Out-of-Plane Vibration of a Clamped Circular Ring Segment,'' J. Acoust. Soc. Am., **35**, No. 6, pp. 933–934.

[9] Suzuki, K., Kosawada, T., and Takahashi, S., 1983, ''Out-of-Plane Vibrations of Curved Bars With Varying Cross-Section,'' Bull. JSME, **26**, No. 212, pp. 268–275.

[10] Royster, L. H., 1966, ''Effect of Linear Taper on the Lowest Natural Extensional Frequency of Elastic Arcs,'' ASME J. Appl. Mech., **33**, pp. 211–212.

[11] Reddy, M. N., 1968, ''Lateral Vibrations of Plane Curved Bars,'' J. Struct. Div. ASCE, **94**, No. 10, pp. 2197–2211.

[12] Bickford, F. C., and Strom, B. T., 1975, ''Vibration of Plane Curved Beams,'' J. Sound Vib., **39**, No. 2, pp. 135–146.

[13] Kawakami, M., Sakiyama, T., Matsuda, H., and Morita, C., 1995, ''In-Plane and Out-of-Plane Free Vibrations of Curved Beams With Variable Sections,'' J. Sound Vib., **187**, No. 3, pp. 381–401.

[14] Fung, Y. C., 1965, *Foundations of Solid Mechanics*, Prentice-Hall, Englewood Cliffs, NJ.

[15] Timoshenko, S., 1955, *Strength of Materials: Part I*, 3rd Ed., D. Van Nostrand, Princeton.

[16] Lee, L. S. S., 1975, ''Vibrations of an Intermediately Supported U-Bend Tube,'' ASME J. Eng. Ind., **97**, pp. 23–32.

[17] Meirovitch, L., 1967, *Analytical Methods in Vibrations*, Macmillan, New York.

[18] Lee, S. Y., and Kuo, Y. H., 1992, ''Exact Solutions for the Analysis of General Elastically Restrained Non-uniform Beams,'' ASME J. Appl. Mech., **59**, Part 2, pp. S189–S196.

[19] Rao, J. S., 1992, *Advanced Theory of Vibration*, Wiley Eastern Limited, New Delhi.

[20] Volterra, E., and Morell, J. D., 1961, ''Lowest Natural Frequency of Elastic Arc for Vibrations Outside the Plane of Initial Curvature,'' ASME J. Appl. Mech., **28**, pp. 624–627.

Dynamic Analysis of a One-Dimensional Poroviscoelastic Column

M. Schanz
Institute of Applied Mechanics,
Technical University Braunschweig,
D-38023 Braunschweig, Germany

A. H.-D. Cheng
Department of Civil and
Environmental Engineering,
University of Delaware,
Newark, DE 19716

The response due to a dynamic loading of a poroviscoelastic one-dimensional column is treated analytically. Biot's theory of poroelasticity is generalized to poroviscoelasticity using the elastic-viscoelastic correspondence principle in the Laplace domain. Damping effects of the solid skeletal structure and the solid material itself are taken into account. The fluid is modeled as in the original Biot's theory without any viscoelastic effects. The solution of the governing set of two coupled differential equations known from the purely poroelastic case is converted to the poroviscoelastic solution using the developed elastic-viscoelastic correspondence in Laplace domain. The time-dependent response of the column is achieved by the ''Convolution Quadrature Method'' proposed by Lubich. Some interesting effects of viscoelasticity on the response of the column caused by a stress, pressure, and displacement loading are studied. [DOI: 10.1115/1.1349416]

1 Introduction

For a wide range of fluid infiltrated materials, such as water saturated soils, oil impregnated rocks, or air filled foams, the elastic theory is a crude approximation. Due to presence of a second, interacting continuum, a different theory is necessary. The theory of porous materials containing a viscous fluid, known as the theory of poroelasticity, was introduced by Biot [1]. In subsequent years, this theory was extended to the anisotropic case ([2]), and also to dynamics ([3]). Following this development, the dynamic as well as the quasi-static analysis of a fully saturated porous continuum is possible. A comprehensive review of the quasi-static theory in rock mechanics can be found in the work of Detournay and Cheng [4].

In addition to the effect of the viscous fluid diffusion in the pores, the solid constituent, its skeleton, and its interaction with partially entrapped fluid can introduce time-dependent behavior as viscoelastic material. Further on, the rheology of pore fluid can exhibit viscoelastic behavior as well. This effect, however, will not be taken into account in the study here. The implementation of the solid viscoelastic effects in the theory of poroelasticity was first introduced by Biot [5]. Further work on this topic was done in the quasi-static case in [6] and in dynamics in [7], to cite a few. The last cited paper generalized Biot's theory to partially saturated continua.

Recently, a representation of the poroviscoelastic theory based on rheological modeling at micromechanical level was published by Abousleiman et al. [8]. It was argued that to have a physically consistent model, the rheology for the solid constituent and the skeletal structure should be clearly separated, and then combined to form a bulk continuum model. Based on this model, originally in quasi-static range, the current work examines its dynamic responses. The set of the governing differential equations for the dynamic case are deduced for a one-dimensional column. The corresponding analytical solution for one-dimensional column for the poroelastodynamic case has been presented by Schanz and Cheng [9]. The extension to poroviscoelasticity of this solution will be done in Laplace domain with the help of the elastic-viscoelastic correspondence principle.

With this solution, the frequency-dependent response of this column due to an impulsive load can be studied with respect to the influence of the viscoelasticity by taking the real part of the complex Laplace variable to zero. Then, the response of an arbitrary dynamical loaded system in time domain is given by the convolution integral of the impulse response function and the time-dependent loading. This convolution integral is numerically evaluated by the so-called ''Convolution Quadrature Method'' proposed by Lubich [10]. The weights of this quadrature formula are determined from the Laplace transformed impulse response function and a linear multistep method. In this method, no solution in time domain of the original problem is necessary. Through a series of stringent tests that includes a comparison with the highly acclaimed Dubner-Abate-Durbin-Crump method (e.g., [11] or [12]), our experience indicates that the Lubich method is one of the most robust in performing the inversion of wave-like functions that involves a significant number of cycles resulting from impact loading. This method has been, among other applications, successfully applied to a time domain formulation of the boundary element method ([13]).

2 Governing Equations

Following Biot's approach to model the behavior of porous media, the constitutive equations can be expressed as ([1])

$$\sigma_{ij} = 2G\epsilon_{ij} + \left(K - \frac{2}{3}G\right)\epsilon_{kk}\delta_{ij} - \alpha\delta_{ij}p \tag{1a}$$

$$\zeta = \alpha\epsilon_{kk} + \frac{\phi^2}{R}p, \tag{1b}$$

in which σ_{ij} denotes the total stress, p the pore pressure, ϵ_{ij} the strain of the solid frame, ζ the variation of fluid volume per unit reference volume, and δ_{ij} the Kronecker delta. In the above, the sign conventions for stress and strain follow that of elasticity, namely, tensile stresses and strains are denoted positive. The Latin indices takes the values 1, 2, 3 or 1, 2 in three-dimensional or two-dimensional cases, respectively, where summation convention is implied over repeated indices. The bulk material is defined by the material constants shear modulus G and the drained bulk compression modulus K. Biot's effective stress coefficient α, the porosity ϕ, and R complete the set of material parameters.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, December 12, 1999; final revision, July, 2000. Associate Editor: D. A. Siginer. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

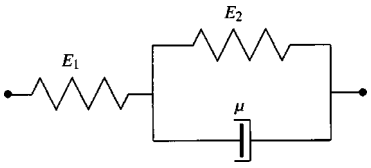


Fig. 1 One-dimensional rheological three-parameter model

In the constitutive equations above the only damping effects taken into account are caused by the interaction of the viscous fluid and the elastic solid. Introducing additionally viscoelasticity is done by means of the elastic-viscoelastic correspondence principle, as shown by Biot [5]. In a typical implementation in Laplace domain, the material constants shown in (1) are replaced by the corresponding functions of the Laplace variable, i.e., they become time-dependent. However, this approach provides little physical insight into the rheological models introduced, because the effective stress coefficient α, or R, or the pair of coefficients Q and R in the partial stress formulation of Biot [5], have no simple relation to the compression or shear behavior of the constituents. Rather, considerations of constitutive relation at micromechanical level ([4]) lead to a more rational model for our purpose

$$\alpha = 1 - \frac{K}{K_s} \tag{2}$$

and

$$R = \frac{\phi^2 K_f K_s^2}{K_f(K_s - K) + \phi K_s(K_s - K_f)}, \tag{3}$$

where K_s denotes the compression modulus of the solid grains and K_f the compression modulus of the fluid. With these expressions we are able to discuss how to implement viscoelastic behavior from a physical point of view.

Next, viscoelastic constitutive equations are introduced. From the two most common representations of viscoelastic constitutive equations, the hereditary integral or the differential operator formulation ([14]), the differential operator formulation is suitable to our purpose. The simplest model ensuring causal behavior is the three-parameter model, sometimes referred to as Kelvin-Voigt model (see Fig. 1). When the system is subjected to a step load, it instantly deforms in an elastic state characterized by the spring constant E_1. As time progresses, the resistance offered by the dash-pot diminishes and the system softens. At large times, the apparent spring constant becomes $E = E_1 E_2/(E_1 + E_2)$, which is smaller than the initial modulus E_1. The speed of the creep is regulated by the dash-pot viscosity μ. A characteristic time scale for the creep can be defined as $q = \mu/E_2$. The appropriate constitutive relation is given as

$$p = \frac{\mu}{E_1 + E_2}, \quad E = \frac{E_1 E_2}{E_1 + E_2}, \quad q = \frac{\mu}{E_2}$$
$$p\frac{\mathrm{d}}{\mathrm{d}t}\sigma + \sigma = E\left(\varepsilon + q\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon\right). \tag{4}$$

To find the elastic-viscoelastic correspondence, the differential Eq. (4) is transformed to Laplace domain

$$\hat{\sigma}(ps + 1) = E\hat{\epsilon}(1 + qs), \tag{5}$$

with $\mathcal{L}\{f(t)\} = \hat{f}(s)$ denotes the Laplace transform, with the complex Laplace variable s. Compared with Hook's law the elastic-viscoelastic correspondence is clearly observed,

$$E \to E\frac{1 + qs}{1 + ps}, \tag{6}$$

where the right-hand side is often called complex modulus.

In the explicit form of (2) and (3), the Kelvin-Voigt model can be applied to each of the moduli, corresponding to different physical effects. In detail:

- Replacing G by the complex modulus $\hat{G}(s)$ models a viscoelastic shear behavior of the solid frame.
- Replacing K_s by the complex modulus $\hat{K}_s(s)$ models a viscoelastic behavior of solid grains against volumetric deformation. This is necessary if the material has its own damping mechanism.
- Replacing K by the complex modulus $\hat{K}(s)$ models a viscoelastic behavior of the solid skeleton against volumetric deformation. Such a behavior can be caused, e.g., by micropores which are not connected to the main part of the fluid. The fluid in micropores can propagate through microcracks in the material causing damping due to the time required to reach localized equilibrium.
- Replacing K_f by the complex modulus $\hat{K}_f(s)$ models a viscoelastic behavior of the fluid. This, however, will not be attempted here for the following reasons: First, most pore fluids such as water or air are not viscoelastic. Second, a viscoelastic fluid can have shear stresses, which will interact with the surrounding solid. These effects are not modeled in Biot's theory. An arbitrarily generalization will not lead to a consistent theory.

Summarizing, in the following, a time-dependent compression and shear modulus of the solid $\hat{K}_s(s)$ and $\hat{G}(s)$ and a time-dependent bulk modulus $\hat{K}(s)$ are taken into account. This leads to the poroviscoelastic constitutive equations in Laplace domain as

$$\hat{\sigma}_{ij} = 2\hat{G}\hat{\epsilon}_{ij} + \left(\hat{K} - \frac{2}{3}\hat{G}\right)\hat{\epsilon}_{kk}\delta_{ij} - \hat{\alpha}\,\delta_{ij}\hat{p} \tag{7a}$$

$$\hat{\zeta} = \hat{\alpha}\hat{\epsilon}_{kk} + \frac{\phi^2}{\hat{R}}\hat{p}, \tag{7b}$$

with

$$\hat{\alpha}(s) = 1 - \frac{\hat{K}(s)}{\hat{K}_s(s)}$$

and

$$\hat{R}(s) = \frac{\phi^2 K_f \hat{K}_s^2(s)}{K_f(\hat{K}_s(s) - \hat{K}(s)) + \phi\hat{K}_s(s)(\hat{K}_s(s) - K_f)}. \tag{8}$$

Note, every formerly constant which is now indicated with $\hat{(\ )}$ is a function of s, respectively, of time. In the following, it is assumed that $\hat{K}_s(s)$, $\hat{G}(s)$, and $\hat{K}(s)$ are modeled as a three parameter model using the correspondence relation (6)

$$\hat{K}(s) = K\frac{1 + q_k s}{1 + p_k s}, \quad \hat{K}_s(s) = K_s\frac{1 + q_{ks} s}{1 + p_{ks} s}, \quad \hat{G}(s) = G\frac{1 + q_g s}{1 + p_g s}. \tag{9}$$

This completes the constitutive equations for a poroviscoelastic model. In the following, the functional argument (s) will be dropped for brevity.

Now, the governing set of differential equations are achieved by inserting (7) in the Laplace transformed dynamic equilibrium

$$\hat{\sigma}_{ij,j} + \hat{F}_i = \rho s^2 \hat{u}_i + \phi\rho_f s^2 \hat{v}_i, \tag{10}$$

and in the continuity equation

$$s\hat{\zeta} + \hat{q}_{i,i} = 0, \tag{11}$$

where $\rho = \rho_s(1 - \phi) + \phi\rho_f$ is the bulk density, with ρ_s and ρ_f denoting the solid and fluid density, respectively. As well the displacements of the solid are denoted by $\hat{u}_i$ and the relative fluid to solid displacements by $\hat{v}_i$. In Eqs. (10) and (11) and in the

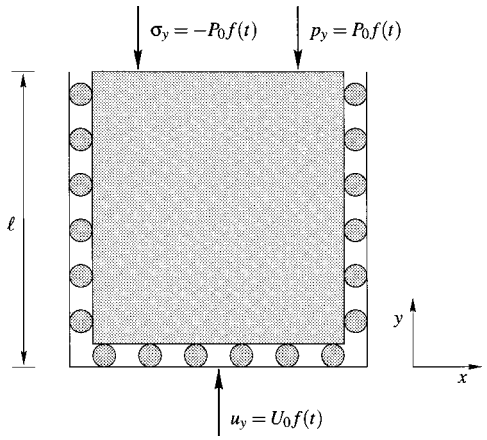


Fig. 2 One-dimensional column under dynamic loading

following, vanishing initial conditions for all variables are assumed. $\hat{F}_i$ are the bulk body forces which are neglected in the following, as only perturbation from the hydrostatic state is sought, and $\hat{q}_i = \phi s \hat{v}_i$ denotes the specific flux of the fluid. Finally, the derivative with respect to the spatial variable x_i is abbreviated by $(\,)_{,i}$.

Proceeding in taking Darcy's law

$$\hat{q}_i = -\kappa\left(\hat{p}_{,i} + s^2\rho_f\hat{u}_i + \frac{\rho_a + \phi\rho_f}{\phi}s^2\hat{v}_i\right) \tag{12}$$

into account, where κ denotes the permeability and ρ_a the apparent mass density, the final set of differential equations for the displacements $\hat{u}_i$ and the pore pressure $\hat{p}$ are achieved,

$$\hat{G}\hat{u}_{i,jj} + \left[\hat{K} + \frac{1}{3}\hat{G}\right]\hat{u}_{j,ij} - (\hat{\alpha} - \beta)\hat{p}_{,i} - s^2(\rho - \beta\rho_f)\hat{u}_i = 0 \tag{13}$$

$$\frac{\beta}{s\rho_f}\hat{p}_{,ii} - \frac{\phi^2 s}{\hat{R}}\hat{p} - (\hat{\alpha} - \beta)s\hat{u}_{i,i} = 0, \tag{14}$$

with the abbreviation

$$\beta = \frac{\phi^2 s\kappa\rho_f}{\phi^2 + s\kappa(\rho_a + \phi\rho_f)}. \tag{15}$$

For simplicity, the apparent mass density ρ_a is assumed to be frequency independent as $\rho_a \approx 0.66\phi\rho_f$ ([15]). With this set of equations the dynamic behavior of a poroviscoelastic continuum is completely defined.

3 Analytical Solution in One Dimension

A one-dimensional column of length l as sketched in Fig. 2 is considered. It is assumed that the sidewalls and the bottom are rigid, frictionless, and impermeable. Hence, the displacements normal to the surface are blocked and the column is otherwise free to slide parallel to the wall. At the top, the stress σ_y and the pressure p are prescribed. Due to these restrictions only the displacement u_y and the pore pressure p remain as degrees-of-freedom. This one-dimensional example can be used to study the influence of poroelastic parameters on wave propagation, or it can be used for actual application of finite and also semi-infinite columns by setting the layer depth l large. Here, we are particular interested in observing the interplay of the two compressional waves, a fast and a slow wave.

The governing set of differential Eqs. (13) and (14) is reduced to two scalar, coupled ordinary differential equations

$$\hat{E}\hat{u}_{,yy} - (\hat{\alpha} - \beta)\hat{p}_{,y} - s^2(\rho - \beta\rho_f)\hat{u}_y = 0 \tag{16}$$

$$\frac{\beta}{s\rho_f}\hat{p}_{,yy} - \frac{\phi^2 s}{\hat{R}}\hat{p} - (\hat{\alpha} - \beta)s\hat{u}_{y,y} = 0, \tag{17}$$

with the modulus $\hat{E} = \hat{K} + (4/3)\hat{G}$. The boundary conditions are

$$\hat{u}_y(y=0) = U_0, \quad \hat{q}(y=0) = 0$$

and

$$\hat{\sigma}(y=l) = -P_0, \quad \hat{p}(y=l) = P_0, \tag{18}$$

where an impulse function for the temporal behavior $f(t) = \delta(t)$, with $\delta(t)$ denoting the Dirac distribution, is assumed, together with vanishing initial conditions. Each of the nonzero boundary conditions in (18) represents a different type of loading. Due to the neglected body forces, this is a system of homogeneous ordinary differential equations with inhomogeneous boundary conditions. Such a system has been solved for the nonviscoelastic case in [9]. Inserting in these solutions the elastic-viscoelastic correspondences (9) leads to the solution of the poroviscoelastic problem above. As we are dealing with a linear problem the superposition principle is valid. Therefore, the solution can be divided in the three different load cases:

Stress Boundary Conditions. $\hat{u}_y(y=0) = 0$, $\hat{\sigma}(y=l) = -P_0$ and $\hat{p}(y=l) = 0$

$$\hat{u} = \frac{P_0}{\hat{E}(d_1\lambda_3 - d_3\lambda_1)}\left[\frac{d_3(e^{-\lambda_1(l-y)} - e^{-\lambda_1(l+y)})}{1 + e^{-2\lambda_1 l}} - \frac{d_1(e^{-\lambda_3(l-y)} - e^{-\lambda_3(l+y)})}{1 + e^{-2\lambda_3 l}}\right] \tag{19}$$

$$\hat{p} = \frac{P_0 d_1 d_3}{\hat{E}(d_1\lambda_3 - d_3\lambda_1)}\left[\frac{d_3(e^{-\lambda_1(l-y)} + e^{-\lambda_1(l+y)})}{1 + e^{-2\lambda_1 l}} - \frac{d_1(e^{-\lambda_3(l-y)} + e^{-\lambda_3(l+y)})}{1 + e^{-2\lambda_3 l}}\right] \tag{20}$$

Pressure Boundary Conditions. $\hat{u}_y(y=0) = 0$, $\hat{\sigma}(y=l) = 0$ and $\hat{p}(y=l) = P_0$

$$\hat{u} = \frac{P_0}{\hat{E}(d_1\lambda_3 - d_3\lambda_1)}\left[\frac{(\hat{E}\lambda_3 - \hat{\alpha}d_3)(e^{-\lambda_1(l-y)} - e^{-\lambda_1(l+y)})}{1 + e^{-2\lambda_1 l}} - \frac{(\hat{E}\lambda_1 - \hat{\alpha}d_1)(e^{-\lambda_3(l-y)} - e^{-\lambda_3(l+y)})}{1 + e^{-2\lambda_3 l}}\right] \tag{21}$$

$$\hat{p} = \frac{P_0}{\hat{E}(d_1\lambda_3 - d_3\lambda_1)}\left[\frac{d_1(\hat{E}\lambda_3 - \hat{\alpha}d_3)(e^{-\lambda_1(l-y)} + e^{-\lambda_1(l+y)})}{1 + e^{-2\lambda_1 l}} - \frac{d_3(\hat{E}\lambda_1 - \hat{\alpha}d_1)(e^{-\lambda_3(l-y)}e^{-\lambda_3(l+y)})}{1 + e^{-2\lambda_3 l}}\right] \tag{22}$$

Displacement Boundary Conditions. $\hat{u}_y(y=0) = U_0$, $\hat{\sigma}(y=l) = 0$ and $\hat{p}(y=l) = 0$

$$\hat{u} = \frac{U_0}{\hat{E}(\lambda_3^2 - \lambda_1^2)}\left[\frac{(\hat{E}\lambda_3^2 + s^2(\hat{\alpha}\rho_f - \rho))(e^{-\lambda_1(2l-y)} + e^{-\lambda_1 y})}{1 + e^{-2\lambda_1 l}} - \frac{(\hat{E}\lambda_3^2 + s^2(\hat{\alpha}\rho_f - \rho))(e^{-\lambda_3(2l-y)} + e^{-\lambda_3 y})}{1 + e^{-2\lambda_3 l}}\right] \tag{23}$$

Table 1 Material data of Berea sandstone, a soil, and a sediment

	$K\left[\frac{N}{m^2}\right]$	$G\left[\frac{N}{m^2}\right]$	$\rho\left[\frac{kg}{m^3}\right]$	φ	$K_s\left[\frac{N}{m^2}\right]$	$\rho_f\left[\frac{kg}{m^3}\right]$	$K_f\left[\frac{N}{m^2}\right]$	$\kappa\left[\frac{m^4}{Ns}\right]$
rock	8×10^9	6×10^9	2458	0.19	3.6×10^{10}	1000	3.3×10^9	1.9×10^{-10}
soil	2.1×10^8	9.8×10^7	1884	0.48	1.1×10^{10}	1000	3.3×10^9	3.55×10^{-9}
sediment	3.7×10^7	2.2×10^7	1396	0.76	3.6×10^{10}	1000	2.3×10^9	1×10^{-8}

$$\hat{p}=\frac{U_0}{\hat{E}(\lambda_3^2-\lambda_1^2)}\left[\frac{d_1(\hat{E}\lambda_3^2+s^2(\hat{\alpha}\rho_f-\rho))(e^{-\lambda_1(2l-y)}-e^{-\lambda_1 y})}{1+e^{-2\lambda_1 l}}\right.$$
$$\left.-\frac{d_3(\hat{E}\lambda_3^2+s^2(\hat{\alpha}\rho_f-\rho))(e^{-\lambda_3(2l-y)}-e^{-\lambda_3 y})}{1+e^{-2\lambda_3 l}}\right] \quad (24)$$

The corresponding stresses and fluxes are calculated with the one-dimensional form of the constitutive Eq. (7*a*)

$$\hat{\sigma}(s,y)=\hat{E}\hat{u}_{y,y}-\hat{\alpha}\hat{p} \quad (25)$$

and the one-dimensional form of Darcy's law (12)

$$\hat{q}(s,y)=-\frac{\beta}{s\rho_f}(\hat{p}_{,y}+s^2\rho_f\hat{u}_y). \quad (26)$$

With the solutions above, the frequency-dependent harmonic response of a one-dimensional poroviscoelastic column can be studied by taking $s=-i\omega$. However, the time-dependent response due to an arbitrary excitation $f(t)$, is achieved by the convolution integral, e.g., for the displacements

$$u_y(t,y)=\int_0^t \mathcal{L}^{-1}\{\hat{u}_y(s,y)\}(\tau,y)f(t-\tau)d\tau, \quad (27)$$

where $\mathcal{L}^{-1}$ is the inverse Laplace transform operator. Another way to obtain solution of arbitrary transient input is to take the advantage of the property of Laplace transform

$$u_y(t,y)=\mathcal{L}^{-1}\{\hat{u}_y(s,y)\hat{f}(s)\} \quad (28)$$

where $\hat{f}(s)$ is the Laplace transform of the boundary condition $f(t)$.

We now have two possibilities to evaluate the response in time domain. We can either multiply the impulse response functions (19)–(24) by the input excitation in Laplace domain, $\hat{f}(s)$, with a subsequent numerical inverse transformation as indicated in (28), or we use the ''Convolution Quadrature Method'' proposed by Lubich [10] to directly tackle (27). The first choice, with all its advantages and disadvantages, is the traditional approach (see, e.g., [12] or [11]). But, in this case here, where the one function in the convolution integral (27) is available in Laplace domain and the other function in time domain, it is preferable to take the Convolution Quadrature Method. This method approximates the convolution integral (27) numerically by a quadrature formula

$$u_y(n\Delta t)=\sum_{k=0}^{n}\omega_{n-k}(\Delta t)f(k\Delta t), \quad n=0,1,\ldots,N, \quad (29)$$

whose weights $\omega_{n-k}(\Delta t)$ are determined with the help of the Laplace transformed impulse response functions $\hat{u}_y(s,y)$ and a function $\gamma(s)$ that defines the linear multistep method

$$\omega_n(\Delta t)-\frac{\mathcal{R}^{-n}}{L}\sum_{l=0}^{L-1}\hat{u}_y\left(\frac{\gamma(\mathcal{R}e^{il2\pi/L})}{\Delta t}\right)e^{-inl2\pi/L}. \quad (30)$$

More details of the method and the definition of parameters can be found in Appendix A. In the following, this method is used to perform the time-dependent responses, choosing a backward differentiation formula of order 2 (BDF 2) as the underlying multistep method.

4 Results in Frequency and Time Domain

With the analytical solution developed in Section 3, the influence of different damping mechanisms is studied. Three very different materials, ranked in descending order of stiffness, a rock (Berea sandstone) ([16]), a soil (coarse sand) ([17]), and a seabed sediment ([18]), are chosen to cover a wide range of material properties. The material data are given in Table 1. We observe that the stiffness of the material, in terms of frame bulk modulus and shear modulus, spans more than two orders of magnitude. The value of bulk density decreases as porosity increases. The fluid bulk modulus for sediment is different because sea water was referred in [18]. The bulk moduli of solid grain are about the same. We should point out that for the soil case, dense sand saturated with silicon oil was used in [17]. In Table 1, however, the fluid was changed to water with other material coefficients consistently converted. Finally, we observe that the permeability also spans more than two orders of magnitude.

In the constitutive Eq. (7) the bulk modulus $\hat{K}$, the shear modulus $\hat{G}$, and the compression modulus of the solid itself $\hat{K}_s$ are each chosen to be viscoelastic, modeled by a three-parameter model. For each of them, the values of p and q need to be given. However, to the authors' best knowledge, no such data have been reported in the literature. Therefore, the same set of data is somewhat arbitrarily chosen for the three materials. To compare the influence of viscoelasticity in different moduli on the dynamic response, four different cases are considered:

Case 1: Only the bulk compression modulus $\hat{K}(s)$ is modeled viscoelastic: $p_k=1[\frac{1}{s}]$, $q_k=1.5[\frac{1}{s}]$ and $p_{ks}=p_g=q_{ks}=q_g=0[\frac{1}{s}]$.

Case 2: Only the shear modulus $\hat{G}(s)$ is modeled viscoelastic: $p_g=1[\frac{1}{s}]$, $q_g=1.5[\frac{1}{s}]$ and $p_{ks}=p_k=q_{ks}=q_k=0[\frac{1}{s}]$.

Case 3: Only the compression modulus of the solid material $\hat{K}_s(s)$ is modeled viscoelastic: $p_{ks}=1[\frac{1}{s}]$, $q_{ks}=1.5[\frac{1}{s}]$ and $p_k=p_g=q_k=q_g=0[\frac{1}{s}]$.

Case 4: The purely poroelastic case without any viscoelasticity: $p_{ks}=q_{ks}=p_k=p_g=q_k=q_g=0[\frac{1}{s}]$.

Before solving the transient problems, the frequency response of a column with length $l=1$ m is first considered. In Fig. 3 the absolute value of the displacements $\hat{u}_y(\omega,y=l)$ at the top of the column are plotted versus frequency ω for the three materials. As boundary condition, a constant step pressure loading (without total stress) is assumed. In Fig. 3 we observe resonance peaks as expected. The first resonance frequency is around 2000 Hz for the sediment, which increases to about 5000 Hz for the rock. The various curves correspond to different assumptions of viscoelasticity, referred to as case 1 to 4 in the above. It is found that the sediment response is least affected by viscoelastic effect—there is basically no shift in eigenfrequencies and only a slight damping in response amplitude. This is in accordance with our model, be-

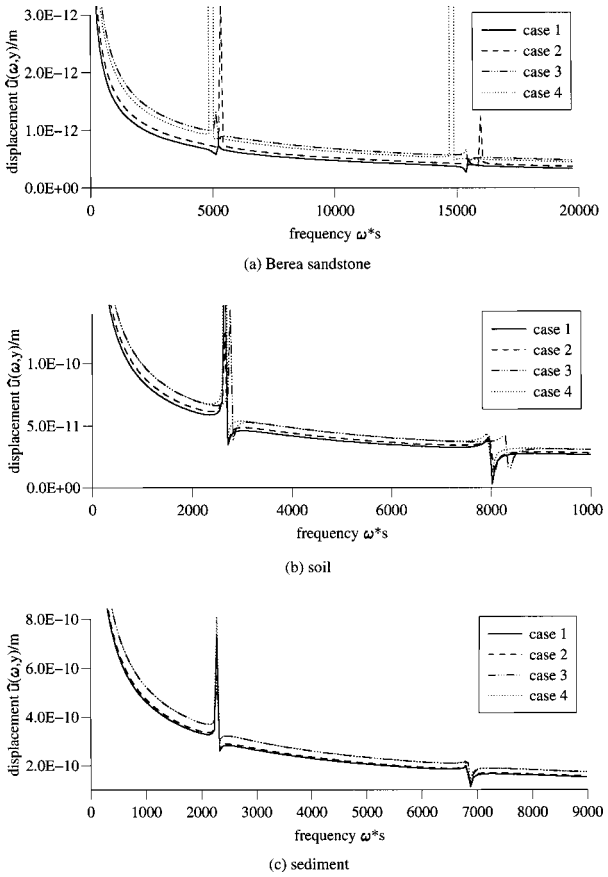


Fig. 3 Absolute value of the displacements $|\hat{u}_y(\omega, y=l)|$ at the top of the column versus frequency ω

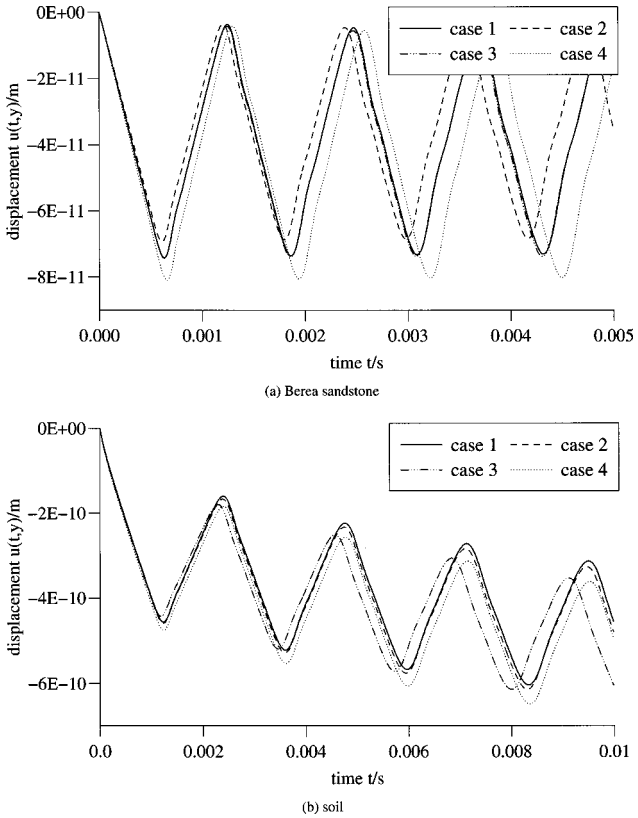


Fig. 4 Displacements u_y $(t, y=l)$ at the top of the column versus time t

cause the sediment bulk property is dominated by the fluid, which is elastic. The viscoelastic solid hence contributes to only a secondary influence. The soil response is also less influenced. There is a small shifting of eigenfrequencies, and a somewhat larger damping than the sediment case. The largest effects are found in the rock material. Not only there exists larger damping, particularly on the resonance peaks, but also significant shift of eigenfrequencies occurs. We further observe that for all materials, the largest damping results from the viscoelasticity of bulk compression modulus. For soil, the largest shift of eigenfrequencies results from the viscoelastic effect of $\hat{K}_s$, compared to rock where $\hat{G}$ has the most influence. This shows that the effect of each modulus is different in different materials.

For the frequency response of the other two boundary conditions, a stress and a displacement loading, the influence of viscoelasticity exhibits similar trend. Hence it is enough to show the results for just this boundary condition.

Now, the time-dependent behavior is considered. Due to the relative insensitivity of sediment response to viscoelasticity, only results for the two other materials are presented. In Fig. 4 the displacements $u(t, y=l)$ at the top of the column, caused by a step stress loading $\sigma(t, y=l) = -H(t) N/\mathrm{m}^2$, are depicted versus time. In each of the curves, a different time-step size is used for the Lubich method, due to different wave speeds of the materials. For the Berea sandstone $\Delta t = 1 \times 10^{-5}$ s and for the soil $\Delta t = 2 \times 10^{-5}$ s are used, with $N = 500$ time steps. As with any numerical method, too large a time-step size leads to worse results due to inadequate approximation of the time history of the displacements.

In Fig. 4, the rock displacements show an oscillation similar to that for an elastic material, whereas for the soil, the oscillation is combined with a settlement, due to the well-known consolidation effect. We notice that the wave speed is modified in both materials. Case 4, the case without viscoelasticity, has the slowest wave speed, by observing the time it takes the wave to transverse the column. This is not surprising, because by setting the two parameters p and q in (6) to zero, case 4 has the smallest modulus. In the viscoelastic cases, the apparent modulus of the material is between $1.5E$ for small time (or high frequency), and E for large time (or low frequency), due to the p and q values used. Hence the wave speed of the viscoelastic and the elastic cases should not be directly compared. However, among the viscoelastic cases, we can compare and observe that different modulus has different effect on the two materials. The fastest wave in the rock is associated with the viscoelasticity of shear modulus. The fastest wave in soil, on the other hand, is observed to be associated with the solid compression modulus. The oscillation amplitude is found to be the smallest also in these two cases, respectively, for soil and rock. These are consistent with the observation in frequency domain.

We further tested cases with an increased damping value p, which is observed to enhance the damping effect. But as mentioned before, no measured damping values are available. It is not possible to say whether these assumed damping values are realistic or not. Therefore, these results are not presented here.

We next investigate wave propagation in this one-dimensional column with the aim of capturing the two compressional waves, a fast and a slow wave. These two waves have been identified for the poroelastic case as presented by Schanz and Cheng [9]. To clearly observe these two waves, a semi-infinite column is used to eliminate reflections at the ends that can confuse the arrival of the two different waves. An observation of pressure is made at 5 m below the top surface, where a step stress loading is applied.

Our experience in the poroelastic cases ([9]) has indicated that the second compressional wave dissipates rapidly. With the permeability of these used materials, the second wave will not survive with a detectable magnitude at 5 m below surface. To en-

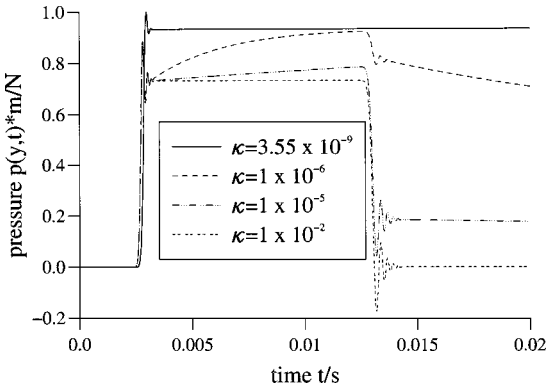


Fig. 5 Pressure $p(t, y=995\,\mathrm{m})$ versus time: wave propagation for different values of κ in an "infinite" soil column

hance the observation of the second wave, artificially large permeabilities, or different materials, must be used. For the present purpose, the permeabilities are arbitrarily increased.

In Fig. 5 the pressure at 5 m below surface in an infinitely long soil column, modeled by $l=1000$ m, caused by a stress Heaviside boundary condition, is plotted versus time for different values of the permeability κ, ranging from 10^{-9} to $10^{-2}\,\mathrm{m}^4/(N\,s)$, and $p_k = p_g = p_{ks} = 1.5(1/s)$. Other material properties are referred to Table 1.

Let us first examine the case with highest permeability, $\kappa = 10^{-2}\,\mathrm{m}^4/(N\,s)$. We observe a step rise in pressure that indicates the arrival of the first wave around 2.5 ms. The pressure stays roughly constant until at 13 ms. At that time, the second wave arrives and negates the positive pressure. Since there is no boundary reflection, the identity of the second wave is clearly established. We should point out that the small fluctuation around the pressure front is an artifact of the numerical method, which generally cannot be avoided. It is, however, small enough to be tolerated. With decreasing permeability, the first wave arrives at increasingly larger amplitude, and the second wave at smaller amplitude. In the case of $\kappa = 3.55\times10^{-9}\,\mathrm{m}^4/(N\,s)$, the curve is flat after the arrival of the first wave, which means that the second wave arrives with an undetectable amplitude due to viscous damping. These dynamic behaviors are similar to those in the poroelastic cases without viscoelasticity, as discussed in more detail in [9].

Once the general dynamic behavior is established, the influence of viscoelasticity in the individual modulus is studied in Fig. 6. As in Fig. 5, the pressure $p(t, y=995\,\mathrm{m})$ due to a stress Heaviside step loading is plotted versus time. But here the cases 1 to 4 defined in the beginning of this section are examined. To enhance the observation of the second wave, the largest permeability case used in the preceding example, $\kappa = 10^{-2}\,\mathrm{m}^4/(N\,s)$, is used here. Similar to the investigation above, the viscoelasticity of different modulus has different effects on the two materials. First of all, we observe that the wave velocities are modified, much more so for the second wave than for the first wave. The arrivals of the first waves are close to each other. Nevertheless, in both materials case 4 gives the slowest first wave. In rock, case 2 has the fastest first wave, and in soil, it is case 3. These are consistent with earlier observations. The second wave, on the other hand, is more complicated. In most cases the second wave of the viscoelastic cases travels faster than the nonviscoelastic one, case 4. However, in case 3, where only the solid grain modulus is modeled viscoelastic, the first wave becomes faster, but the second wave becomes slower than case 4. We also observe that there is significant amplitude reduction of the first wave for the rock material when viscoelasticity is present, except for case 3, where the amplitude increases. For the soil, there is little change in amplitude.

5 Conclusions

In the presented work, Biot's theory of porous media is extended to poroviscoelasticity by means of the elastic-viscoelastic

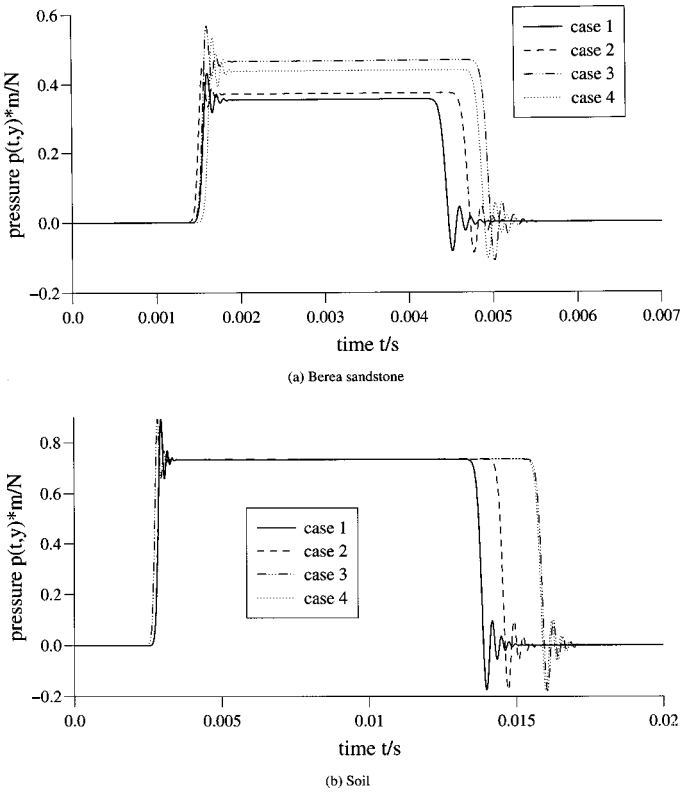


Fig. 6 Pressure $p(t, y=995\,\mathrm{m})$ versus time: wave propagation for different damping cases

correspondence principle. A physically more appealing model is implemented that separates the viscoelasticity to three components—that due to the compression of solid frame, of solid grain, and the shearing of solid frame. A three-parameter rheological model is applied to each of them. The fluid is modeled as viscous and Newtonian, as in the classical theory. Next, an analytical solution of a one-dimensional column is derived in Laplace domain. Then, with the Convolution Quadrature Method the time-dependent behavior is achieved.

Three widely different materials, a rock, a soil, and a sediment, are used in the analysis. The viscoelastic effect is found to be stronger in rock and soil, than in sediment. The rock is shown to be more influenced by the shear modulus whilst the soil is more affected by the compression modulus of the grains. In the frequency domain, shifting of resonance frequencies and damping of resonance peaks are observed. In the time domain with a step stress loading, viscoelastic effect generally leads to an increase in wave speed for both the fast and the slow waves, and a decrease in amplitude, except for case 3 in rock. This shows the conclusions drawn here are not entirely general, and are material dependent.

Acknowledgment

Martin Schanz was supported by the German Research Foundation (DFG) under grant SCHA 527/4-1. The financial support is gratefully acknowledged.

Appendix A

Convolution Quadrature Method. The ''Convolution Quadrature Method'' developed by Lubich numerically approximates a convolution integral

$$y(t)=\int_0^t f(t-\tau)g(\tau)d\tau \rightarrow y(n\Delta t)=\sum_{k=0}^{n}\omega_{n-k}(\Delta t)g(k\Delta t),$$
$$n=0,1,\ldots,N, \tag{31}$$

by a quadrature rule whose weights are determined by the Laplace transformed function $\hat{f}$ and a linear multistep method. This method was originally published in [10] and [19]. Application to the boundary element method may be found in [20]. Here, a brief overview of the method is given.

In formula (31) the time t is divided in N equal steps Δt. The weights $\omega_n(\Delta t)$ are the coefficients of the power series

$$\hat{f}\left(\frac{\gamma(z)}{\Delta t}\right)=\sum_{n=0}^{\infty}\omega_n(\Delta t)z^n, \tag{32}$$

with the complex variable z. The coefficients of a power series are usually calculated with Cauchy's integral formula. After a polar coordinate transformation, this integral is approximated by a trapezoidal rule with L equal steps $2\pi/L$. This leads to

$$\omega_n(\Delta t)=\frac{1}{2\pi i}\int_{|z|=\mathcal{R}}\hat{f}\left(\frac{\gamma(z)}{\Delta t}\right)z^{-n-1}dz$$
$$\approx\frac{\mathcal{R}^{-n}}{L}\sum_{l=0}^{L-1}\hat{f}\left(\frac{\gamma(\mathcal{R}e^{il\,2\pi/L})}{\Delta t}\right)e^{-inl\,2\pi/L}, \tag{33}$$

where $\mathcal{R}$ is the radius of a circle in the domain of analyticity of $\hat{f}(z)$.

The function $\gamma(z)$ is the quotient of the characteristic polynomials of the underlying multistep method, e.g., for a BDF 2, $\gamma(z)=3/2-2z+(1/2)z^2$. The used linear multistep method must be $A(\alpha)$-stable and stable at infinity ([19]). Experience shows that the BDF 2 is the best choice ([21]). Therefore, t is used in all calculations in this paper.

If one assumes that the values of $\hat{f}(z)$ in (33) are computed with an error bounded by ϵ, then the choice $L=N$ and $\mathcal{R}^N=\sqrt{\epsilon}$ yields an error in ω_n of size $\mathcal{O}(\sqrt{\epsilon})$ ([10]). Several tests conducted by the authors lead to the conclusion that the parameter $\epsilon=10^{-10}$ is the best choice for the kind of functions dealt with in this paper ([13]). The assumption $L=N$ results in N^2 coefficients $\omega_n(\Delta t)$ to be calculated. Due to the exponential function at the end of formula (33) this can be done very fast using the technique of the Fast Fourier Transformation (FFT).

References

[1] Biot, M. A., 1941, ''General Theory of Three-Dimensional Consolidation,'' J. Appl. Phys., **12**, pp. 155–164.
[2] Biot, M. A., 1955, ''Theory of Elasticity and Consolidation for a Porous Anisotropic Solid,'' J. Appl. Phys., **26**, pp. 182–185.
[3] Biot, M. A., 1956, ''Theory of Propagation of Elastic Waves in a Fluid-Saturated Porous Solid. I. Low-Frequency Range, II. Higher Frequency Range,'' J. Acoust. Soc. Am., **28**, No. 2, pp. 168–191.
[4] Detournay, E., and Cheng, A. H.-D., 1993, *Fundamentals of Poroelasticity*, Vol. II (Comprehensive Rock Engineering: Principles, Practice & Projects), Pergamon Press, Tarrytown, NY, Chapter 5, pp. 113–171.
[5] Biot, M. A., 1956, ''Theory of Deformation of a Porous Viscoelastic Anisotropic Solid,'' J. Appl. Phys., **27**, No. 5, pp. 459–467.
[6] Wilson, R. K., and Aifantis, E. C., 1982, ''On the Theory of Consolidation With Double Porosity,'' Int. J. Eng. Sci., **20**, pp. 1009–1035.
[7] Vgenopoulou, I., and Beskos, D. E., 1992, ''Dynamic Behavior of Saturated Poroviscoelastic Media,'' Acta Mech., **95**, pp. 185–195.
[8] Abousleiman, Y., Cheng, A. H.-D., Jiang, C., and Roegiers, J.-C., 1996, ''Poroviscoelastic Analysis of Borehole and Cylinder Problems,'' Acta Mech., **109**, No. 1–4, pp. 199–219.
[9] Schanz, M., and Cheng, A. H.-D., 2000, ''Transient Wave Propagation in a One-Dimensional Poroelastic Column,'' Acta Mech., **145**, No. 1-4, pp. 1–8.
[10] Lubich, C., 1988, ''Convolution Quadrature and Discretized Operational Calculus. I.,'' Numer. Math., **52**, pp. 129–145.
[11] Narayanan, G. V., and Beskos, D. E., 1982, ''Numerical Operational Methods for Time-Dependent Linear Problems,'' Int. J. Numer. Methods Eng., **18**, pp. 1829–1854.
[12] Cheng, A. H.-D., Sidauruk, P., and Abousleiman, Y., 1994, ''Approximate Inversion of the Laplace Transform,'' Mathematica J. **4**, No. 2, pp. 76–82.
[13] Schanz, M., and Antes, H., 1997, ''Application of 'Operational Quadrature Methods' in Time Domain Boundary Element Methods,'' Meccanica, **32**, No. 3, pp. 179–186.
[14] Christensen, R. M., 1971, *Theory of Viscoelasticity*, Academic Press, New York.
[15] Bonnet, G., and Auriault, J.-L., 1985, ''Dynamics of Saturated and Deformable Porous Media: Homogenization Theory and Determination of the Solid-Liquid Coupling Coefficients,'' N. Boccara and M. Daoud, eds., *Physics of Finely Divided Matter*, Springer-Verlag, Berlin, pp. 306–316.
[16] Cheng, A. H.-D., Badmus, T., and Beskos, D. E., 1991, ''Integral Equations for Dynamic Poroelasticity in Frequency Domain With BEM Solution,'' J. Eng. Mech. **117**, No. 5, pp. 1136–1157.
[17] Kim, Y. K., and Kingsbury, H. B., 1979, ''Dynamic Characterization of Poroelastic Materials,'' Exp. Mech., **19**, pp. 252–258.
[18] Badiey, M., Cheng, A. H.-D., and Mu, Y., 1998, ''From Geology to Geoacoustics—Evaluation of Biot-Stoll Sound Speed and Attenuation for Shallow Water Acoustics,'' J. Acoust. Soc. Am., **103**, No. 1, pp. 309–320.
[19] Lubich, C., 1988, ''Convolution Quadrature and Discretized Operational Calculus. II,'' Numer. Math., **52**, pp. 413–425.
[20] Schanz, M., and Antes, H., 1997, ''A New Visco- and Elastodynamic Time Domain Boundary Element Formulation,'' Comput. Mech., **20**, No. 5, pp. 452–459.
[21] Schanz, M., 1999, ''A Boundary Element Formulation in Time Domain for Viscoelastic Solids,'' Commun. Numer. Meth. Eng. **15**, pp. 799–809.

Bifurcations of Eigenvalues of Gyroscopic Systems With Parameters Near Stability Boundaries

A. P. Seyranian
Institute of Mechanics,
Moscow State Lomonosov University,
Michurynski pr. 1,
Moscow 117192, Russia
e-mail: seyran@imec.msu.ru

W. Kliem
Department of Mathematics,
Technical University of Denmark,
Building 303,
DK-2800 Kgs. Lyngby, Denmark
e-mail: w.kliem@mat.dtu.dk

This paper deals with stability problems of linear gyroscopic systems $M\ddot{x}+G\dot{x}+Kx=0$ with finite or infinite degrees-of-freedom, where the system matrices or operators depend smoothly on several real parameters. Explicit formulas for the behavior of eigenvalues under a change of parameters are obtained. It is shown that the bifurcation (splitting) of double eigenvalues is closely related to the stability, flutter, and divergence boundaries in the parameter space. Normal vectors to these boundaries are derived using only information at a boundary point: eigenvalues, eigenvectors, and generalized eigenvectors, as well as first derivatives of the system matrices (or operators) with respect to parameters. These results provide simple and constructive stability and instability criteria. The presented theory is exemplified by two mechanical problems: a rotating elastic shaft carrying a disk, and an axially moving tensioned beam. [DOI: 10.1115/1.1356417]

Dedicated to Prof. Peter Lancaster on the occasion of his 70th birthday.

1 Introduction

The theory of gyroscopic systems has a history which is more than 100 years old. One of the first investigations in this field was carried out by Thomson and Tait [1]. The possibility of gyroscopic stabilization of unstable conservative systems has to be taken into account in the dynamics of all kinds of rotating bodies such as elastic shafts, satellites, spaceships, etc. Nonrotating systems like fluid conveying pipes can also be influenced by gyroscopic forces.

Here we restrict ourselves to mentioning only a few of the numerous books and papers on this subject: Hagedorn [2], Huseyin et al. [3], Walker [4], Barkwell and Lancaster [5], Seyranian [6], Veselic [7], Seyranian et al. [8], Lancaster and Zizler [9], Kliem and Seyranian [10], and Mailybaev and Seyranian [11]. One will find further references in these articles, as well as—concerning older literature—in books by Müller [12], Huseyin [13], and Merkin [14].

The mathematical background of the present work is related to the classical papers on perturbation theory for eigenvalues of operators by Vishik and Lyusternik [15] and Lidskii [16]. A comprehensive review of this theory with some extensions has been given by Moro et al. [17]. The results of [15] were applied to vibrational systems and extended to multiple parameters by Seyranian [18].

Concerning the question of behavior of eigenvalues under change of a single parameter, Renshaw and Mote [19] formulated a conjecture on the stability of gyroscopic systems near the zero eigenvalue (divergence boundary). Lancaster and Kliem [20] showed by two counterexamples that this conjecture is not generally true. Renshaw [21] (for the special case of 2×2 systems with zero eigenvalue) and Hryniv et al. [22] reformulated the Renshaw-Mote conjecture with an additional condition and showed its validity. In a recent paper, Parker [23] deals with continuous gyroscopic systems (rotors) with the angular velocity as parameter. When studying stability he concentrates on a perturbation analysis in the vicinity of the zero eigenvalue (divergence boundary). To avoid nondifferentiability of the eigenvalue λ, he introduces a formulation of the system in terms of λ^2 (which is differentiable). However, this procedure does not work in the case of eigenvalues near the flutter boundary. Note that most of the cited papers deal with one-parameter problems.

Our paper is organized as follows. After a general formulation of the problem with *several parameters* we derive perturbation formulas for simple and semi-simple eigenvalues. Then we study a semi-simple double eigenvalue, and a double eigenvalue with a single eigenvector in detail. These cases are mainly concerned with eigenvalue behavior near the *stability-flutter boundary*. Afterwards we concentrate on the eigenvalue locus near the *stability-divergence boundary* (double zero eigenvalue). The degenerate case explaining the criterion by Hryniv et al. [22] is investigated as well. In all cases *explicit expansions* for eigenvalues near the stability boundaries are obtained and analyzed. Finally we illustrate the developed results with mechanical examples.

2 Eigenvalue Problem for a Gyroscopic System

The equation of motion for a linear gyroscopic system is

$$M\ddot{x}+G\dot{x}+Kx=0, \tag{1}$$

where M, G, and K are linear differential operators in some domain P if we are modeling a continuum, or real $m\times m$ matrices in the case of a discrete system. Dots mean differentiation with respect to time of the vector x of generalized coordinates (or deflection function). The system operators or matrices represent time-independent mass, gyroscopic, and potential forces, respectively. We introduce the inner product by $(v,w)=\int_P v\bar{w}dP$ and assume that the operators M and K are self-adjoint and G is skew-adjoint. This means that $(Mv,w)=(v,Mw)$, $(Kv,w)=(v,Kw)$, and $(Gv,w)=-(v,Gw)$ for all admissible functions v and w satisfying the boundary conditions. In the case of matrices, M and K are symmetric, $M^T=M$, $K^T=K$ and G is skew-symmetric, $G^T=-G$. Besides we consider M to be positive definite, (Mu,u)

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 1, 2000; final revision, June 25, 2000. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

>0. It is assumed that the system operators or matrices smoothly depend on a vector of real parameters $p=(p_1,p_2,\ldots,p_n)^T$.

For the sake of simplicity in the following we drop to mention every time both the operator and the matrix possibility, since the theory is essentially the same. Therefore we will give an outline of the theory in matrix formulation. Examples will be given for continua as well as for discrete systems.

Consider the eigenvalue problem corresponding to (1)

$$(\lambda^2 M+\lambda G+K)u=0, \tag{2}$$

where λ is an eigenvalue, and u a corresponding eigenvector. The eigenvalues λ are determined from the characteristic equation

$$\det(\lambda^2 M+\lambda G+K)=0. \tag{3}$$

It is well known ([13,24]), that if λ is an eigenvalue of (2), then $-\lambda$, $\bar{\lambda}$ and $-\bar{\lambda}$ are also eigenvalues (a bar denotes complex conjugate). This means symmetry of the eigenvalues with respect to the imaginary and real axis and is also called Hamiltonian symmetry according to the fact that gyroscopic systems are Hamiltonian. Then it follows that the stability of the gyroscopic system can only be attained when all eigenvalues under a change of the parameters reach the imaginary axis. If all the roots of Eq. (3) are purely imaginary and simple, $\lambda_j=i\omega_j$, then the general solution of Eq. (1) has the form

$$x=\sum_{j=1}^{m}\left(c_j u_j e^{i\omega_j t}+\bar{c}_j\bar{u}_j e^{-i\omega_j t}\right) \tag{4}$$

implying the stability of the system. In (4) c_j are arbitrary complex constants. With a change of the parameters $p_1,p_2,\ldots,p_n$, simple purely imaginary eigenvalues move along the imaginary axis. They cannot leave this axis due to the Hamiltonian symmetry of the eigenvalues. This means that the stability of the gyroscopic system can only be lost when some simple eigenvalues meet on the imaginary axis (become multiple).

For the sake of convenience we introduce the inner product of vectors a and b in the complex space C^m by

$$(a,b)=b^*a=\sum_{j=1}^{m} a_j\bar{b}_j. \tag{5}$$

A star after a symbol means the transposed and complex conjugate quantity.

Along with (2) we consider the adjoint eigenvalue problem with eigenvector v

$$(\lambda^2 M+\lambda G+K)^* v=0. \tag{6}$$

If we introduce the notation

$$L:=\lambda^2 M+\lambda G+K \tag{7}$$

then we have

$$L^*=\bar{\lambda}^2 M^T+\bar{\lambda}G^T+K^T=\bar{\lambda}^2 M-\bar{\lambda}G+K. \tag{8}$$

Note that if the eigenvalue is purely imaginary, $\lambda=i\omega$, then the matrix operator L is Hermitian, i.e., $L=L^*$. This follows directly from (7) and (8). For this case the adjoint eigenvalue problems (2) and (6) coincide, such that the eigenvectors of the two problems can be chosen to coincide as well,

$$v=u. \tag{9}$$

3 Perturbation of Eigenvalues

We assume that at a point $p=p_0$ in the parameter space R^n the gyroscopic system possesses a purely imaginary eigenvalue $\lambda_0=i\omega_0$, simple or multiple. Our task is to study the behavior of eigenvalues in the vicinity of the initial point p_0.

For this purpose we consider a parameter variation $p=p_0+\varepsilon e$, where $e=(e_1,\ldots,e_n)$ is a direction vector of unit norm $|e|=1$, and ε is a small positive number. Since the matrices of the gyroscopic system (1) depend smoothly on the vector of parameters p, we obtain

$$\begin{gathered} M=M_0+\varepsilon M_1+\varepsilon^2 M_2+\ldots, \quad G=G_0+\varepsilon G_1+\varepsilon^2 G_2+\ldots, \\ K=K_0+\varepsilon K_1+\varepsilon^2 K_2+\ldots \end{gathered} \tag{10}$$

where

$$\begin{gathered} M_0=M(p_0), \quad G_0=G(p_0), \quad K_0=K(p_0), \\ M_1=\sum_{j=1}^{n}\frac{\partial M}{\partial p_j}e_j, \quad G_1=\sum_{j=1}^{n}\frac{\partial G}{\partial p_j}e_j, \\ K_1=\sum_{j=1}^{n}\frac{\partial K}{\partial p_j}e_j, \quad M_2=\frac{1}{2}\sum_{j,k=1}^{n}\frac{\partial^2 M}{\partial p_j\partial p_k}e_j e_k \end{gathered} \tag{11}$$

with respective expressions for G_2, K_2, etc.

Due to the variation of the vector p, the eigenvalue λ and corresponding eigenvector u take increments which can be expressed as series in integer or fractional powers of ε, depending on the Keldysh chain (see Gohberg et al. [25]).

Simple Eigenvalue. We assume that at $p=p_0$ the eigenvalue $\lambda_0=i\omega_0\neq 0$ is a simple root of the characteristic Eq. (3) with the eigenvector u_0. According to the perturbation theory of nonselfadjoint operators ([15,16]), in this case the eigenvalues and eigenvectors can be expanded in integer power series of ε

$$\lambda=i\omega_0+\varepsilon\lambda_1+\varepsilon^2\lambda_2+\ldots, \quad u=u_0+\varepsilon w_1+\varepsilon^2 w_2+\ldots. \tag{12}$$

For the first term λ_1 we find ([18])

$$\lambda_1=-\frac{(L_1u_0,u_0)}{(G_0u_0,u_0)+2\lambda_0(M_0u_0,u_0)} \tag{13}$$

where we used the notation

$$\begin{aligned} L_1=\sum_{j=1}^{n}\frac{\partial L}{\partial p_j}e_j&=\sum_{j=1}^{n}\left(\lambda_0^2\frac{\partial M}{\partial p_j}+\lambda_0\frac{\partial G}{\partial p_j}+\frac{\partial K}{\partial p_j}\right)e_j \\ &=\lambda_0^2 M_1+\lambda_0 G_1+K_1. \end{aligned} \tag{14}$$

Formula (13) can be transformed to a more appropriate form if we multiply the numerator and the denominator of (13) by $\lambda_0=i\omega_0$ and use the equality

$$\lambda_0^2(M_0u_0,u_0)+\lambda_0(G_0u_0,u_0)+(K_0u_0,u_0)=0 \tag{15}$$

which follows from (2). Then we get

$$\lambda_1=\frac{-i\omega_0^3(M_1u_0,u_0)-\omega_0^2(G_1u_0,u_0)+i\omega_0(K_1u_0,u_0)}{\omega_0^2(M_0u_0,u_0)+(K_0u_0,u_0)}. \tag{16}$$

Due to symmetry properties of the matrices M, G, and K and Eq. (11) we have $M_1^T=M_1$, $K_1^T=K_1$, $G_1^T=-G_1$. This means that the quantities (M_1u_0,u_0) and (K_1u_0,u_0) are real, and the quantity (G_1u_0,u_0) is purely imaginary. Then from (16) we deduce that λ_1 is purely imaginary (or zero) in accordance with the above conclusion that purely imaginary simple eigenvalues can not leave the imaginary axis. They are differentiable with respect to parameters.

Multiple Eigenvalue: Semi-Simple Case. We assume that at $p=p_0$ the eigenvalue $\lambda_0=i\omega_0$ with algebraic multiplicity r possesses a full number of linear independent eigenvectors $u_1,u_2,\ldots,u_r$. If the other eigenvalues are purely imaginary and simple, then the solution of (1) at $p=p_0$ takes the form

$$x=\sum_{j=1}^{r}\left(c_j u_j e^{i\omega_0 t}+\bar{c}_j\bar{u}_j e^{-i\omega_0 t}\right)+\sum_{j=r+1}^{m}\left(c_j u_j e^{i\omega_j t}+\bar{c}_j\bar{u}_j e^{-i\omega_j t}\right) \tag{17}$$

with c_j being arbitrary complex constants. Expression (17) shows that the motion of the gyroscopic system (1) at $p=p_0$ is stable.

To study the behavior of eigenvalues near the point p_0 we consider a parameter variation $p=p_0+\varepsilon e$. According to the results of the perturbation theory ([15,16]), in the case under study the eigenvalues must be expanded into integer powers of ε

$$\lambda = i\omega_0 + \varepsilon\lambda_1 + \varepsilon^2\lambda_2 + \ldots . \tag{18}$$

The first-order coefficients λ_1 are determined by the Eq. ([18])

$$\det\left[(L_1 u_j, u_k) + \lambda_1\left(\frac{\partial L}{\partial\lambda}u_j, u_k\right)\right] = 0, \quad j,k = 1,2,\ldots,r \tag{19}$$

where for the eigenvectors of the adjoint problem we substituted $v_j = u_j$, $j=1,2,\ldots,r$, and the notation L_1 from (14) is used. The validity of expansion (18) is ensured by the assumption that the matrix $[(L_1 u_j, u_k)]$ is *nonsingular* and the roots λ_1 of equation (19) are all *distinct*, see [16,17]. An example violating this condition was reported ([20]) in the one-parameter case with an expansion $\lambda = \pm i\varepsilon^{3/2} + \ldots$ different from (18). With the above assumption we rule out such nongeneric cases. It should also be mentioned that without the extra assumption of distinct roots λ_1 of (19) the expansion (18) has to be changed into the weaker form

$$\lambda = i\omega_0 + \varepsilon\lambda_1 + o(\varepsilon).$$

Equation (19) can also be written as

$$\det\left[(L_1 u_j, u_k) - i\lambda_1\left(i\frac{\partial L}{\partial\lambda}u_j, u_k\right)\right] = 0, \quad j,k = 1,\ldots,r. \tag{20}$$

The matrices $[(L_1 u_j, u_k)]$ and $[(i\partial L/\partial\lambda u_j, u_k)]$ are Hermitian, since the operators L_1 and $i\partial L/\partial\lambda$ are Hermitian. If one of the two matrices is positive definite, then there exists a basis $u_1^*, u_2^*, \ldots, u_r^*$, in which both matrices are diagonal and real. This means that in this basis we have

$$\begin{aligned}(L_1 u_1^*, u_1^*) - i\lambda_1^{(1)}(i\partial L/\partial\lambda u_1^*, u_1^*) \\ = 0, \ldots, (L_1 u_r^*, u_r^*) - i\lambda_1^{(r)}(i\partial L/\partial\lambda u_r^*, u_r^*) = 0.\end{aligned} \tag{21}$$

If a Hermitian operator H is positive definite, then the so-called Gram matrix $[(Hu_j, u_k)], j,k=1,\ldots,r$ is positive definite for linear independent vectors $u_1, \ldots, u_r$ ([26]). In the case when $i\partial L/\partial\lambda$ is definite, we can solve (21) and get purely imaginary values $\lambda_1^{(1)}, \ldots, \lambda_1^{(r)}$ all different from zero since $[(L_1 u_j, u_k)]$ was assumed nonsingular. This means splitting (bifurcation) of the semi-simple eigenvalue λ_0 along the imaginary axis. Thus, we obtain a simple sufficient criterion for stability:

If the operator $i\partial L/\partial\lambda$ is positive (or negative) definite, then the multiple eigenvalue $\lambda_0 = i\omega_0$ splits into r purely imaginary eigenvalues for any vector of variation e in the parameter space, ensuring stability of the gyroscopic system (1) near the initial point p_0 (except for directions e for which some roots λ_1 are equal to each other).

However, in the general case when the matrix $i\partial L/\partial\lambda$ is not definite, the stability of the gyroscopic system can be lost when we consider small changes of parameters. We will study such a possibility thoroughly for the case of a double eigenvalue ($r=2$).

Double Eigenvalue: Semi-Simple Case. We consider a double eigenvalue $\lambda_0 = i\omega_0 \neq 0$ with two linearly independent eigenvectors u_1 and u_2. For the first-order coefficient λ_1 we have the quadratic Eq. (20). We assume that the eigenvectors satisfy the orthogonality condition

$$(i\partial L/\partial\lambda u_1, u_2) = 0. \tag{22}$$

This diagonalization can always be done since the matrix $[(i\partial L/\partial\lambda u_j, u_k)]$ is Hermitian.

For convenience of the analysis we introduce the following notation:

$$\begin{aligned}(L_1 u_1, u_1) = a_1, \quad (L_1 u_2, u_2) = a_2, \quad (L_1 u_1, u_2) = a_{12}, \\ (i\partial L/\partial\lambda u_1, u_1) = b_1, \quad (i\partial L/\partial\lambda u_2, u_2) = b_2.\end{aligned} \tag{23}$$

The quantities a_1, a_2, b_1, b_2 are real numbers while a_{12} is complex. With this notation we can write Eq. (20) in the form

$$\det\begin{bmatrix} a_1 - i\lambda_1 b_1 & a_{12} \\ \overline{a}_{12} & a_2 - i\lambda_1 b_2 \end{bmatrix} = 0 \tag{24}$$

and we find

$$\lambda_1 = \frac{-i(a_1 b_2 + a_2 b_1) \pm \sqrt{-(a_1 b_2 - a_2 b_1)^2 - 4 b_1 b_2 |a_{12}|^2}}{2 b_1 b_2}. \tag{25}$$

if we assume $b_1 b_2 \neq 0$.

Let us first consider the case, when b_1 and b_2 have the same sign, $b_1 b_2 > 0$, i.e., the matrix $[(i\partial L/\partial\lambda u_j, u_k)]$ is definite. Then it follows immediately that the solutions (25) yield stable splitting of the double eigenvalue $\lambda_0 = i\omega_0$ along the imaginary axis for arbitrary variation e in the parameter space. This case was dealt with in the previous subsection.

The case $b_1 b_2 < 0$ leads to an unstable bifurcation of the double eigenvalue $\lambda_0 = i\omega_0$ into two eigenvalues with nonzero real parts, if the discriminant of the quadratic Eq. (24) is positive, i.e., if

$$-4 b_1 b_2 |a_{12}|^2 > (a_1 b_2 - a_2 b_1)^2. \tag{26}$$

Now we introduce real vectors f_1, f_2 and $f_3 \in R^n$ with the coordinates

$$\begin{aligned}f_1^j + i f_2^j = 2\sqrt{-b_1 b_2}(\partial L/\partial p_j u_1, u_2), \\ f_3^j = b_2(\partial L/\partial p_j u_1, u_1) - b_1(\partial L/\partial p_j u_2, u_2).\end{aligned} \tag{27}$$

Using (14), (23), and (27) we rewrite (26) as

$$\langle f_1, e\rangle^2 + \langle f_2, e\rangle^2 > \langle f_3, e\rangle^2, \tag{28}$$

where the inner product $\langle\cdot,\cdot\rangle$ of vectors in R^n is defined in the usual way.

Note that the vectors f_1, f_2, and f_3 depend only on the information at the initial point, and do not depend on the vector of variation e.

If e satisfies inequality (28), it belongs to the instability domain, but if the left-hand side of (28) is less than the right-hand side, then e belongs to the stability domain. The boundary surface between the stability and instability domains is determined by the equality

$$\langle f_1, e\rangle^2 + \langle f_2, e\rangle^2 = \langle f_3, e\rangle^2. \tag{29}$$

This is a conical surface in the parameter space R^n, dividing the stability and instability domains. But on this surface expansions (18) are not guaranteed since on the surface the two roots λ_1 of Eq. (25) are equal.

Consider the two-dimensional parameter space $p=(p_1, p_2)$ and a vector of variation $e=(\cos\alpha, \sin\alpha)$, see Fig. 1. Then we get for the stability domain the inequality

$$\begin{aligned}(f_1^1\cos\alpha + f_1^2\sin\alpha)^2 + (f_2^1\cos\alpha + f_2^2\sin\alpha)^2 \\ \leqslant (f_3^1\cos\alpha + f_3^2\sin\alpha)^2.\end{aligned} \tag{30}$$

Introducing the coefficients

$$\begin{aligned}A = (f_1^1)^2 + (f_2^1)^2 - (f_3^1)^2, \quad B = f_1^1 f_1^1 + f_2^1 f_2^2 - f_3^1 f_3^2, \\ C = (f_1^2)^2 + (f_2^2)^2 - (f_3^2)^2, \quad D = B^2 - AC,\end{aligned} \tag{31}$$

inequality (30) yields the quadratic inequality

$$C\tan^2(\alpha) + 2B\tan(\alpha) + A \leqslant 0. \tag{32}$$

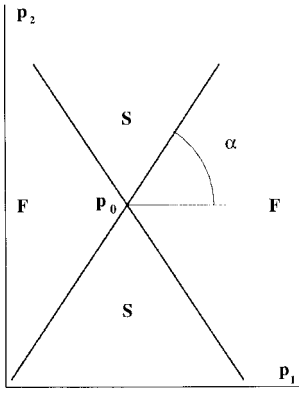


Fig. 1 Stability (S) and flutter (F) domains in two-dimensional parameter space in the vicinity of the initial point p_0

For example, if $C>0$, $D>0$, then the stability domain is given by

$$(-B-\sqrt{D})/C \leqslant \tan\alpha \leqslant (-B+\sqrt{D})/C. \tag{33}$$

The stability domain (S) in the vicinity of the initial point p_0 for Cases (1) and (3) is shown qualitatively in Fig. 1. These and other possibilities for the stability domain are depicted in Table 1.

Double Eigenvalue With a Single Eigenvector. This case is also described in the literature as non-semisimple eigenvalues, or strong interaction. We assume that at $p=p_0$ the eigenvalue $\lambda_0=i\omega_0$ is a double root of the characteristic Eq. (3), corresponding to a Keldysh chain of length 2 ([25]):

$$Lu_0=0, \tag{34}$$

$$Lw_0=-(\partial L/d\lambda)u_0. \tag{35}$$

It means that there exists only one eigenvector u_0, and w_0 is a generalized (or associated) eigenvector; the partial derivative of L is taken at λ_0. A solution of (35) exists only if the right-hand side of (35) is orthogonal to the solution of the adjoint eigenvalue problem (6), which is $v_0=u_0$ according to (9). Thus we have

$$(\partial L/\partial\lambda u_0,u_0)=0. \tag{36}$$

Assuming that the other eigenvalues λ_j are purely imaginary and simple, a solution of (1) at $p=p_0$ takes the form ([25]):

$$x=c_1u_0e^{i\omega_0t}+\overline{c}_1\overline{u}_0e^{-i\omega_0t}+c_2(u_0t+w_0)e^{i\omega_0t}$$
$$+\overline{c}_2(\overline{u}_0t+\overline{w}_0)e^{-i\omega_0t}+\sum_{j=3}^{m}(c_ju_je^{i\omega_jt}+\overline{c}_j\overline{u}_je^{-i\omega_jt}). \tag{37}$$

Here $c_j, j=1,\ldots,m$ are arbitrary complex constants. The solution (37) is unstable (onset of instability) due to the presence of the secular term $te^{i\omega_0t}$.

With the intention to study the behavior of eigenvalues near the point p_0 we consider a variation $p=p_0+\varepsilon e$. According to the perturbation theory ([15,16]) in this particular case the expansions of eigenvalues and eigenvectors contain fractional powers of ε,

Table 1 The stability domain for different combinations of parameters

(1)	$C>0$, $D>0$	$(-B-\sqrt{D})/C \leqslant \tan\alpha \leqslant (-B+\sqrt{D})/C$
(2)	$C>0$, $D<0$	Isolated stability point
(3)	$C<0$, $D>0$	$\tan\alpha \leqslant (-B+\sqrt{D})/C$, $\tan\alpha \geqslant (-B-\sqrt{D})/C$
(4)	$C<0$, $D<0$	No instability domain

$$\lambda=\lambda_0+\varepsilon^{1/2}\lambda_1+\varepsilon\lambda_2+\ldots, \quad u=u_0+\varepsilon^{1/2}w_1+\varepsilon w_2+\ldots. \tag{38}$$

For the first coefficient λ_1 we find ([18])

$$\lambda_1^2=-\frac{(L_1u_0,u_0)}{(\partial L/\partial\lambda w_0,u_0)+(M_0u_0,u_0)}, \tag{39}$$

where notation (14) has been used. For the eigenvalue of the adjoint problem we have substituted $v_0=u_0$, see (9). It is easy to prove that the right-hand side of (39) is a real number. First we estimate the term

$$\begin{aligned}(\partial L/\partial\lambda w_0,u_0)&=([G_0+2\lambda_0M_0]w_0,u_0)\\&=(w_0,[G_0+2i\omega_0M_0]^*u_0)\\&=(w_0,[-G_0-2\lambda_0M_0]u_0)=(w_0,-\partial L/\partial\lambda u_0)\\&=(w_0,Lw_0)=(Lw_0,w_0).\end{aligned} \tag{40}$$

The chain of equalities (40) holds due to (35) and the Hermitian property $L^*=L$.

Note that the validity and convergence of expansion (38) is ensured by assumption that the right-hand side of (39) is not zero, which means $(L_1u_0,u_0)\neq 0$, see [15–16].

For the one-parameter case this condition also appears in [22] as sufficient for instability in a one-side vicinity of p_0. With (40) and (14) equality (39) becomes

$$\lambda_1^2=-\frac{\sum_{j=1}^{n}(\partial L/\partial p_ju_0,u_0)e_j}{(Lw_0,w_0)+(M_0u_0,u_0)}. \tag{41}$$

The operators L, $\partial L/\partial p_j$, and M_0 are Hermitian. Hence the right-hand side of (41) is a real number.

For convenience we introduce a real vector $h\in R^n$ with the coordinates

$$h_j=-\frac{(\partial L/\partial p_ju_0,u_0)}{(Lw_0,w_0)+(M_0u_0,u_0)}. \tag{42}$$

The vector h depends only on the information at the initial point p_0 and does not depend on the vector of variation e. Then (41) is rewritten as

$$\lambda_1^2=\langle h,e\rangle. \tag{43}$$

Eventually, from (38) we end up with

$$\lambda=i\omega_0\pm\sqrt{\varepsilon\langle h,e\rangle}+O(\varepsilon). \tag{44}$$

Equation (44) clearly shows nondifferentiability of a double eigenvalue with respect to parameters. For all variations e satisfying the inequality $\langle h,e\rangle>0$ the double eigenvalue bifurcates into two simple, one of them having a positive real part. If we assume the remaining eigenvalues simple and purely imaginary this means onset of flutter ($\omega_0\neq 0$) or divergence ($\omega_0=0$) instability of the gyroscopic system. And if we take variations e satisfying the inequality $\langle h,e\rangle<0$, the double eigenvalue splits into two purely imaginary eigenvalues which means stability of the system. This shows that the point p_0 belongs to the boundary surface between the stability and instability domains in the parameter space, and the vector h is the normal to the boundary and lies in the instability domain, see Fig. 2.

Note that expansions (38) and (44) become invalid if $\lambda_1^2=\langle h,e\rangle=0$, i.e., when the vector of variation e belongs to the tangent plane of the stability boundary.

Double Zero Eigenvalue. (1) First we consider a double semi-simple eigenvalue $\lambda_0=0$ with two real eigenvectors u_1 and u_2. Using (11) and (14) in (19) and taking there $\lambda_0=0$ we obtain the quadratic equation for the first term coefficient λ_1

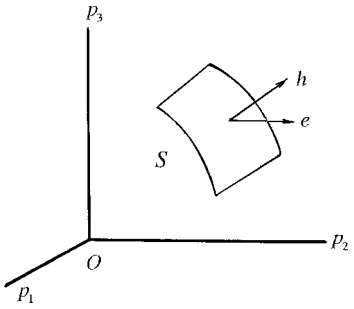


Fig. 2 Normal vector *h* and vector of variation *e* at a point of the stability boundary

$$\det\begin{bmatrix} (K_1u_1,u_1) & (K_1u_1,u_2)+\lambda_1(G_0u_1,u_2) \\ (K_1u_2,u_1)+\lambda_1(G_0u_2,u_1) & (K_1u_2,u_2) \end{bmatrix}=0. \tag{45}$$

The terms (G_0u_j,u_j), $j=1,2$ disappear because the eigenvectors u_1 and u_2 are real and $G_0^T=-G_0$. Besides, we have

$$(K_1u_2,u_1)=(K_1u_1,u_2), \quad (G_0u_2,u_1)=-(G_0u_1,u_2) \tag{46}$$

due to the symmetry properties of K_1 and G_0. Therefore (45) results in

$$\lambda_1^2(G_0u_1,u_2)^2+(K_1u_1,u_1)(K_1u_2,u_2)-(K_1u_1,u_2)^2=0. \tag{47}$$

We introduce real vectors k_1, k_2, and k_{12} with the coordinates

$$k_1^j=(\partial K/\partial p_j u_1,u_1), \quad k_2^j=(\partial K/\partial p_j u_2,u_2),$$
$$k_{12}^j=(\partial K/\partial p_j u_1,u_2), \quad j=1,\ldots,n \tag{48}$$

and assume the normality condition

$$(G_0u_1,u_2)=1. \tag{49}$$

Then using (11) and substituting (48) and (49) into (47) we get

$$\lambda_1=\pm\sqrt{\langle k_{12},e\rangle^2-\langle k_1,e\rangle\langle k_2,e\rangle}. \tag{50}$$

If the vector of variation e satisfies the inequality

$$\langle k_{12},e\rangle^2-\langle k_1,e\rangle\langle k_2,e\rangle<0 \tag{51}$$

the system is stable, and if e satisfies (51) with the opposite sign $>$, the gyroscopic system becomes unstable (divergence). The equality sign in (51) corresponds to the boundary between stability and divergence domains.

(2) Consider now the case of a double eigenvalue $\lambda_0=0$ with a single eigenvector, i.e., a Keldysh chain of length 2. Substituting (7) into (34) and (35) with $\lambda_0=0$ we obtain

$$K_0u_0=0, \quad K_0w_0=-G_0u_0. \tag{52}$$

Since u_0 is a real vector and $G_0^T=-G_0$, the orthogonality condition (36) is always satisfied, $(G_0u_0,u_0)=0$. Then Eqs. (39) and (41) imply

$$\lambda_1^2=-\frac{(K_1u_0,u_0)}{(K_0w_0,w_0)+(M_0u_0,u_0)} \tag{53}$$

and the coordinates of the normal vector h to the stability boundary according to (42) take the form

$$h_j=-\frac{(\partial K/\partial p_j u_0,u_0)}{(K_0w_0,w_0)+(M_0u_0,u_0)}, \quad j=1,\ldots,n. \tag{54}$$

Now bifurcation of eigenvalues is given by (44) with $\omega_0=0$. The stability condition is $\langle h,e\rangle<0$, and the instability (divergence) condition is $\langle h,e\rangle>0$. It is easy to see that the stability condition applied to 2×2 matrices with one parameter agrees with that of [21].

Example 1. Consider a simple model of a massless elastic shaft rotating with a constant angular velocity and carrying a disk with mass. The shaft is subjected to a constant axial compression force. In nondimensional variables with respect to a rotating reference system this rotor is described by a gyroscopic system (1) with the system matrices, see [13]

$$M=\begin{bmatrix}1&0\\0&1\end{bmatrix}, \quad G=2\omega\begin{bmatrix}0&-1\\1&0\end{bmatrix},$$
$$K=\begin{bmatrix}c_1-\eta-\omega^2&0\\0&c_2-\eta-\omega^2\end{bmatrix}. \tag{55}$$

Here c_1 and c_2 are the elastic rigidities of the shaft in two principal directions, η represents the compression force, and ω is the angular velocity of the shaft rotation. A stability analysis of this system depending on four parameters c_1, c_2, η and ω was presented in [8].

Now we want to study stability of the system (1), (55) with respect to the parameters c_1 and c_2, taking fixed values of the other two parameters $\eta=3$ and $\omega=2$.

Consider the point $(c_1,c_2)=(7,7)$. According to (2), (3) at this point we have a double eigenvalue $\lambda_0=0$ with two linear independent eigenvectors

$$u_1=(\alpha,0)^T, \quad u_2=(0,\beta)^T, \quad \alpha\neq0, \quad \beta\neq0. \tag{56}$$

This is the semi-simple case. Using the normality condition (49) and (55), we find $2\omega\alpha\beta=1$. As $\omega=2$ we can take $\alpha=1$, $\beta=1/4$. According to (18), (50) we have

$$\lambda=\pm\varepsilon\sqrt{\langle k_{12},e\rangle^2-\langle k_1,e\rangle\langle k_2,e\rangle}+O(\varepsilon^2) \tag{57}$$

where e is a vector in the c_1, c_2-plane. Calculating the vectors k_1, k_2, k_{12} with the use of (48), (55), and (56) we obtain

$$k_1=(1,0)^T, \quad k_2=(0,1/16)^T, \quad k_{12}=(0,0)^T. \tag{58}$$

Following (51), (57), and (58) we find that the stability domain in the vicinity of the point $(c_1,c_2)=(7,7)$ is given by the inequality

$$\langle k_1,e\rangle\langle k_2,e\rangle>0. \tag{59}$$

If we take both possibilities $\langle k_1,e\rangle>0$, $\langle k_2,e\rangle>0$ and $\langle k_1,e\rangle<0$, $\langle k_2,e\rangle<0$ we get the result that the vector e belonging to the

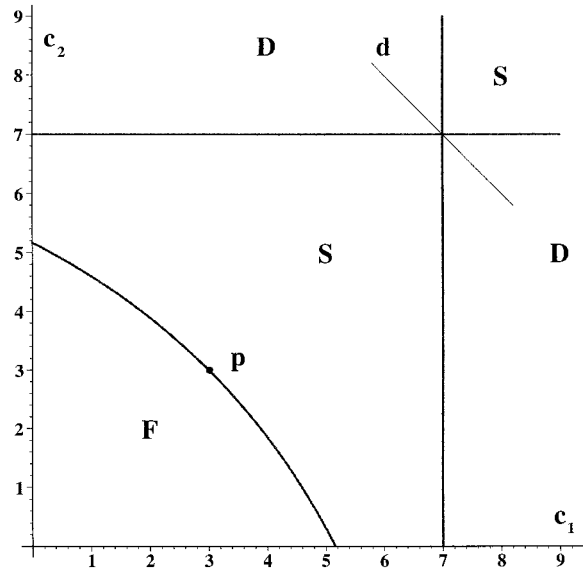


Fig. 3 Stability and instability domains in the two-dimensional parameter space of the rotating shaft with $\omega=2$ and $\eta=3$. S denotes stability, D-divergence, and F-flutter.

stability domain (S) lies in the first and third quadrants, see Fig. 3, and for e in the two other quadrants we have divergence instability (D).

For comparing with the one-parameter case of ([21,22]) we choose $c_1=p$ and $c_2=14-p$ which corresponds to line d in Fig. 3. Then the stability condition presented in these papers is not satisfied since the matrix $[(L_1u_j,u_k)]=\left[\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right]$ is indefinite, and nothing can be said about stability or instability. But according to our theory this case means divergence instability.

Another situation concerns points (c_1,c_2,η,ω) in the four-parameter space satisfying the equation

$$(c_1-c_2)^2+8\omega^2(c_1+c_2-2\eta)=0 \tag{60}$$

which, according to [8], is a boundary between stability and flutter. The eigenvalues on this boundary are double and equal to

$$\lambda=\pm i\sqrt{(c_1+c_2)/2-\eta+\omega^2}. \tag{61}$$

The eigenvalue $\lambda_0=i\sqrt{(c_1+c_2)/2-\eta+\omega^2}$ possesses a single eigenvector u_0 and, according to (35), a generalized (associated) eigenvector w_0 given by

$$u_0=(i2\omega\sqrt{2(c_1+c_2-2\eta+2\omega^2)},\quad c_1-c_2-4\omega^2)^T,$$
$$w_0=(0,\ (-3c_1-c_2+4\eta)/\lambda_0)^T. \tag{62}$$

Then vector h from (42) has the coordinates

$$h_1=-4\omega^2(c_1+c_2+2\omega^2-2\eta)/d,\quad h_2=8\omega^2(c_1-\eta-\omega^2)/d,$$
$$h_3=-4\omega^2(c_1-c_2-4\omega^2)/d,$$
$$h_4=\omega(6c_1^2-2c_2^2-4c_1c_2-8\eta c_1+8\eta c_2)/d,$$
$$d=-3c_1^2+c_2^2+2c_1c_2+4\eta c_1-4\eta c_2+4\omega^2(c_2-c_1+4\omega^2). \tag{63}$$

If we choose a constant angular velocity $\omega=2$, the stability boundary (60) becomes the paraboloid $(c_1-c_2)^2+32(c_1+c_2-2\eta)=0$. Consider, e.g., the vicinity of the point $(c_1,c_2,\eta)=(3,3,3)$ on the paraboloid. For this point the vector h from (42) has according to (63) the coordinates $(h_1,h_2,h_3)^T=(-1/2,-1/2,1)^T$, h is the normal to the paraboloid and lies in the area of flutter instability. If we choose $\eta=3$ constant as well, (60) becomes a parabola in the (c_1,c_2) parameter space with the normal vector $(h_1,h_2)^T=(-1/2,-1/2)^T$ at the point $p=(c_1,c_2)=(3,3)$. Choosing the vector $e=(\cos\alpha,\sin\alpha)^T$, expressions (43) and (44) result in

$$\lambda_1^2=\langle h,e\rangle=-(\cos\alpha+\sin\alpha)/2=-\sin(\alpha+\pi/4)/\sqrt{2},$$
$$\lambda=i\sqrt{(c_1+c_2+2)/2}\pm i\sqrt{\varepsilon\sin(\alpha+\pi/4)/\sqrt{2}}+O(\varepsilon). \tag{64}$$

In the vicinity of the point $(c_1,c_2)=(3,3)$ we get therefore stability for all chosen vectors e satisfying

$$\langle h,e\rangle<0\Leftrightarrow-\pi/4<\alpha<3\pi/4 \tag{65}$$

and flutter instability for

$$\langle h,e\rangle>0\Leftrightarrow3\pi/4<\alpha<7\pi/4 \tag{66}$$

in agreement with the stability map in [8], see Fig. 3.

For $\alpha=-\pi/4$ and $\alpha=3\pi/4$ we get $e=\pm(\sqrt{2}/2,-\sqrt{2}/2)$ which is tangential to the stability boundary, and λ_1 is zero. In this case only an investigation of higher order coefficients $\lambda_2,\lambda_3,\ldots$ in expansion (38) can give an answer.

Example 2. Consider an axially moving beam modeling band saws, belts, magnetic tapes, etc., see [27,23]. The nondimensional equation of motion for free response is

$$y_{tt}+2\nu y_{xt}+y_{xxxx}+(\nu^2-\kappa^2)y_{xx}=0 \tag{67}$$

where ν and κ represent transport speed and axial tension. We assume simply supported boundary conditions.

With $y=e^{\lambda t}u(x)$ we obtain the eigenvalue problem

$$\lambda^2u+2\lambda\nu u_x+u_{xxxx}+(\nu^2-\kappa^2)u_{xx}=0,$$
$$u(0)=u_{xx}(0)=u(1)=u_{xx}(1)=0. \tag{68}$$

The corresponding differential operators in (1), (2), and (7) are

$$M=1,\quad G=2\nu\partial/\partial x,\quad K=\partial^4/\partial x^4+(\nu^2-\kappa^2)\partial^2/\partial x^2,$$
$$L=\lambda^2+2\lambda\nu\partial/\partial x+\partial^4/\partial x^4+(\nu^2-\kappa^2)\partial^2/\partial x^2 \tag{69}$$

and depend on the two parameters ν and κ.

Investigating divergence instability we consider the double eigenvalue $\lambda_0=0$ in Eq. (68) and obtain the critical speeds ν_c with corresponding eigenfunctions u_0, see [23]

$$\nu_c^2=\kappa^2+n^2\pi^2,\quad u_0=\sin n\pi x,\quad n=1,2,\ldots. \tag{70}$$

The divergence boundaries are shown in Fig. 4.

To find the generalized eigenfunction w_0 we write the eigenvalue problem (52) in the form

$$\partial^4w_0/\partial x^4+n^2\pi^2\partial^2w_0/\partial x^2=-2\nu_c\partial u_0/\partial x,$$
$$w_0(0)=(\partial^2w_0/\partial x^2)_{x=0}=0,\quad w_0(1)=(\partial^2w_0/\partial x^2)_{x=1}=0. \tag{71}$$

Using (70) leads to the general solution

$$w_0=(cn\pi/2)x\sin n\pi x-c+c(1+(-1)^{n+1})x+c\cos n\pi x$$
$$+\gamma\sin n\pi x,\quad n=1,2,\ldots \tag{72}$$

where $c=2\nu_c/(n\pi)^3$ and γ is an arbitrary constant.

According to (44), (54) we can find the vector h describing splitting of the double eigenvalue $\lambda_0=0$ near the divergence boundary. The operators in (54) become

$$\partial K/\partial\kappa=-2\kappa\partial^2/\partial x^2,\quad \partial K/\partial\nu=2\nu_c\partial^2/\partial x^2. \tag{73}$$

Computing the inner products with the use of (70), (72) we obtain

$$(\partial K/\partial\kappa u_0,u_0)=-2\kappa\int_0^1\partial^2u_0/\partial x^2u_0dx=\kappa n^2\pi^2,$$
$$(\partial K/\partial\nu u_0,u_0)=2\nu_c\int_0^1\partial^2u_0/\partial x^2u_0dx=-\nu_cn^2\pi^2,$$
$$(Kw_0,w_0)=-(Gu_0,w_0)$$
$$=-\nu_c^2(3n^2\pi^2-16(1+(-1)^{n+1}))/(2n^4\pi^4),$$
$$(Mu_0,u_0)=1/2. \tag{74}$$

Finally, we find the vector h which is normal to the divergence boundaries as

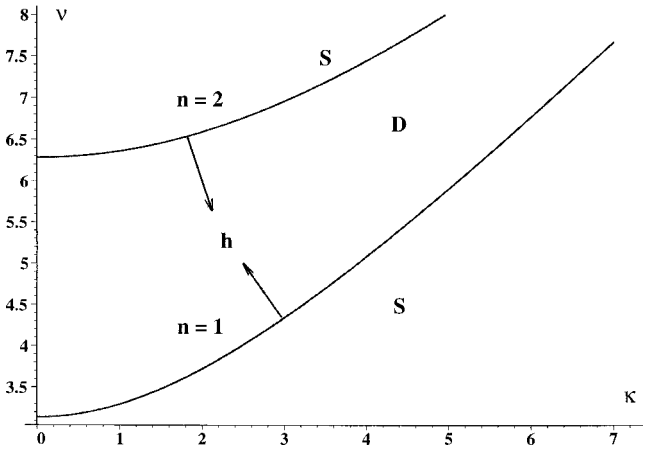


Fig. 4 Stability (S) and divergence (D) domains of the axially moving beam

$$h = \frac{2n^6\pi^6}{n^4\pi^4 - \nu_c^2(3n^2\pi^2 - 16(1+(-1)^{n+1}))}(-\kappa, \nu_c). \quad (75)$$

Here the factor of the vector $(-\kappa, v_c)$ is positive for $n=1$ but negative for $n=2,3,\ldots$.

Since the vector h lies in the divergence domain, we have the following result (see also Fig. 4). The eigenvalue λ is stable in the vicinity below the first critical speed $(n=1)$ and unstable above. The opposite situation occurs for all the other critical speeds $(n=2,3,\ldots)$ where the eigenvalue λ is unstable below the critical speed and stable above. This is in agreement with the result of Parker [23], who investigated the one-parameter problem with two first critical speeds.

The splitting of the zero eigenvalue is according to (44) given by

$$\lambda = \pm\sqrt{\varepsilon\langle h, e\rangle} + O(\varepsilon). \quad (76)$$

It should be noted that we only have considered the stability-divergence boundaries. For fixed κ and sufficiently large ν above the second divergence boundary $(n=2)$, flutter instability occurs, as shown numerically in [27].

Conclusion

We have studied the bifurcations of eigenvalues of gyroscopic systems under a change of several real parameters. These can be stiffness and mass coefficients, geometrical sizes, characteristics of material properties, loads, rotational speeds, etc. Stability of a gyroscopic system can only be lost when eigenvalues meet on the imaginary axis. It follows from this observation that stability or instability (flutter, divergence) depend on the bifurcation of multiple eigenvalues under a change of parameters.

The main distinction and advantage of the present theory in comparison with previous research is that it is a multiparameter study and explicit expressions for the bifurcations of eigenvalues are given, leading to stability as well as to instability conditions. The obtained formulas allow to construct approximations of the stability domain in the parameter space using information only at a point on the boundary. Without the given approach, such a construction is a difficult and time consuming problem in the multiparameter case. The formulas can also be useful for applying gradient methods to stabilize the system.

The results of the developed perturbation theory are essentially similar for systems with finite degrees-of-freedom as well as for distributed systems governed by differential operators. We have demonstrated this by studying two mechanical problems.

References

[1] Thomson, W., and Tait, P. G., 1879, *Treatise on Natural Philosophy*, Vol. I, Part I, Cambridge University Press, Cambridge, U.K.

[2] Hagedorn, P., 1975, ''Über die Instabilität konservativer Systeme mit gyroskopischen Kräften,'' Arch. Ration. Mech. Anal., **58**, No. 1, pp. 1–9.

[3] Huseyin, K., Hagedorn, P., and Teschner, W., 1983, ''On the Stability of Linear Conservative Gyroscopic Systems,'' Z. Angew. Math. Phys., **34**, No. 6, pp. 807–815.

[4] Walker, J. A., 1991, ''Stability of Linear Conservative Gyroscopic Systems,'' ASME J. Appl. Mech., **58**, pp. 229–232.

[5] Barkwell, L., and Lancaster, P., 1992, ''Overdamped and Gyroscopic Vibrating Systems,'' ASME J. Appl. Mech., **59**, pp. 176–181.

[6] Seyranian, A. P., 1993, ''Interaction of Vibrational Frequencies of a Gyroscopic System,'' Mech. Solids, **28**, No. 4, pp. 33–41.

[7] Veselic, K., 1995, ''On the Stability of Rotating Systems,'' Z. Angew. Math. Mech., **75**, No. 4, pp. 325–328.

[8] Seyranian, A. P., Stoustrup, J., and Kliem, W., 1995, ''On Gyroscopic Stabilization,'' Z. Angew. Math. Phys., **46**, pp. 255–267.

[9] Lancaster, P., and Zizler, P., 1998, ''On the Stability of Gyroscopic Systems,'' ASME J. Appl. Mech., **65**, pp. 519–522.

[10] Kliem, W., and Seyranian, A. P., 1997, ''Analysis of the Gyroscopic Stabilization of a System of Rigid Bodies,'' Z. Angew. Math. Phys., **48**, No. 5, pp. 840–847.

[11] Mailybaev, A. A., and Seyranian, A. P., 1999, ''The Stability Domains of Hamiltonian Systems,'' J. Appl. Math. Mech., **63**, No. 4, pp. 545–555.

[12] Müller, P. C., 1977, *Stabilität und Matrizen*, Springer-Verlag, Berlin.

[13] Huseyin, K., 1978, *Vibrations and Stability of Multiple Parameter Systems*, Noordhoff International Publishing, Alphen aan den Rijn.

[14] Merkin, D. R., 1987, *Introduction to the Theory of Stability of Motion*, Nauka, Moscow.

[15] Vishik, M. I., and Lyusternik, L. A., 1960, ''The Solution of Some Perturbation Problems for Matrices and Selfadjoint or Non-Selfadjoint Differential Equations I,'' Russ. Math. Surv., **15**, pp. 1–74.

[16] Lidskii, V. B., 1966, ''Perturbation Theory of Non-Conjugate Operators,'' USSR Comput. Math. Math. Phys., **6**, No. 1, pp. 73–85.

[17] Moro, J., Burke, J. V., and Overton, M. L., 1997, ''On the Lidskii-Vishik-Lyusternik Perturbation Theory for Eigenvalues of Matrices With Arbitrary Jordan Structure,'' SIAM J. Matrix Anal. Appl., **18**, No. 4, pp. 793–817.

[18] Seyranian, A. P., 1993, ''Sensitivity Analysis of Multiple Eigenvalues,'' Mech. Struct. Mach., **21**, No. 2, pp. 261–284.

[19] Renshaw, A. A., and Mote, Jr., C. D., 1996, ''Local Stability of Gyroscopic Systems Near Vanishing Eigenvalues,'' ASME J. Appl. Mech., **63**, pp. 116–120.

[20] Lancaster, P., and Kliem, W., 1999, ''Comments on Stability Properties of Conservative Gyroscopic Systems,'' ASME J. Appl. Mech., **66**, pp. 272–273.

[21] Renshaw, A. A., 1998, ''Stability of Gyroscopic Systems Near Vanishing Eigenvalues,'' ASME J. Appl. Mech., **65**, pp. 1062–1064.

[22] Hryniv, R. O., Lancaster, P., and Renshaw, A. A., 1999, ''A Stability Criterion for Parameter Dependent Gyroscopic Systems,'' ASME J. Appl. Mech., **66**, pp. 660–664.

[23] Parker, R. G., 1998, ''On the Eigenvalues and Critical Speed Stability of Gyroscopic Continua,'' ASME J. Appl. Mech., **65**, pp. 134–140.

[24] Ziegler, H., 1968, *Principles of Structural Stability*, Blaisdell, Waltham, MA.

[25] Gohberg, I., Lancaster, P., and Rodman, L., 1982, *Matrix Polynomials*, Academic Press, San Diego, CA.

[26] Horn, R., and Johnson, C. A., 1985, *Matrix Analysis*, Cambridge University Press, Cambridge, UK.

[27] Wickert, J. A., and Mote, Jr., C. D., 1991, ''Response and Discretization Methods for Axially Moving Materials,'' ASME J. Appl. Mech., **44**, pp. S279–S284.

Does a Partial Elastic Foundation Increase the Flutter Velocity of a Pipe Conveying Fluid?

I. Elishakoff
Professor,
Department of Mechanical Engineering,
Florida Atlantic University,
Boca Raton, FL 33431
e-mail: ielishak@me.fau.edu
Mem. ASME

N. Impollonia
Assistant Professor,
Dipartimento di Costruzioni e Tecnologie Avanzate,
Universita di Messina,
Contrada Sperone 98166, Italy

The effect of the elastic Winkler and rotatory foundations on the stability of a pipe conveying fluid is investigated in this paper. Both elastic foundations are partially attached to the pipe. It turns out that the single foundation, either translational or rotatory, which is attached to the pipe along its entire length, increases the critical velocity. Such an intuitively anticipated strengthening effect is surprisingly missing for the elastic column on Winkler foundation subjected to the so-called statically applied follower forces. Yet, partial foundation for the pipe conveying fluid is associated with a nonmonotonous dependence of the critical velocity versus the attachment ratio defined as the length of the partial foundation over the entire length of the pipe. It is concluded that such a paradoxical nonmonotonicity is shared by both the realistic structure (pipe conveying fluid) and in the ''imagined system,'' to use the terminology of Herrmann pertaining to the column under to follower forces. [DOI: 10.1115/1.1354206]

1 Introduction

In 1972, Smith and Herrmann published a study devoted to the stability of the elastic Beck's column under follower force, when the column is attached to the elastic translational Winkler foundation with uniform modulus *k*. They arrived at the paradoxical conclusion that the flutter load does not depend of the translational foundation modulus, irrespective of its magnitude. Due to this surprising conclusion the paper by Smith and Herrmann [1] generated several other studies. These studies differ in their approach to explain the above mathematical result. Sundararajan [2] showed that if the elastic foundation modulus is nonuniform along the column axis then the flutter load depends on the elastic foundation modulus. Other investigators ([3–5]) showed that the noninclusion of the damping terms was the fact responsible for the paradoxical conclusion arrived at by Smith and Herrmann [1]. Note that the Smith and Herrmann conclusion pertaining the Winkler foundation does not hold for the column on rotatory foundation. Becker et al. [6] investigated this problem and concluded that even in the undamped case when the modulus of the rotatory foundation increases the flutter load of the system increases too. Panovko and Gubanova [7] mention, in the fourth edition of their book, that the conclusion by Smith and Herrmann [1] is ''. . . wrong and represents a direct consequence of extremely early transfer to the ideal elastic model. . . . Critical load depends of the stiffness coefficient of the foundation.'' In two publications Koiter [8,9] criticized the very notion of follower forces, since there is no experimental verification (at least until presently) of the existence of the purely statically applied follower forces. He quoted from Herrmann [10]: ''It is a peculiar feature of stability problems of elastic systems subjected to (nonconservative) follower forces that their analysis arose not out of a desire or need to consider a system which presented itself in engineering practice or in the research laboratory, but rather because the *fictitiously applied follower forces* acting on a given system were arbitrarily prescribed to depend in a certain manner on deformation.''

In these circumstances it appears instructive to study the effect of the elastic foundation on the stability of a realistic problem, namely the pipe conveying fluid. This problem is also a nonconservative one, and has been extensively investigated theoretically, experimentally, and numerically. The most recent comprehensive review of research in this area is given by Païdoussis and Li [11] (see also Païdoussis [12]). The effect of elastic foundation on the fluid conveying pipe was investigated in several studies. Becker et al. [13] illustrated the variety of behaviors for different kinds of foundations; Lottati and Kornecki [14] derived numerous results for the case of varying fluid-over-total mass ratio when only the translational Winkler foundation was present.

In this paper we study the effect of the translational and rotatory foundations on the stability of the pipe conveying fluid, with elastic foundation attached to a *part* of the pipe. The analysis includes, as a particular case, the presence of these foundations along the *entire* length of the pipe, and allows the comparison with other studies. We are interested in the effect of elastic foundations on the behavior of realistic problem, rather than ''unexpected behavior of an imagined system,'' in the terminology of Hermann [10] as he characterizes the ''follower forces.''

2 Equation of Motion

Consider a slender cantilever pipe of length *l*, with uniform cross-sectional area *A* and moment of inertia *I*. The material of the pipe is assumed to obey a viscoelastic stress-strain relationship of the Voigt type with E_0 being the elastic Young's modulus, *c* the internal and *b* the external damping coefficient. The mass of the pipe for unit length is denoted by m_p. The pipe that is conveying an incompressible fluid of density ρ with constant velocity *V* rests for a length a $(0 \leqslant a \leqslant l)$ along its axis on an elastic Winkler-type foundation with constant modulus k_1 and damping coefficient d_1 so that the force $f_w(x,t)$ per unit length associated with it is

$$f_w(x,t) = \left(k_1 + d_1 \frac{\partial}{\partial t} \right) U(a-x) w \tag{1}$$

where $U(\cdot)$ is the unit-step function. Additionally, on the same interval, the pipe is supported by a rotary foundation with constant modulus k_2 and damping coefficient d_2 resulting in the restoring distributed moment

$$m(x,t) = \left(k_2 + d_2 \frac{\partial}{\partial t} \right) U(a-x) \frac{\partial w}{\partial x} \tag{2}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 8, 1998; final revision, Aug. 15, 2000. Associate Editor: D. A. Siginer. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

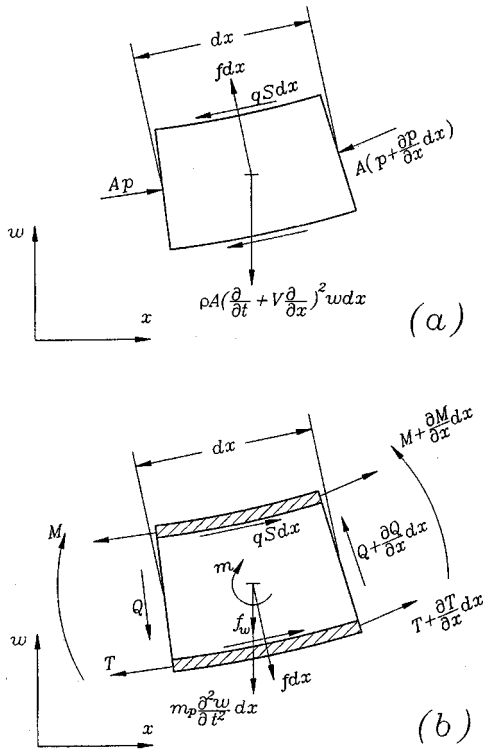


Fig. 1 Forces and moments acting on elements of (*a*) a fluid and (*b*) the pipe

and has the dimension of a force. The basic equation of motion for small transverse vibrations $w(x,t)$ is derivable either from Hamilton's principle ([15]) or through using the dynamic equilibrium approach. The latter way will be adopted in this study, following Gregory and Païdoussis [16] and Blevins [17]. In addition to their derivation, we introduce damping mechanisms and translational and rotary foundations.

Consider an element of a pipe and a fluid as shown in Fig. 1. Let the internal perimeter of the pipe be S. When the fluid flows through the deflecting pipe, it experiences centrifugal acceleration due to the changing curvature of the pipe. The acceleration is opposed by the vertical component of fluid pressure applied to the fluid element and the pressure force f per unit length applied on the fluid element by the pipe walls. A requirement of the balance of forces in two respective directions, namely, along the tangent to the centerline of the deflected element and the perpendicular to it, yields the equations

$$A\frac{\partial p}{\partial x}+qS=0 \tag{3}$$

$$f-pA\frac{\partial^2 w}{\partial x^2}-\rho A\left(\frac{\partial}{\partial t}+V\frac{\partial}{\partial x}\right)^2 w=0 \tag{4}$$

where q is the shear stress on the internal surface of the pipe due to the friction with the fluid. The equations of motion for the pipe in the same directions are similarly found as

$$\frac{\partial T}{\partial x}+qS-Q\frac{\partial^2 w}{\partial x^2}=0 \tag{5}$$

$$\frac{\partial Q}{\partial x}+T\frac{\partial^2 w}{\partial x^2}-f-m_p\frac{\partial^2 w}{\partial t^2}-b\frac{\partial w}{\partial t}-f_w=0 \tag{6}$$

where T is the longitudinal tension in the pipe and Q is the transverse shear force. The relationship between the shear force Q and the bending moment M acting on the section reads

$$Q=m-\frac{\partial M}{\partial x}. \tag{7}$$

The constitutive equation gives

$$M=E_0 I\left(1+c\frac{\partial}{\partial t}\right)\frac{\partial^2 w}{\partial x^2} \tag{8}$$

Equation (6), considering Eqs. (4), (7), and (8) to eliminate forces f and Q, becomes

$$E_0 I\left(1+c\frac{\partial}{\partial t}\right)\frac{\partial^4 w}{\partial x^4}+(pA-T)\frac{\partial^2 w}{\partial x^2}+\rho A\left(\frac{\partial}{\partial t}+V\frac{\partial}{\partial x}\right)^2 w$$
$$+m_p\frac{\partial^2 w}{\partial t^2}+b\frac{\partial w}{\partial t}+f_w-\frac{\partial m}{\partial x}=0. \tag{9}$$

By neglecting the third term in the left-hand side of Eq. (5), as being of the second order in w, the shear stress q is eliminated from Eq. (3) and Eq. (5) to result in

$$\frac{\partial(pA-T)}{\partial x}=0. \tag{10}$$

We assume that at the free end of the pipe the tension is zero and that the fluid pressure is equal to the ambient pressure, $p=T=0$ at $x=l$. Consequently,

$$pA-T=0 \tag{11}$$

everywhere. Thus Eq. (9), after substituting expressions for f_w and m, becomes

$$EI\left(1+c\frac{\partial}{\partial t}\right)\frac{\partial^4 w}{\partial x^4}+\rho A\left(\frac{\partial}{\partial t}+V\frac{\partial}{\partial x}\right)^2 w+m_p\frac{\partial^2 w}{\partial t^2}+b\frac{\partial w}{\partial t}$$
$$+\left(k_1+d_1\frac{\partial}{\partial t}\right)U(a-x)w-\left(k_2+d_2\frac{\partial}{\partial t}\right)U(a-x)\frac{\partial^2 w}{\partial x^2}=0. \tag{12}$$

The boundary conditions that apply to the cantilever pipe read

$$w=0 \quad \text{and} \quad \frac{\partial w}{\partial x}=0, \quad \text{at } x=0$$
$$M=EI\left(1+c\frac{\partial}{\partial t}\right)\frac{\partial^2 w}{\partial x^2}=0, \quad \text{at } x=l. \tag{13}$$

In case of the *partial* foundation in the interval $0\leqslant x\leqslant a<l$, the fourth boundary condition reads

$$Q=EI\left(1+c\frac{\partial}{\partial t}\right)\frac{\partial^3 w}{\partial x^3}=0, \quad \text{at } x=l. \tag{14}$$

When the foundation is attached along the entire length of the beam the fourth boundary condition is more complicated,

$$Q=EI\left(1+c\frac{\partial}{\partial t}\right)\frac{\partial^3 w}{\partial x^3}-\left(k_2+d_2\frac{\partial}{\partial t}\right)\frac{\partial^2 w}{\partial x^2}=0 \quad \text{at } x=l. \tag{15}$$

It is remarkable that in the latter case the rotatory foundation term k_2 appears in the boundary condition, since the expression for the shear force in Eq. (7) contains the term $m(x,t)$, responsible for the effect of the rotatory foundation. We seek solution of Eq. (12) in the form

$$w(x,t)=e^{\Omega t}z(x) \tag{16}$$

The following dimensionless quantities are introduced:

$$\xi=\frac{x}{L}; \quad m_f=\rho A; \quad \omega=\Omega\sqrt{\frac{m_f+m_p}{E_0 I}}\,l^2;$$

$$\mu=\sqrt{\frac{m_f}{m_f+m_p}};\quad \alpha=\frac{a}{l}$$

$$\nu=Vl\sqrt{\frac{m_f}{E_0I}};\quad \chi_1=\frac{k_1l^4}{E_0I};\quad \chi_2=\frac{k_2l^2}{E_0I};\quad \beta=\frac{bl^2}{\sqrt{E_0I(m_f+m_p)}} \tag{17}$$

$$\gamma=\frac{c}{l^2}\sqrt{\frac{E_0I}{m_f+m_p}};\quad \delta_1=\frac{d_1}{l^2}\sqrt{\frac{E_0I}{m_f+m_p}};\quad \delta_2=\frac{d_2}{l^2}\sqrt{\frac{E_0I}{m_f+m_p}}.$$

Equations (12) and (15) become

$$L(z)\equiv(1+\omega\gamma)z^{IV}+[\nu^2-\chi_2(1+\omega\delta_2)U(\alpha-\xi)]z''+2\omega\nu\mu z' +[\omega\beta+(\chi_1+\omega\delta_1)U(\alpha-\xi)+\omega^2]z=0 \tag{18}$$

$$z(0)=0;\quad z'(0)=0;\quad z''(1)=0 \tag{19}$$

$$z'''(1)=0;\quad (\text{if } \alpha=1) \tag{20}$$

$$(1+\omega\gamma)z'''(1)-(\chi_2+\omega\delta_2)z'(1)=0;\quad (\text{if } \alpha<1) \tag{21}$$

where the operator L is defined in Eq. (18), and the primes indicate derivative with respect to the dimensionless coordinate ξ. The behavior of the nonconservative system governed by the above non-self-adjoint problem is dictated by the value of the real part of the dimensional exponents Ω_j in Eq. (16) or their dimensionless counterparts ω_j in Eq. (17). From Eq. (16) it is evident that if $\mathrm{Re}(\Omega_j)>0$ the transverse deflections are time-wise unbounded, and the flutter phenomenon occurs. On the other hand, negative values of $\mathrm{Re}(\Omega_j)$ lead to decaying (and, hence, stable) oscillations of the pipe. The analysis requires the evaluation of sets of control parameters for which the stability of the system is lost. The value of the fluid velocity ν is taken as a control parameter while all the others are fixed. The minimum value ν for which the pipe undergoes unstable oscillations is referred as the *critical velocity*. In the following the nondimensional critical velocity ν_{cr} is evaluated numerically by both exact and approximate analyses. The approximate analysis may appear to be superfluous at the first glance, yet its results are utilized as initial guesses for the iterative solution scheme required in the exact setting.

3 Exact Solution

Instead of using Eq. (16), we represent the solution in two regions:

$$w(x,t)=\begin{cases} e^{\Omega t}y_1(x); & \text{for } 0\leqslant x<a \\ e^{\Omega t}y_2(x); & \text{for } a\leqslant x<l. \end{cases} \tag{22}$$

Equation (18) becomes

$$L_1(y_1)\equiv(1+\omega\gamma)y_1^{IV}+[\nu^2-(\chi_2+\omega\delta_2)]y_1''+2\omega\nu\mu y_1' +[\omega\beta+(\chi_1+\omega\delta_1)+\omega^2]y_1=0;\ \text{for } 0\leqslant\xi<\alpha \tag{23a}$$

$$L_2(y_2)\equiv(1+\omega\gamma)y_2^{IV}+\nu^2y_2''+2\omega\nu\mu y_2'+(\omega\beta+\omega^2)y_2=0; \quad \text{for } \alpha\leqslant\xi\leqslant 1. \tag{23b}$$

The continuity conditions on the interface between the two regions read

$$\begin{aligned} y_1(\alpha)&=y_2(\alpha)\\ y_1'(\alpha)&=y_2'(\alpha)\\ y_1''(\alpha)&=y_2''(\alpha)\end{aligned} \tag{24}$$

$$(1+\omega\gamma)y_1'''(\alpha)-(\chi_2+\omega\delta_2)z'(1)=(1+\omega\gamma)y_2'''(\alpha).$$

The boundary conditions (19)–(21) are recast in the following manner:

$$y_1(0)=0;\quad y_1'(0)=0;\quad y_2''(1)=0;\quad y_2'''(1)=0. \tag{25}$$

The solutions of the differential Eqs. (23) are given by

$$y_1(\xi)=\sum_{j=1}^{4}C_je^{\lambda_j\xi};\quad y_2(\xi)=\sum_{j=5}^{8}C_je^{\lambda_j\xi} \tag{26}$$

with the conventional modifications in the event that two or more roots coincide. In Eq. (26) λ_j, $j=1,2,\ldots,8$, are the roots of the characteristic equations

$$(1+\omega\gamma)\lambda^4-[\nu^2-(\chi_2+\omega\delta_2)]\lambda^2+2\omega\nu\mu\lambda+\omega\beta +(\chi_1+\omega\delta_1)+\omega^2=0;\quad (j=1,\ldots,4) \tag{27a}$$

$$(1+\omega\gamma)\lambda^4-\nu^2\lambda^2+2\omega\nu\mu\lambda+\omega\beta+\omega^2=0;\quad (j=5,\ldots,8). \tag{27b}$$

Upon introduction of Eq. (26) in the conditions (24), (25), one obtains the system of linear equations in terms of the vector **C** of constants C_j

$$\sum_{j=1}^{4}C_j=0;\ \sum_{j=1}^{4}C_j\lambda_j=0;\ \sum_{j=5}^{8}C_j\lambda_j^2e^{\lambda_j}=0;\ \sum_{j=5}^{8}C_j\lambda_j^3e^{\lambda_j}=0$$

$$\sum_{j=1}^{4}C_je^{\lambda_j\alpha}=\sum_{j=5}^{8}C_je^{\lambda_j\alpha};\quad \sum_{j=1}^{4}C_j\lambda_je^{\lambda_j\alpha}=\sum_{j=5}^{8}C_j\lambda_je^{\lambda_j\alpha} \tag{28}$$

$$\sum_{j=1}^{4}C_j\lambda_j^2e^{\lambda_j\alpha}=\sum_{j=5}^{8}C_j\lambda_j^2e^{\lambda_j\alpha};\quad \sum_{j=1}^{4}(1+\omega\gamma)C_j\lambda_j^3e^{\lambda_j\alpha} -(\chi_2+\omega\delta_2)\lambda_je^{\lambda_j\alpha}=(1+\omega\gamma)\sum_{j=5}^{8}C_j\lambda_j^3e^{\lambda_j\alpha}.$$

These equations can be written in matrix form as

$$\mathbf{EC}=\mathbf{0} \tag{29}$$

where **E** is the matrix with the following elements different form zero:

$$\begin{aligned} &E_{1j}=1,\ (j=1,\ldots,4);\quad E_{2j}=\lambda_j,\ (j=1,\ldots,4);\\ &E_{3j}=\lambda_j^2e^{\lambda j},\ (j=5,\ldots,8);\quad E_{4j}=\lambda_j^3e^{\lambda j},\ (j=5,\ldots,8);\\ &E_{5j}=e^{\lambda_j\alpha},\ (j=1,\ldots,4);\quad E_{5j}=-e^{\lambda_j\alpha},\ (j=5,\ldots,8);\\ &E_{6j}=\lambda_je^{\lambda_j\alpha},\ (j=1,\ldots,4);\quad E_{6j}=-\lambda_je^{\lambda_j\alpha},\ (j=5,\ldots,8);\\ &E_{7j}=\lambda_j^2e^{\lambda_j\alpha},\ (j=1,\ldots,4);\quad E_{7j}=-\lambda_j^2e^{\lambda_j\alpha},\ (j=5,\ldots,8);\\ &E_{8j}=(1+\omega\gamma)\lambda_j^3e^{\lambda_j\alpha}-(\chi_2+\omega\delta_2)\lambda_je^{\lambda_j\alpha},\ (j=1,\ldots,4);\\ &E_{8j}=-(1+\omega\gamma)\lambda_j^3e^{\lambda_j\alpha},\ (j=5,\ldots,8). \end{aligned} \tag{30}$$

In order for a nontrivial solution $\Sigma C_j\neq 0$ of system in Eq. (28) to exist, the determinant of the matrix **E** must vanish. This requirement leads, after some algebra, to a characteristic equation written in the following form:

$$\mathrm{Det}(\mathbf{E})=\Psi(\lambda_1,\lambda_2,\ldots,\lambda_8,\chi_2,\delta_2,\gamma)=0. \tag{31}$$

Substitution of the values of λ_j from Eqs. (27*a*) and (27*b*) into Eq. (31) yields the characteristic equation in terms of ω. At $\nu=0$ the real parts of the eigenvalues ω_k ($k=1,2,\ldots$) are negative. As ν increases their magnitudes vary, and at a certain critical velocity ν_{cr} the real part of at least one of the frequencies changes sign from negative to positive, indicating the onset of instability. At the boundary of instability the real part of this frequency vanishes, and its imaginary part is termed the critical frequency ω_{cr}. The method of solution is based on an optimization procedure that keeps varying the velocity ν and the imaginary part $\mathrm{Im}(\omega)$ of the eigenvalue while $\mathrm{Re}(\omega)=0$, until a minimum is reached for the real function $\mathrm{Abs}(\mathrm{Det}(\mathbf{E}))$. The correspondent value of ν is the desired critical velocity only if the minimum of the latter function

is equal to zero so that Eq. (31) holds; moreover, in order ν to qualify as ν_{cr} there should be no lesser value of ν for which the same condition is verified. Analogous numerical procedure was utilized by Lottati and Kornecki [14].

In order to ensure convergence of the procedure to the exact solution it is advantageous to start the process from a known approximate solution that is sufficiently close to the exact one. To this aim the Galerkin approximation is resorted to in order to furnish a representation for the starting point. As will be demonstrated, the combined use of the optimization procedure in conjunction with the Galerkin method for the initial guess turns out to be an efficient way to arrive at the exact solution.

4 Auxiliary, Approximate Solution for the Initial Guess

According to the Galerkin method the displacement function $z(\xi)$ appearing in Eq. (16) is expanded in series

$$z(\xi)=\sum_{j=1}^{N} H_j z_j(\xi) \tag{32}$$

where H_j are unknown coefficients, $z_j(\xi)$ are comparison functions. The comparison functions $z_j(\xi)$, representing a complete set, read

$$z_j(\xi)=\frac{(j+3)(j^2+3j+2-P)}{j+1}\xi^{j+1}-\frac{2(j+3)(j^2+2j-P)}{j+2}\xi^{j+2}+(j^2+j-P)\xi^{j+3},\quad (j=1,2,\ldots) \tag{33}$$

where

$$P=\begin{cases} -\dfrac{(\chi_2+\omega\delta_2)}{1+\omega\gamma}, & \text{for } \alpha=1 \\ 0, & \text{otherwise.}\end{cases} \tag{34}$$

The coefficients in front of ξ^n $(n=j+1,j+2,j+3)$ are chosen so to be reducible to known Duncan polynomials ([18]) for which the parameter P is identically zero. Substituting Eq. (33) into Eq. (18) and multiplying the result by the comparison function $z_k(\xi)$ and integrating over the entire length of the beam, results in the set of N algebraic equations

$$\sum_{k=1}^{N} A_{jk}H_k=0;\quad j=1,2,\ldots,N \tag{35}$$

where N is the number of retained term. In Eq. (35),

$$A_{jk}=\int_0^1 z_j L(z_k)\,d\xi. \tag{36}$$

To fulfill the requirement of nontriviality $(\Sigma H_k^2\neq 0)$ the determinant of the matrix **A** has to be set equal to zero. This leads to the characteristic equation

$$a_s\omega^s+a_{s-1}\omega^{s-1}+\ldots+a_1\omega+a_0=0. \tag{37}$$

The degree s of the polynomial equals $4N$ in the case of a rotatory foundation attached to the pipe along its *entire* length; it equals $2N$ in the case of *partial* foundation. This is explained as follows: From Eq. (18) we deduce that in case of fully attached rotatory foundation the expression $L(z_k)$ results in ω^3, since z_j also contain ω in the first power, the result is the appearance of ω^4 in A_{jk}. For N-term approximation, consequently, we get terms with ω^{4N}. In case of *partial* foundation P vanishes automatically so that $L(z_k)$ results in ω^2 and z_j does not contain ω. Hence A_{jk} contain the terms ω^2; for N term approximation we get terms containing ω^{2N}. The stability analysis is then reduced to the numerical evaluation of ω from Eq. (37). Since the fluid velocity is the only control parameter, by its gradual increase we evaluate the roots ω of Eq. (37) until the real part of one of them approaches a zero value from below, this implies that the critical velocity is reached. Alternatively, the application of the Hurwitz criterion ([19]) to Eq. (37) yields the sought critical velocity. The values of ν_{cr} so determined are utilized as input data for the exact analysis described in Section 3.

5 Numerical Results

To gain some insight into the behavior of the system, the critical velocity ν_{cr}, as a function of the foundation attachment ratio α, is extensively investigated for numerous cases. For a better understanding on the contribution of each type of foundation on the stability of the pipe, the cases of purely Winkler as well as purely rotatory foundation will be considered first. Finally, the general case of combined presence of both foundations will be addressed.

5.1 Winkler Foundation Alone. Let us investigate the case $\chi_2=0$, so that only the Winkler foundation is present. Figure 2 illustrates the dimensionless critical velocity ν_{cr} for different values of the mass ratio μ and the internal damping coefficient γ. Solid line corresponds to the pipe without internal damping $(\gamma=0)$; the dotted line indicates the case $\gamma=0.001$; the dash-dotted line is associated with $\gamma=0.005$, whereas the dashed line depicts the case $\gamma=0.01$. Inspection shows that the presence of a full foundation has a stabilizing effect, in the sense that the critical velocity associated with a full foundation $(\alpha=1)$ is greater than that associated with a pipe without an elastic foundation $(\alpha=0)$, although this feature is more pronounced for $\mu\geq 0.3$. Specifically, in the case $\gamma=0.001$, for $\mu=0.1$, critical velocity of the pipe without foundation equals 4.362, whereas its counterpart for the full foundation equals 4.420, constituting an increment of 1.33 percent. For $\mu=0.3$ the corresponding values are 4.749 and 5.538, respectively, resulting in 16.61 percent increase. For $\mu=0.5$ and $\mu=0.7$, respectively, the increases are 26.71 percent and 7.93 percent. Investigation of the effect of partial foundation, which apparently was not conducted prior to this study, for the pipe conveying fluid, reveals some interesting effects. It turns out that, surprisingly, the critical velocity does not always vary monotoni-

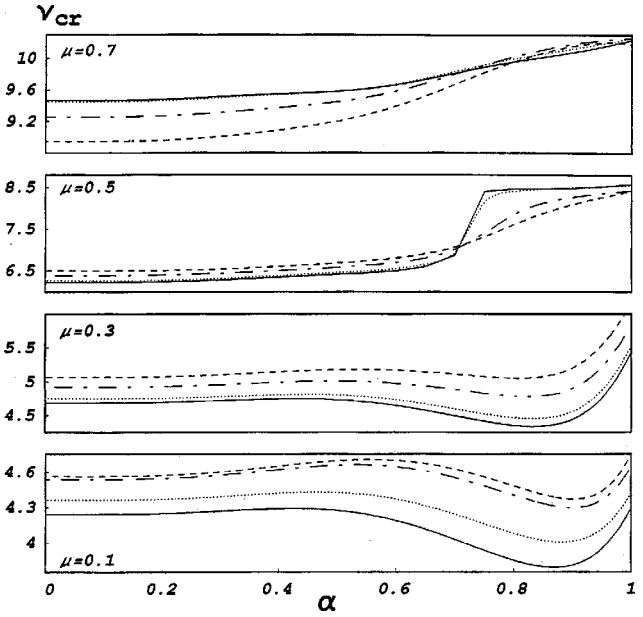


Fig. 2 Dimensionless critical velocity ν_{cr} as a function of the attachment ratio α of a Winkler foundation with modulus $\chi_1=200$ and damping coefficient $\delta_1=0.01$, for different values of the mass ratio μ and of the internal damping (solid: $\gamma=0$; dotted: $\gamma=0.001$; dash-dotted: $\gamma=0.005$; dashed: $\gamma=0.01$). External damping $\beta=0.001$.

cally when the attachment length a of the foundation increases from zero to l. In fact, for smaller values of the mass ratio, namely $\mu=0.1$ or $\mu=0.3$, increase in the attachment length of the foundation beyond a certain threshold leads to a *drop* of the critical velocity. It starts to increase again once the attachment length increases beyond another value. For example, for $\mu=0.1$ and $\gamma=0.001$ (dotted line), in the range $0.479\leqslant\alpha\leqslant0.883$ the dependence ν_{cr} versus the attachment ratio is a decreasing function with ν_{cr} reaching its minimum equal to 4.005 at $\alpha=0.833$; this value is 8.2 percent less than the critical velocity ($\nu_{cr}=4.362$) of the pipe with no foundation. The internal damping increases the critical velocity for low values of μ (inspect the case with $\mu=0.1$ or $\mu=0.3$), whereas its contribution becomes smaller, although still stabilizing, for $\mu=0.5$ in the range $0\leqslant\alpha\leqslant0.7$. Internal damping reduces the critical velocity for $\mu=0.5$ beyond the attachment ratio 0.7. It affects likewise in the case $\mu=0.7$, for attachment ratio below 0.75. Such a reduction is usually characterized in the literature as destabilizing effect, although Bolotin and Zhinzher [20] warn that this term is ''not a very appropriate concept.'' This may suggest that the terms stabilizing effect and destabilizing effect are scientific slang, although widely used. Recent paper by Semler, Alighanbari and Païdoussis [21] provides new physical insight into this phenomenon. Note that the critical velocity appreciably increases as the mass ratio grows. This result was apparently first pointed out by Gregory and Païdoussis [16]. Note that for the follower force analog of this problem, without internal damping but with partial foundation, was considered by Elishakoff and Wang [22], unaware at that time of the Koiter's [8] criticism on the very existence of the follower forces. In this case, the dependence of the critical load versus the attachment ratio had a single maximum; moreover, when the attachment is full, the flutter load equals to that of the column without foundation. As we see the consideration of the realistic problem (that of the pipe containing fluid) has *less* amount of surprising effects than its unrealistic counterpart, the column with follower forces, yet intuitively unexplainable effects remain.

The contribution of the Winkler foundation modulus χ_1 is considered in Fig. 3 ($\mu=0.1$). It shows that the amplitude of this parameter does not alter the nonmonotonous behavior of the critical velocity as a function of the length of the partial foundation. The stabilizing effect of a full Winkler foundation is insignificant, although assuredly present, for mass ratio $\mu=0.1$. Namely, at $\alpha=0$, $\nu_{cr}=4.362$; at $\alpha=1$, $\nu_{cr}=4.363$ for $\chi_1=10$ whereas $\nu_{cr}=4.421$ for $\chi_1=200$. Figure 4 portrays the dependence ν_{cr} versus

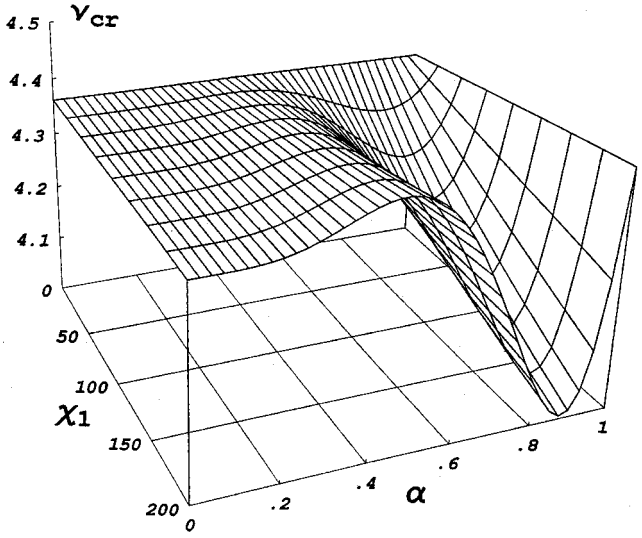


Fig. 3 Dimensionless critical velocity ν_{cr} versus the attachment ratio α and the Winkler foundation modulus χ_1 ($\delta_1=0.01$). Internal damping $\gamma=0.001$; external damping $\beta=0.001$; mass ratio $\mu=0.1$.

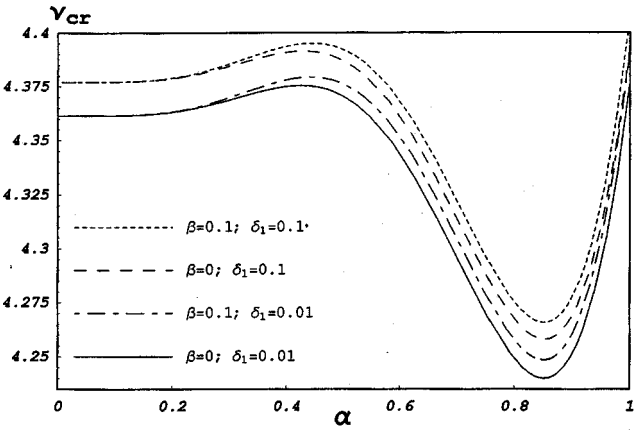


Fig. 4 Dimensionless critical velocity ν_{cr} as a function of the attachment ratio α of a Winkler foundation with modulus $\chi_1=50$, for different values of the damping coefficients β and δ_1. Internal damping $\gamma=0.001$; mass ratio $\mu=0.1$.

α as influenced by the external damping β and the damping associated to the foundation. As one can see they increase the flutter velocity albeit not in large extent.

5.1 Rotatory Foundation Alone. Consider now a pipe resting on a purely rotatory foundation ($\chi_1=0$; $\chi_2\neq0$). As in the case of a purely Winkler foundation, the fully attached rotatory foundation has a stabilizing effect. Indeed, for the fully attached pipes with $\mu=0.3$ and $\gamma=0.001$ (Fig. 5, dotted line), purely rotatory foundation ($\chi_1=0$; $\chi_2=10$) increases the critical velocity by 37.9 percent. Note that the Winkler foundation with $\chi_1=10$ and $\chi_2=0$ results in 0.84 percent enhancement. The nonmonotonic behavior characteristic of purely Winkler foundation persists, although it is less pronounced (see Fig. 5). The effect of the internal damping is similar in both the cases, and will not be recapitulated in this subsection.

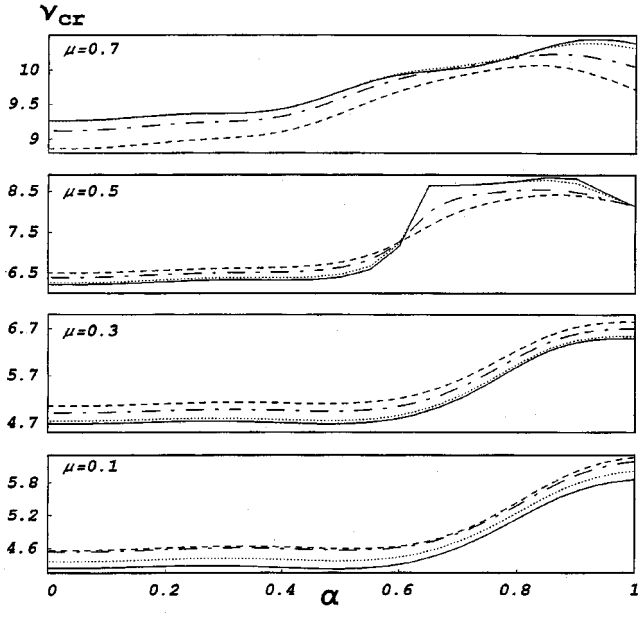


Fig. 5 Dimensionless critical velocity ν_{cr} as a function of the attachment ratio α of a rotatory foundation with modulus $\chi_2=10$ and damping coefficient $\delta_2=0.01$, for different values of the mass ratio μ and of the internal damping (solid: $\gamma=0$; dotted: $\gamma=0.001$; dash-dotted: $\gamma=0.005$; dashed: $\gamma=0.01$). External damping $\beta=0.001$.

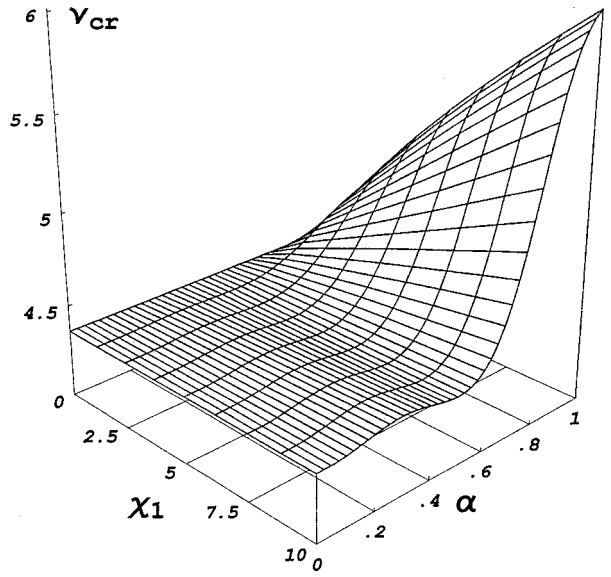


Fig. 6 Dimensionless critical velocity ν_{cr} versus the attachment ratio α and the rotatory foundation modulus χ_2 ($\delta_2 = 0.01$). Internal damping $\gamma = 0.001$; external damping $\beta = 0.001$; mass ratio $\mu = 0.1$.

A paradoxical effect of the partial foundation is evident from Fig. 5. For $\mu = 0.5$ the maximum value of the critical velocity is not reached when the rotatory foundation is fully attached to the pipe. For $\mu = 0.5$ and $\gamma = 0.01$ (dashed line), the maximum critical velocity is reached for the attachment equal to 0.874. The maximum critical velocity for a purely rotatory foundation with $\mu = 0.7$ and $\gamma = 0.01$ is achieved at $\alpha = 0.852$, and drops thereafter. Still, the critical velocity for full foundation is greater than its counterpart for the pipe without foundation.

Figure 6 illustrates, in the case $\mu = 0.1$, the dependence of the critical velocity versus modulus of rotatory elastic foundation χ_2, and the attachment ratio α. For the interval $0 \leqslant \alpha \leqslant 0.6$ the surface appears to be quite flat, with sharp increase recorded thereafter. The effect of the presence of the rotatory foundation is pronounced only when $\alpha > 0.6$. The influence of the external damping and of the damping associated to the rotatory foundation is reported in Fig. 7. They both have a stabilizing effect, although that due to the external damping is larger.

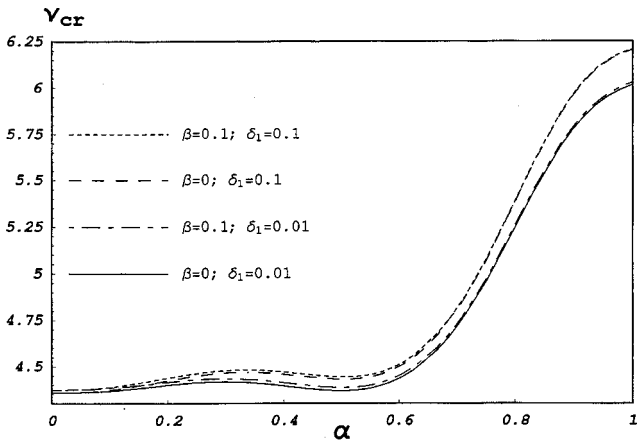


Fig. 7 Dimensionless critical velocity ν_{cr} as a function of the attachment ratio α of a rotatory foundation with modulus $\chi_2 = 10$, for different values of the damping coefficients β and δ_2. Internal damping $\gamma = 0.001$; mass ratio $\mu = 0.1$.

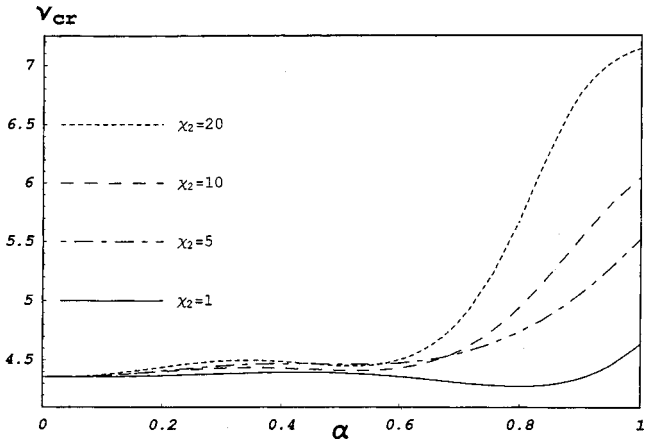


Fig. 8 Dimensionless critical velocity ν_{cr} as a function of the attachment ratio α of a generalized foundation with moduli $\chi_1 = 100$, χ_2 and damping coefficient $\delta_1 = 0.01$, $\delta_2 = 0.01$. Internal damping $\gamma = 0.001$; external damping $\beta = 0.001$; mass ratio $\mu = 0.1$.

5.2 Both Types of Foundations Present. The stability of a pipe resting on a generalized foundation, where the restoring forces and the restoring moments act simultaneously, is of most interest. Figures 8 and 9 depict the critical velocity for χ_1 fixed at 10 while χ_2 is varied. They show that the stability behavior is essentially like that of pipes on purely Winkler or purely rotatory foundation, depending on the ratio of the foundation moduli χ_1/χ_2 and the attachment ratio α. Such an effect of the ratio χ_1/χ_2 is anticipated; the influence of the attachment ratio is less obvious. As we see until the value $\alpha = 0.6$ is reached, the curves associated with different values of χ_2 are in close vicinity between each other. Once $\alpha > 0.6$ pronounced effect of the rotatory modulus occurs. Figure 9 depicts a sharp increase in the critical velocity for $\mu = 0.5$ at specific attachment ratios. For example, $\nu_{cr} = 7.273$ for $\alpha = 0.6$ and $\chi_2 = 10$, at a slight increase in the attachment ratio ($\alpha = 0.65$) the critical velocity increases by 18.86 percent and equals 8.646. As the fully attached rotatory or Winkler foundations, acting alone or in concert, have a stabilizing effect, one can wonder if the presence of a fully attached foundation of one kind only, with the other only partially attached, guarantees an increase of the flutter velocity. The answer to this in-

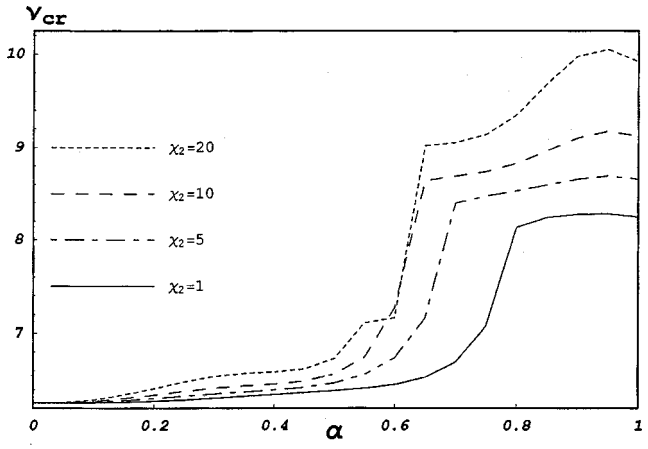


Fig. 9 Dimensionless critical velocity ν_{cr} as a function of the attachment ratio α of a generalized foundation with moduli $\chi_1 = 100$, χ_2 and damping coefficient $\delta_1 = 0.01$, $\delta_2 = 0.01$. Internal damping $\gamma = 0.001$; external damping $\beta = 0.001$; mass ratio $\mu = 0.5$.

quiry is negative. For example, a decrease of the flutter velocity occurs in the case when a fully attached soft rotatory foundation ($\chi_2=1$) is present along with a stiff Winkler foundation ($\chi_1=100$) with attachment ratio $\alpha\approx 0.8$.

6 Conclusions

The effect of the partial foundations was studied on the flutter velocity of the pipe conveying fluid. Unlike the Beck's column on Winkler foundation subjected to purely statically applied ''follower'' load, the dynamic stability behavior of the pipe conveying fluid is influenced by presence of the elastic foundations. The question posed in the title of the paper is answered at this stage as follows: Linear analysis predicts that if elastic foundations acting alone or in combination are attached at the entire length of the pipe, the critical velocity increases; if, however, the attachment is partial the estimate of the critical velocity may either decrease or increase depending on the attachment ratio and other parameters of the system. The news of two kinds appear to be generated. The good news is that the paradoxical result characteristic of the Beck's column with Winkler foundation is absent in the realistic system. Namely, if the foundations are attached along the entire length of the pipe the critical velocity increases. The bad news is the uncovered unexpected nonmonotonous dependence of the critical velocity versus the attachment ratio. This phenomenon could be explained due to following reasons. (1) Possibly, the damping mechanisms known to us do not exhaust all possible damping phenomena; one may speculate that in the future, a new damping mechanism will be uncovered whose inclusion in the linear analysis will remove the above nonmonotonous behavior. Since in words of St. Augustine, we ''know not what we know not,'' the further speculation on this direction appears to be unproductive at this stage at least. (2) ''Small'' nonmonotonicity exhibited in Fig. 9 for $\mu=0.1$ could well be attributed to the structural model used (Bernoulli-Euler theory); use of refined theories may not be associated with nonmonotonicity in some ranges of the parameters. (3) Yet another explanation could lie in recognition of the fact that the pipe undergoes large displacement and the nonlinear analysis of the problem is called for. This latter assessment expressed by the first author was shared by Koiter [23]. A study of the title problem in nonlinear setting is underway and will be reported elsewhere. (4) One cannot rule out the possibility that even the nonlinear setting will not remove the nonmonotonous dependence of the stability characteristics on the elastic foundations. Indeed, to quote Thompson [24], the pipe conveying fluid belongs to the class of problems where '' . . . conventional structural theorems can be not only violated, but actually reversed under fluid loading due to its essentially nonconservative character.''

Acknowledgments

The first author appreciates discussions he had with Professor Dr. Ir. W. T. Koiter, Prof. A. Kornecki, Prof. M. P. Païdoussis, Dr. I. Lottati, Dr. Y. J. Ren, and Dr. X. Qiu on the topic studied in this paper. The study was conducted when N. Impollonia was a Visiting Research Scholar at the Department of Mechanical Engineering of the Florida Atlantic University under the auspices of the University of Messina, Italy. This support is gratefully appreciated.

References

[1] Smith, T. E., and Herrmann, G., 1972, ''Stability of a Beam on an Elastic Foundation Subjected to a Follower Force,'' ASME J. Appl. Mech., **39**, pp. 628–629.
[2] Sundararajan, C., 1974, ''Stability of Columns on Elastic Foundations Subjected to Conservative and Non-Conservative Forces,'' J. Sound Vib., **37**, No. 1, pp. 79–85.
[3] Anderson, G. L., 1976, ''The Influence of a Wieghardt Type Elastic Foundation on the Stability of Some Beams Subjected to Distributed Tangential Forces,'' J. Sound Vib., **44**, No. 1, pp. 103–118.
[4] Celep, Z., 1980, ''Stability of a Beam on Elastic Foundation Subjected to a Non-Conservative Load,'' ASME J. Appl. Mech., **47**, pp. 111–120.
[5] Voloshin, I. I., and Gromov, V. G., 1977, ''On a Stability Criterion for a Bar on an Elastic Base Acted on by a Following Force,'' Mekhanika Tverdogo Tela (Mechanics of Solids), **12**, No. 4, pp. 169–171.
[6] Becker, M., Hauger, W., and Winzen, W., 1977, ''Influence of Internal and External Damping on the Stability of Beck's Column on an Elastic Foundation,'' J. Sound Vib., **54**, No. 3, pp. 468–472.
[7] Panovko, Ya. G., and Gubanova, S. V., 1987, *Stability and Oscillations of Elastic Systems, Paradoxes, Fallacies and New Concepts*, 4th Russian ed., Nauka Publishers, Moscow, pp. 131–132 (English translation of the first edition, 1965, Consultants Bureau, New York).
[8] Koiter, W. T., 1985, ''Elastic Stability,'' Z. Flugwiss. Weltraumforsch., **9**, pp. 205–210.
[9] Koiter, W. T., 1996, ''Unrealistic Follower Forces,'' J. Sound Vib., **194**, No. 4, pp. 636–638.
[10] Herrmann, G., 1967, ''Stability and Equilibrium of Elastic Systems Subjected to Non-Conservative Forces,'' Appl. Mech. Rev., **20**, pp. 103–108.
[11] Païdoussis, M. P., and Li, G. X., 1993, ''Pipes Conveying Fluid: A Model Dynamical Problem,'' Journal of Fluids and Structures, **7**, pp. 137–204.
[12] Païdousis, M. P., 1998, *Fluid-Structure Interaction*, Academic Press, London.
[13] Becker, M., Hauger, W., and Winzen, W., 1978, ''Exact Stability of Uniform Cantilevered Pipes Conveying Fluid or Gas,'' Arch. Mech., **30**, pp. 757–768.
[14] Lottati, I., and Kornecki, A., 1986, ''The Effect of an Elastic Foundation and of Dissipative Forces on the Stability of Fluid-Conveying Pipes,'' J. Sound Vib., **109**, No. 2, pp. 327–338.
[15] Benjamin, T. B., 1961, ''Dynamics of a System of Articulated Pipes Conveying Fluid. I. Theory,'' Proc. R. Soc. London, Ser. A, **A261**, pp. 457–486.
[16] Gregory, R. W., and Païdoussis, M. P., 1966, ''Unstable Oscillation of Tubular Cantilevers Conveying Fluid. I. Theory,'' Proc. R. Soc. London, Ser. A, **A293**, pp. 512–527.
[17] Blevins, R. D., 1990, *Flow Induced Vibrations*, Van Nostrand Reinhold, New York, pp. 384–414.
[18] Duncan, W. J., 1937, ''Galerkin's Method in Mechanics and Differential Equations,'' Aeronautical Research Committee, Reports and Memoranda, No. 1798.
[19] Krall, A. M., 1968, *Stability Techniques for Continuous Linear Systems*, Gordon and Breach, New York, pp. 41–56.
[20] Bolotin, V. V., and Zhinzher, N. I., 1969, ''Effects of Damping on Stability of Elastic System Subjected to Nonconservative Forces,'' Int. J. Solids Struct., **5**, pp. 965–989.
[21] Semler, C., Alighanbari, H., and Païdoussis, M. P., 1998, ''A Physical Explanation of the Destabilizing Effect of Damping,'' ASME J. Appl. Mech., **65**, pp. 642–648.
[22] Elishakoff, I., and Wang, X., 1987, ''Generalization of Smith-Herrmann Problem With the Aid of Computerized Symbolic Algebra,'' J. Sound Vib., **117**, No. 3, pp. 537–542.
[23] Koiter, W. T., 1995, private communication with I. Elishakoff.
[24] Thompson, J. M. T., 1982, ''Paradoxical Mechanics Under Fluid Flow,'' Nature (London), **296**, pp. 169–171.

Optimal Fiber Orientation in Locally Transversely Isotropic Creeping Structures

D. N. Robinson

W. K. Binienda
Assoc. Mem. ASME

Civil Engineering Department,
University of Akron,
Akron, OH 44325-3905

An approximate method is developed for comparing various fiber configurations in a composite structure with the objective of achieving optimal resistance to creep failure. The class of composite structures addressed has a single family of long or continuous strong fibers embedded in a creeping matrix material, e.g., a polymer, metal, etc. Thus, the structure is locally transversely isotropic with the fiber orientation generally varying throughout the structure. The proposed method, intended as an early design tool, is based on an upper bound on creep rupture time and an associated representative failure stress. The latter is evaluated and compared for different fiber configurations, thereby identifying that with optimal creep rupture resistance. This approach allows a substantial saving in computational time by avoiding a detailed analysis of the actual failure process. Application is made to a fiber-reinforced thick-walled cylindrical pressure vessel.
[DOI: 10.1115/1.1354623]

Introduction

When strongly reinforced composite structures operate at temperatures in the creep range of their matrix material, they suffer time-dependent deformation and eventually fail. The failure time is critically influenced by the fiber configuration in the composite structure. Design engineers need simple methods of assessing the effects of fiber orientation on creep rupture, especially for use in the early stages of design.

A creep damaging structure initially incurs damage in regions of relatively high stress, leading to local softening and subsequent redistribution of stress as the damage zone spreads throughout the structure. Detailed numerical calculations of the stress redistribution and corresponding failure are known to require substantial computing time even for relatively simple structures. Repeating such calculations for several fiber configurations in support of an optimal design can lead to prohibitive computational cost. The objective of this research is to provide a simple, yet reliable, method of assessing various fiber configurations in a composite structure with an aim toward identifying that with optimal resistance to creep failure.

An approximate method is proposed that makes use of a constitutive model for a creeping, damaging, anisotropic material introduced by Robinson, Binienda, and Miti-Kavuma [1]. This model applies to composite structures having a single family of strong fibers embedded in a creeping matrix (i.e., a locally transversely isotropic material). It incorporates an isochronous damage function that depends on an invariant specifying the maximum tensile stress normal to the local fiber-matrix interface. This stress component is identified in Robinson et al. [1] as the principally damaging traction in a continuous fiber, metal-matrix composite material. Here, we assert that this stress component is similarly damaging for any strongly reinforced creeping composite with long or continuous fibers, including glass and carbon fiber-reinforced polymers (GFRP and CFRP).

Instead of using the constitutive law of Robinson et al. [1] in a detailed analysis of creep and creep damage of a given composite structure, it is used in bounding the rupture time as proposed earlier by Leckie and Wojewodski [2], Goodall and Cockroft [3], Ponter [4], and Robinson and Wei [5]. Upper bounds on the rupture time are found by considering an imaginary structure, identical in shape and fiber configuration to that of the real one under consideration and subjected to the same loads. The material of the fictitious structure is nondamaging and perfectly-plastic with a yield function that coincides with the isochronous damage function of the actual damaging material, i.e., of the Robinson et al. [1] model. The bound calculation requires only the determination of a (time-independent) limit-load solution for the imaginary structure under the given loads, circumventing a detailed analysis of the time-dependent failure process in the actual damaging structure.

Following Robinson and Wei [5], a global representative failure stress is calculated based on the upper bound on failure time. This is computed for each fiber configuration of interest in a composite structure; the configuration with the smallest representative failure stress has, relatively, the best creep failure resistance. Of course, this amounts to comparing various fiber configurations on the basis of upper bounds on their failure time. Ultimately, this must be corroborated experimentally.

Application of the present method is made to a thick-walled cylindrical pressure vessel under interior pressure; we compare two fiber orientations, circumferential and axial. As expected intuitively, circumferential fiber orientation is ''optimal,'' i.e., the better of the two configurations examined.

Constitutive Equations—Creep Damage

As the proposed method is based on the constitutive model of Robinson et al. [1], we begin by stating its essential features relating to the present work. Consider a composite structure having long or continuous fibers in a creeping matrix material (Fig. 1). The strong fibers generally vary in direction throughout the structure; the fiber direction at each point is denoted by a unit vector field $d_i(x_k)$. A symmetric orientation tensor $D_{ij}=d_i d_j$ is defined as in earlier work, cf., Spencer [6], Rogers [7], Robinson and Duffy [8], and Robinson et al. [1].

The composite material is considered pseudo-homogeneous and transversely isotropic with its overall properties discernible by experiment; it is elastic, nonlinear viscous (creeping) and creep damaging. The constitutive relations for creep and creep damage are

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 12, 1999; final revision, Aug. 1, 2000. Associate Editor: I. M. Daniel. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

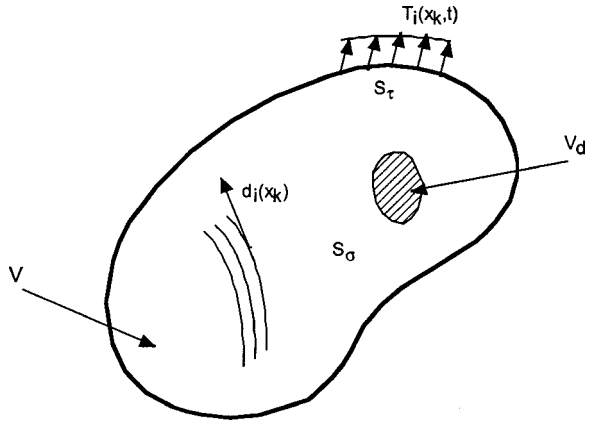


Fig. 1 Structure with strong reinforcement fibers

$$\frac{\dot{\varepsilon}_{ij}^{c}}{\dot{\varepsilon}_o}=\phi^n\frac{\partial\phi}{\partial(\sigma_{ij}/\sigma_o)}\frac{1}{\psi^n} \tag{1}$$

$$\dot{\psi}=-\frac{1}{(p+1)t_o}\Delta^{\nu}\frac{1}{\psi^p} \tag{2}$$

in which σ_{ij} is the (Cauchy) stress, $\dot{\varepsilon}_{ij}^{c}$ is the creep rate of deformation, σ_o is a reference stress chosen in the stress range of interest and $\dot{\varepsilon}_o, n, p, t_o$ and ν are material constants obtained from uniaxial, transverse creep rupture tests at a given temperature, cf., Robinson et al. [1]. The scalar ψ is the material continuity (Kachanov [9]); it equals unity for a material element entirely intact and zero for an element having lost load-carrying capacity entirely.

The stress σ_{ij} and material orientation D_{ij} enter through the functions ϕ and Δ defined in Robinson et al. [1]. As we are principally concerned with creep failure, the flow function ϕ is not of immediate interest and will not be restated here. The isochronous damage function Δ is

$$\Delta(\sigma_{ij},D_{ij})=\frac{N}{\sigma_o} \tag{3}$$

in which

$$N(\sigma_{ij},D_{ij})=\left\langle\frac{1}{2}(J_1-J_o)+\sqrt{T}\right\rangle \tag{4}$$

where the angular brackets in (4) are the Macaulay brackets.

The invariants contained in (4) are defined as follows:

$$J_1=\sigma_{ii}\quad J_o=D_{ij}\sigma_{ji}\quad J_2=\frac{1}{2}s_{ij}s_{ji}$$
$$I=D_{ij}s_{ji}\quad I_o=D_{ij}s_{jk}s_{ki}\quad T=J_2+\frac{1}{4}I^2-I_o. \tag{5}$$

s_{ij} are the components of the deviatoric stress.

Physically, the invariant T in (5) is the square of the maximum transverse shear stress at material element in the structure. N in (4) represents the maximum transverse tensile stress, i.e., the maximum tensile stress normal to the local fiber-matrix interface.

Global Structural Damage

The structural problem of concern is illustrated in Fig. 1. The structure is fixed over part of its bounding surface S_U and loaded with tractions $T_i(x_k,t)$ over the portion S_T. The tractions are supposed to be applied relatively abruptly and held constant thereafter. The conditions are isothermal with the ambient temperature in the creep range of the matrix material.

At time τ local failure occurs at a point (or points) in the structure where $\psi\to 0$. A damage front spreads from this point occupying a volume V_D (Fig. 1) at time $t>\tau$. $V'=V-V_D$ designates the portion of the structure that can sustain stress at time $t>\tau$.

As in Leckie and Wojewodski [2] and Robinson and Wei [5], we introduce a global damage measure based on the Kachanov continuity ψ, i.e.,

$$\Psi=\frac{1}{V}\int_V\psi^{p+1}dV \tag{6}$$

which has the properties, $\Psi=1$ for an undamaged structure and $\Psi=0$ for a failed structure.

We note that as $\psi=0$ in V_D

$$\int_{V'}\psi^{p+1}dV=\int_V\psi^{p+1}dV \tag{7}$$

for all t.

Using (2), we write the rate of global damage as

$$-\frac{d\Psi}{dt}=\frac{1}{t_oV}\int_V\Delta^{\nu}(\sigma_{ij},D_{ij})dV \tag{8}$$

where the constant t_o is the time to rupture under the reference transverse tensile stress σ_o, cf. Robinson et al. [1].

Integrating (8) over $0\to t$ while Ψ varies from $1\to\Psi$, there results

$$\Psi(t)=1-\frac{1}{t_0}\int_0^t\left[\frac{1}{V}\int_V\Delta^{\nu}(\sigma_{ij},D_{ij})dV\right]dt \tag{9}$$

where $\sigma_{ij}(x_k,t)$ is the (redistributing) stress field at time t.

Bounds on Global Damage-Representative Failure Stress

Guided by the work of Goodall and Cockroft [3] and Ponter [4], we specify a fictitious perfectly plastic material whose yield function is

$$\Phi(\sigma_{ij},D_{ij})=\Delta^{\nu}-(Y/\sigma_o)^{\nu} \tag{10}$$

where Δ is the isochronous damage function (3) of the real material and Y is a uniaxial yield stress in transverse tension. As earlier, σ_o is a transverse reference stress.

The flow law for the fictitious perfectly-plastic material is

$$\frac{\dot{\varepsilon}_{ij}^{P}}{\dot{\varepsilon}_o}=\dot{\lambda}\Delta^{\nu-1}\Omega_{ij};\ \text{if}\ \Phi=0\ \text{and}\ \Omega_{ij}\dot{\sigma}_{ij}=0 \tag{11}$$

$$\frac{\dot{\varepsilon}_{ij}^{P}}{\dot{\varepsilon}_o}=0;\ \text{if}\ \Phi<0\ \text{or}\ \Phi=0\ \text{and}\ \Omega_{ij}\dot{\sigma}_{ij}<0 \tag{12}$$

in which $\dot{\lambda}>0$ and

$$\Omega_{ij}=\frac{\partial\Delta}{\partial(\sigma_{ij}/\sigma_o)}=\frac{1}{2}(\delta_{ij}-D_{ij})+\frac{1}{2\sqrt{T}}\left[s_{ij}+\frac{1}{2}(D_{ij}+\delta_{ij})-D_{jk}s_{ki}-D_{ki}s_{jk}\right]. \tag{13}$$

For the given tractions T_i on S_T of the fictitious (and real) structure, Y is chosen so that $T_i\equiv T_i^L$ is the limit load for the yield condition (10). Note that Y can be inhomogeneous, i.e., a function of position.

At the limit load T_i^L, the limit stress field σ_{ij}^L in the fictitious structure is constant (time-independent) in V_P, that part of the volume V in which plastic flow occurs. The strain rate in V_P is

fully plastic, compatible and related to σ_{ij}^L through (11)–(13). The remaining part of the volume is rigid with $\Delta^\nu < (Y/\sigma_o)^\nu$ and $\dot{\varepsilon}_{ij}^P = 0$.

We assume that the isochronous failure surfaces $\Delta = \text{const.}$ are convex in stress space, cf., Robinson et al. [1].[1] As Δ in (3)–(4) is homogeneous of degree one in stress, Δ^ν is a similarly convex and, by definition

$$\Delta^\nu(\sigma_{ij}^a, D_{ij}) - \Delta^\nu(\sigma_{ij}^b, D_{ij}) - \frac{\partial}{\partial \sigma_{ij}} \Delta^\nu(\sigma_{ij}^b, D_{ij})(\sigma_{ij}^a - \sigma_{ij}^b) \geqslant 0 \tag{14}$$

for any two stress states σ_{ij}^a and σ_{ij}^b. We identify $\sigma_{ij}^a = \sigma_{ij}$ in (14) as the stress field in the actual solution of the real structure. We take $\sigma_{ij}^b = \sigma_{ij}^L$ as the limit load stress field in the fictitious structure. With these identifications made, multiplying (14) by $\dot{\lambda} > 0$ and integrating over V_P we obtain

$$\int_{V_P} \Delta^\nu(\sigma_{ij}, D_{ij}) dV \geqslant \int_{V_P} \Delta^\nu(\sigma_{ij}^L, D_{ij}) dV \tag{15}$$

where the contribution from the third term in (14) vanishes by the theorem of virtual work, inasmuch as both of the statically admissible stress fields σ_{ij}^a and σ_{ij}^b are in equilibrium with the same tractions on S_T. Further, as $\Delta^\nu \geqslant 0$, we can rewrite (15) as

$$\int_V \Delta^\nu(\sigma_{ij}, D_{ij}) dV \geqslant \int_{V_P} \Delta^\nu(\sigma_{ij}^L, D_{ij}) dV. \tag{16}$$

Using (16) and recognizing the time-independence of σ_{ij}^L, (9) becomes

$$\Psi(t) \leqslant 1 - \frac{t}{t_o}\left[\frac{1}{V}\int_{V_P} \Delta^\nu(\sigma_{ij}^L, D_{ij}) dV\right]. \tag{17}$$

In the creep failure limit, $t \to t_R$ as $\Psi \to 0$, (17) gives

$$\frac{t_R}{t_o} \leqslant \frac{1}{\dfrac{1}{V}\displaystyle\int_{V_P} \Delta^\nu(\sigma_{ij}^L, D_{ij}) dV} = \frac{t_U}{t_o} = \frac{1}{\left(\dfrac{\sigma_R}{\sigma_o}\right)^\nu}. \tag{18}$$

In (18) we denote the upper bound on the rupture time t_R as t_U. Consistent with (2), we identify σ_R as the representative failure stress, i.e., the uniaxial transverse tensile stress having the failure time t_U, that is

$$\frac{\sigma_R}{\sigma_o} = \left[\frac{1}{V}\int_{V_P} \Delta^\nu(\sigma_{ij}^L, D_{ij}) dV\right]^{1/\nu}. \tag{19}$$

Application of the method requires the representative failure stress (19) to be calculated and compared for various fiber configurations in a given structure. That having the smallest representative failure stress is the optimal fiber configuration. Although relating only to upper bounds on the actual failure time, we assert that this method provides a meaningful comparative measure of the resistance to creep failure of the different fiber configurations.

Application—Thick-Walled Cylinder

As an example problem we choose a fiber-reinforced thick-walled cylinder under interior pressure p. The associated design problem may be stated as: What is the configuration of a single family of fibers (with other features essentially constant, e.g., fiber density, temperature, etc.) that leads to the longest creep rupture time of the cylinder for a given pressure p?

A direct approach to answering this question would be to choose several fiber configurations, and for each, conduct a full numerical, incremental structural analysis allowing for redistribution of stress as the damage region spreads throughout the cylinder and accounting for near singular response resulting from Kachanov-like creep damage evolution (cf. Zienkiewicz and Taylor [10]). Thus determining the time to failure for each fiber configuration, the ''optimal'' configuration is that having the greatest failure time. Note that in order to conduct these complete solutions, we must have the creep/creep rupture behavior fully characterized, viz., (1)–(5), as well as the anisotropic elastic behavior. The direct approach can require substantial material characterization and computing time.

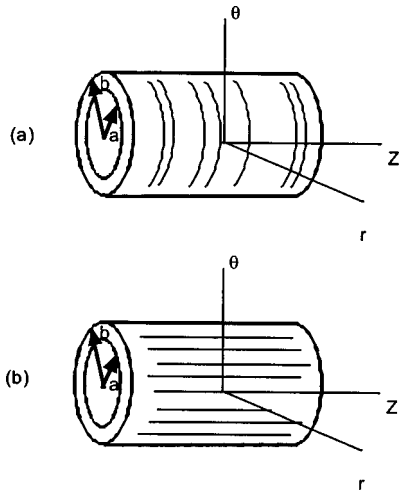


Fig. 2 Thick-walled cylinder under interior pressure (*a*) circumferential reinforcement, (*b*) axial reinforcement

In applying the proposed approximate technique we may similarly choose several fiber configurations, and for each, determine a representative failure stress σ_R using (19). This requires finding a (time-independent) limit load stress field σ_{ij}^L for each fiber arrangement. That configuration having the smallest σ_R is ''optimal.'' Here, for illustrative purposes we shall compare only two fiber-reinforcement configurations in the cylinder, circumferential and axial. In each case the limit load stress fields are relatively easily found. For off axis reinforcement of the cylinder and in more general structural problems, the determination of the limit load stress field σ_{ij}^L can also be relatively difficult. Numerical (finite element) methods have been developed and are available for solving the limit analysis problem, cf., Zienkiewicz and Taylor [10].

We now apply the proposed technique to a thick-walled cylinder under interior pressure p. The cylinder has closed ends with inner and outer radii denoted by a and b, respectively (Fig. 2). In terms of the cylindrical coordinates (z, θ, r) of Fig. 2, the nonzero stress components are

$$\sigma_z, \sigma_\theta \text{ and } \sigma_r. \tag{20}$$

The cylinder is taken to be in a condition of generalized plane strain, i.e., $\dot{\varepsilon}_z = \text{const.}$ The equilibrium equation relating σ_θ and σ_r is

$$\frac{d}{dr}(r\sigma_r) = \sigma_\theta. \tag{21}$$

That for axial equilibrium, involving σ_z, is

$$\pi a^2 p = \int_a^b \sigma_z 2\pi r dr. \tag{22}$$

The compatibility equation is

[1]It is known from experiment that $\Delta = \text{const.}$ surfaces are generally convex in stress space. Specifically, for the pseudo-homogeneous, highly anisotropic class of materials addressed here, such experimental evidence is not definitive. Experiments regarding the question of convexity are in progress.

$$\frac{d}{dr}(r\dot{\varepsilon}_\theta)=\dot{\varepsilon}_r \tag{23}$$

and the boundary conditions are

$$\sigma_r(a)=-p \quad \text{and} \quad \sigma_r(b)=0. \tag{24}$$

The constitutive equations for the imaginary, nondamaging, perfectly plastic material are (10)–(13).

Circumferential Reinforcement. We first consider a circumferentially reinforced cylinder (Fig. 2(*a*)). The components of the orientation tensor are

$$D_{22}\equiv D_\theta=1 \quad \text{otherwise} \quad D_{ij}=0. \tag{25}$$

Calculating the relevant invariants in (5), we determine Δ from (3) and Ω_{ij} from (13), i.e.,

$$\Delta=\frac{\sigma_z}{\sigma_o} \quad \text{and} \quad \Omega_{ij}=\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}. \tag{26}$$

We now construct a statically admissible stress field σ_{ij}^L consistent with a lower bound on the limit load for the fictitious plastic cylinder. The stress components $\sigma_\theta^L(r)$ and $\sigma_r^L(r)$ cannot be found uniquely; they are only required to be statically admissible, i.e., any forms satisfying equilibrium (21) and the boundary conditions (24).

For the axial stress component we take

$$\sigma_z^L=\frac{p}{\left(\frac{b}{a}\right)^2-1} \tag{27}$$

which satisfies the equilibrium Eq. (22). Further, we take the yield stress Y as

$$Y=\sigma_z^L=\frac{p}{\left(\frac{b}{a}\right)^2-1}. \tag{28}$$

Thus, $\sigma_{ij}^L=(\sigma_z^L,\sigma_\theta^L,\sigma_r^L)$ constitutes a lower bound stress field; i.e., it is statically admissible, and through (28) and (26) satisfies the yield condition (10).

Using the flow law (11), we calculate the plastic strain rate field $\dot{\varepsilon}_{ij}^P$ associated with the lower bound stress field σ_{ij}^L. There results

$$\frac{\dot{\varepsilon}_z^P}{\dot{\varepsilon}_o}=\dot{\lambda}\left(\frac{\sigma_z^L}{\sigma_o}\right)^{\nu-1}; \quad \dot{\varepsilon}_\theta^P=\dot{\varepsilon}_r^P=0. \tag{29}$$

This satisfies the compatibility Eq. (23) identically. Further, with $\dot{\lambda}=$const. in (29) the generalized plane strain condition $\dot{\varepsilon}_z=$const. is satisfied. With these conditions met, σ_{ij}^L qualifies as the limit load stress field for the fictitious perfectly-plastic cylinder having a yield stress Y given by (28).

Recognizing that $V_P=V$, we calculate the representative failure stress for circumferentially oriented fibers from (19), i.e.,

$$\sigma_{RC}=\frac{p}{\left(\frac{b}{a}\right)^2-1}. \tag{30}$$

Axial Reinforcement. Next, we consider axial reinforcement (Fig. 2(*b*)). Now, the components of the orientation tensor are

$$D_{11}\equiv D_z=1 \quad \text{otherwise} \quad D_{ij}=0. \tag{31}$$

Evaluating the appropriate invariants in (5) leads to Δ and Ω_{ij} as

$$\Delta=\frac{\sigma_\theta}{\sigma_o} \quad \text{and} \quad \Omega_{ij}=\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}. \tag{32}$$

Following a similar procedure as that for circumferential reinforcement, we construct a statically admissible stress field σ_{ij}^L. We take

$$\sigma_r^L(r)=p\left[\frac{\left(\frac{a}{r}\right)^{1/(\nu-1)}}{\left(\frac{b}{a}\right)^{(\nu-2)/(\nu-1)}-1}+\frac{\left(\frac{a}{r}\right)}{\left(\frac{a}{b}\right)^{(\nu-2)/(\nu-1)}-1}\right] \tag{33}$$

and

$$\sigma_\theta^L(r)=p\left[\frac{\left(\frac{\nu-2}{\nu-1}\right)}{\left(\frac{b}{a}\right)^{(\nu-2)/(\nu-1)}-1}\right]\left(\frac{a}{r}\right)^{1/(\nu-1)}. \tag{34}$$

These are easily shown to satisfy the equilibrium Eq. (21) and the boundary conditions (24). The stress component $\sigma_z^L(r)$ is not determined uniquely; it can assume any form satisfying the equilibrium Eq. (22).

Now, we take

$$Y(r)=\sigma_\theta^L(r)=p\left[\frac{\left(\frac{\nu-2}{\nu-1}\right)}{\left(\frac{b}{a}\right)^{\nu-2/\nu-1}-1}\right]\left(\frac{a}{r}\right)^{1/\nu-1} \tag{35}$$

as the (inhomogeneous) uniaxial yield stress of the fictitious plastic material.

Thus, as before, $\sigma_{ij}^L=(\sigma_z^L,\sigma_\theta^L,\sigma_r^L)$ constitutes a lower bound stress field, i.e., it is statically admissible and from (34) and (32) satisfies the yield condition (10).

Again using (11), we calculate the plastic strain rate field $\dot{\varepsilon}_{ij}^P$ associated with the lower bound stress field σ_{ij}^L. Thus,

$$\frac{\dot{\varepsilon}_\theta}{\dot{\varepsilon}_o}=\dot{\lambda}\left(\frac{\sigma_\theta^L}{\sigma_o}\right)^{\nu-1} \quad \frac{\dot{\varepsilon}_z}{\dot{\varepsilon}_o}=\frac{\dot{\varepsilon}_r}{\dot{\varepsilon}_o}=0. \tag{36}$$

With $\dot{\lambda}=$const. and using (34), (36) satisfies compatibility (23) and the generalized plane-strain condition $\dot{\varepsilon}_z=$const. With these conditions met, σ_{ij}^L qualifies as the limit stress field for the imaginary perfectly-plastic structure. In this case, the fictitious plastic material has an inhomogeneous yield stress given by (35); under this choice we again recognize $V_P=V$.

As earlier, we use (19) to calculate the representative failure stress, giving

$$\sigma_{RA}=p\left[\frac{2}{\left(\frac{b}{a}\right)^2-1}\right]^{1/\nu}\left[\frac{\left(\frac{\nu-2}{\nu-1}\right)}{\left(\frac{b}{a}\right)^{(\nu-2)/(\nu-1)}-1}\right]^{(\nu-1)/\nu}. \tag{37}$$

The better fiber configuration in terms of creep failure resistance is the smaller of (30) and (37). The ratio σ_{RA}/σ_{RC} is

$$\frac{\sigma_{RA}}{\sigma_{RC}}=2^{1/\nu}\left[\left(\frac{\nu-2}{\nu-1}\right)\frac{\left(\frac{b}{a}\right)^2-1}{\left(\frac{b}{a}\right)^{(\nu-2)/(\nu-1)}-1}\right]^{(\nu-1)/\nu}. \tag{38}$$

This ratio is plotted in Fig. 3 versus the material parameter ν for specified ratios of outer to inner radius of the cylinder b/a. Evidently, as $\sigma_{RA}/\sigma_{RC}>1$, circumferential reinforcement is the better fiber arrangement. Figure 3 exemplifies the convenience of the

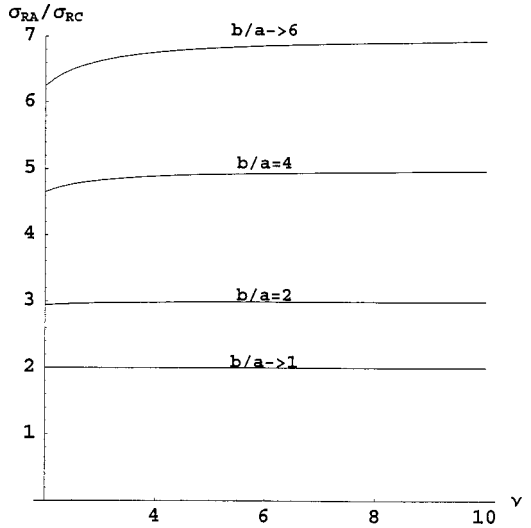


Fig. 3 Ratio of representative failure stresses σ_{RA}/σ_{RC} versus ν for various b/a

representative failure stress as a measure of creep rupture resistance; the ratio σ_{RA}/σ_{RC} is virtually insensitive to the exponent ν.

Limit/Thin-Walled Cylinder

We note that the bracketed term in (38) has the thin-walled limit $(b/a \to 1)$

$$\lim_{b\to a}\left[\left(\frac{\nu-2}{\nu-1}\right)\frac{\left(\frac{b}{a}\right)^2-1}{\left(\frac{b}{a}\right)^{(\nu-2)/(\nu-1)}-1}\right]=2. \tag{39}$$

Into (38) this gives

$$\lim_{b\to a}\frac{\sigma_{RA}}{\sigma_{RC}}=2 \tag{40}$$

for a thin-walled cylinder. This is consistent with the exact solution for the thin-walled cylinder, cf., Robinson et al. [1].

For any homogeneously stressed structure (e.g., the thin-walled tube) in which the stress field does not undergo redistribution, and is thus time-independent, (9) becomes

$$\Psi(t)=1-\frac{t}{t_o}\Delta^{\nu}. \tag{41}$$

In the creep failure limit $t\to t_R$ as $\Psi\to 0$, (41) gives

$$\frac{t_R}{t_o}=\frac{1}{\Delta^{\nu}}=\frac{1}{(\sigma_R/\sigma_o)^{\nu}} \tag{42}$$

thus,

$$\frac{\sigma_R}{\sigma_o}=\Delta(\sigma_{ij},D_{ij}). \tag{43}$$

Evidently, the optimal fiber configuration in a homogeneously stressed structure is that for which Δ is a minimum, or equivalently from (3), that for which the invariant N is minimum. Physically, this is the fiber arrangement that has the least stress normal to the fiber-matrix interface. We conjecture that this fiber configuration is optimal for fast fracture and fatigue as well as for creep failure.

Summary and Conclusions

The proposed method is intended to serve as an early design tool in which a representative failure stress, based on an upper bound on creep failure time, is calculated and compared for various configurations of a single family of fibers in a creeping composite structure. Upon comparing various fiber arrangements, that having the smallest representative failure stress is the optimal choice. The present method avoids potentially costly detailed analyses of the actual creep/creep rupture process for different fiber configurations. It requires the calculation of only a time-independent limit load solution of a fictitious perfectly plastic structure having the actual geometry, fiber configuration and given loads.

An application is made comparing axial and circumferential fiber configurations in a pressurized thick-walled cylinder. As intuitively expected, circumferential fiber orientation is found to be ''optimal'' in creep rupture resistance. The results are examined for the limiting case of a thin-walled cylinder and shown to be consistent with an exact solution obtained earlier in Robinson et al. [1].

The best fiber arrangement for creep failure resistance in a homogeneously stressed structure, where no stress redistribution occurs, is identified as that having the least value of the invariant N, viz., that with the least tensile stress normal to the local fiber-matrix interface. It is conjectured that the same fiber arrangement likewise may be optimal relating to fast fracture and fatigue.

An earlier paper by Robinson and Wei [5] addresses a comparable bounding method in which the fictitious structure is considered nondamaging and viscous with a dissipation potential function that coincides with the isochronous damage function of the actual damaging material. The bound calculation in that case requires only the determination of a (time-independent) steady-state solution for the imaginary viscous structure. Similarly, this avoids a detailed calculation of the actual time-dependent failure process.

As the representative failure stress is based on an upper bound on the failure time, its applicability as a comparative measure for assessing optimal fiber orientations needs to be verified experimentally. An experimental program addressing this and other fundamental features related to this research is currently in progress under funding provided by the National Science Foundation.

The class of structures addressed in this paper is of practical interest in its own right. Moreover, the work admits an extension to structures of particular interest, having two or more families of strong fibers (as angle-ply, cross-ply and woven composites). This extension is a topic of current research of the authors.

References

[1] Robinson, D. N., Binienda, W. K., and Miti-Kavuma, M., 1992, ''Creep and Creep Rupture of Metallic Composites,'' J. Eng. Mech., **118**, No. 8, pp. 1646–1668.

[2] Leckie, F. A., and Wojewodski, W., 1975, ''Estimates of Rupture Life-Constant Load,'' Int. J. Solids Struct., **11**, pp. 1357–1365.

[3] Goodall, J. W., and Cockroft, R. D. H., 1973, ''On Bounding the Life of Structures Subjected to Steady Load and Operating Within the Creep Range,'' Int. J. Mech. Sci., **15**, pp. 251–263.

[4] Ponter, A. R. S., 1977, ''Upper Bounds on the Creep Rupture Life of Structures Subjected to Variable Load and Temperature,'' Int. J. Mech. Sci., **19**, pp. 79–92.

[5] Robinson, D. N., and Wei, Wei, 1996, ''Fiber Orientation in Composite Structures for Optimal Resistance to Creep Failure,'' J. Eng. Mech., **122**, No. 9, pp. 855–859.

[6] Spencer, A. J. M., 1982, ''The Formulation of Constitutive Equations for Anisotropic Solids,'' *Mechanical Behavior of Anisotropic Solids*, J. P. Boehler, ed., Martinus-Nijhoff, Dordrecht.

[7] Rogers, T. G., 1989, ''Rheological Characterization of Anisotropic Materials,'' Composites, **20**, No. 1, pp. 21–38.

[8] Robinson, D. N., and Duffy, S. F., 1990, ''Continuum Deformation Theory for High-Temperature Metallic Composites,'' J. Eng. Mech., **116**, No. 4, pp. 832–844.

[9] Kachanov, L. M. 1986, *Introduction to Continuum Damage Mechanics*, Martinus-Nijhoff, The Netherlands.

[10] Zienkiewicz, O. C., and Taylor, R. L., 1991, *The Finite Element Method, Solid and Fluid Mechanics, Dynamics and Non-Linearity*, McGraw-Hill, New York.

N. Huber
Forschungszentrum Karlsruhe,
Institut für Materialforschung II (IMF II),
Postfach 3640,
D-76021 Karlsruhe, Germany
e-mail: norbert.huber@imf.fzk.de

A. Konstantinidis
Laboratory of Mechanics and Materials,
Aristotle University of Thessaloniki,
GR-54006 Thessaloniki,
Hellas, Greece
e-mail: avraam@mom.gen.auth.gr

Ch. Tsakmakis
Technische Universität Darmstadt,
Institut für Mechanik 1,
Hochschulstrasse 1,
D-64289 Darmstadt, Germany
e-mail: tsakmakis@mechanik.tu-darmstadt.de

Determination of Poisson's Ratio by Spherical Indentation Using Neural Networks—Part I: Theory

When studying analytically the penetration of an indenter of revolution into an elastic half-space use is commonly made of the fraction $E_r = E/(1-\nu^2)$. *Because of this, only* E_r *is determined from the indentation test, while the value of* ν *is usually assumed. However, as shown in the paper, if plastic deformation is involved during loading, the depth-load trajectory depends on the reduced modulus and, additionally, on the Poisson ratio explicitly. The aim of the paper is to show, with reference to a simple plasticity model exhibiting linear isotropic hardening, that the Poisson ratio can be determined uniquely from spherical indentation if the onset of plastic yield is known. To this end, a loading and at least two unloadings in the plastic regime have to be considered. Using finite element simulations, the relation between the material parameters and the quantities characterizing the depth-load response is calculated pointwise. An approximate inverse function represented by a neural network is derived on the basis of these data.*
[DOI: 10.1115/1.1354624]

1 Introduction

The knowledge of Poisson's ratio is of interest when discussing vibration problems ([1–3]), evaluating the toughness of composites (see, e.g., [4]), etc. However, especially for foils and films, it has been claimed to be difficult, if not impossible, to determine the Poisson ratio ([5]). In case of nanoindentation this difficulty arises from the correlation of the unloading stiffness S and the reduced modulus E_r. This results from analytical solutions for an elastic half-space in contact with an indenter of revolution ([6–8]).

It is well known that the indentation test can be employed to obtain mechanical properties of materials. For example, one can determine the Young's modulus (see, e.g., [9–13]), by using an assumed value for Poisson's ratio ν. Olaf [14] has carried out elastic-plastic finite element calculations for pyramidal indenters in order to investigate the sensitivity of the so determined value of Young's modulus with respect to different values of ν. It turned out that ν has only a minor influence on the determined value of E.

When considering spherical indentation of an elastic-plastic material, the load P is commonly plotted against the indentation depth h. Clearly, for a given material, the response of P is a functional of the history of h. However, for a fixed loading history (history of h), one can regard different h-P-plots as resulting from the indentation of materials with different parameters. In other words, the h-P-plots may be interpreted as a function of the material parameters involved in the constitutive theory and the indenter geometry. Such a function can be determined pointwise numerically. Generally, a function of the form

$$P = \hat{P}(E, \nu, \dots) \tag{1}$$

applies for spherical indentation and given loading history, where E denotes the Young's modulus and ν the Poisson ratio.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 26, 1999; final revision, November 1, 2000. Associate Editor: K. T. Ramesh. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

For spherical indentation, the function $\hat{P}$ reduces to the simple form $P \sim E/(1-\nu^2)$, provided the deformation is elastic. However, it cannot be concluded that this relation holds in the case of elastic-plastic deformation. Let E_r denote the reduced modulus, $E_r = E/(1-\nu^2)$, and suppose P to satisfy a relation of the form

$$P = \overline{P}(E_r, \nu, \dots) \tag{2}$$

for the case of elastic-plastic deformations. Then, it is possible to determine ν from (2) in the form

$$\nu = \overline{\nu}(P, E_r, \dots) \tag{3}$$

provided $\overline{P}$ is invertible with respect to ν. Note, that the missing variables in (2) or (3) are the parameters governing the hardening response and the quantities describing the geometry of the problem.

In the present work, a neural network is proposed to represent the function $\overline{\nu}$. Previous work has shown that the problem of elastic-plastic parameter identification from spherical indentation data can only be solved sufficiently accurate by using neural networks, if a priori knowledge is used for the formulation of the input and output quantities of the network ([15,16]). To this end, the available analytical solutions for the indentation of an elastic half-space will be discussed in the first step. Then, these solutions will be generalized to the case of elastic-plastic loading in analogy to (2). From dimensional analysis of the generalized equations, dimensionless-dependent quantities can be derived, which represent the effects of Poisson's ratio and the hardening properties of the material and which are invariant to the associated elastic deformation.

Note that in our paper the assumption is made that the Poisson's ratio remains constant during plastic deformation. This is the case, e.g., for metallic materials which we have in mind.

2 Derivation of the Input Quantities

The solution for elastic contact of nonconforming surfaces has been derived by Hertz [6]. In the case of two elastic solids in contact, which have a spherical geometry in the contact region, the relation between the load P and the approach δ of distant points in the two solids (see [17] p. 93) is given by

$$P = \frac{4}{3} E_r^* \sqrt{R^* \delta^3}, \tag{4}$$

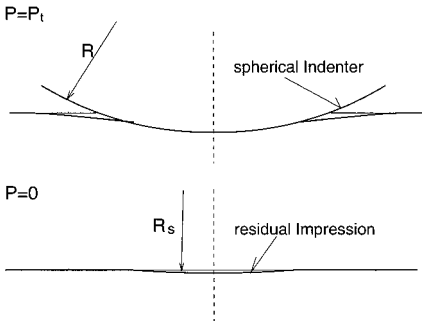


Fig. 1 Sketch of the geometry of spherical indentation with small overall plastic deformation

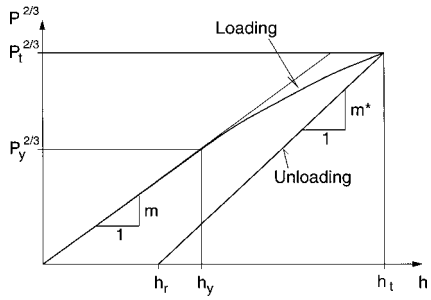


Fig. 2 Sketch of a spherical indentation depth-load response for elastic-plastic deformation

where

$$\frac{1}{E_r^*}=\frac{1-\nu^2}{E}+\frac{1-\nu_i^2}{E_i} \tag{5}$$

and

$$\frac{1}{R^*}=\frac{1}{R}-\frac{1}{R_s}. \tag{6}$$

In these relations, E and ν are Young's modulus and Poisson's ratio of the specimen and E_i and ν_i are Young's modulus and Poisson's ratio of the indenter, respectively. The radius of the indenter is denoted by R and the residual surface of the specimen may have a radius R_s at $P=0$ (see Fig. 1). Note, that for deriving (4) the relation $R_s \gg R$ has been assumed.

To describe elastic-plastic deformation behavior, a plasticity law with yield stress at the beginning of plastic flow k_0 and linear isotropic hardening (modulus E_T) is assumed. Note, that within this work E_T denotes the slope of the hardening rule in the plastic strain-stress diagram. The onset of plastic yield during spherical indentation is given for $\nu=0.3$ by (see Johnson [17] p. 155)

$$P_y^* := P_y|_{\nu=0.3}=\frac{\pi^3 R^2}{6E_r^{*2}}(1.6k_0)^3, \tag{7}$$

where in deriving this formula use has been made of the von Mises yield criterion. Clearly, (4) applies for $\delta=\delta_y^*$,

$$\delta_y^* := \delta_y|_{\nu=0.3}, \tag{8}$$

as well, where δ_y is the value of δ at the beginning of plastic yield, the corresponding value for P being P_y^*:

$$P_y^*=\frac{4}{3}E_r^*\sqrt{R^*\delta_y^{*3}}. \tag{9}$$

Now suppose the specimen to be plane at the beginning, i.e., $R_s=\infty$ and $R^*=R$. Solving (9) with respect to δ_y^* and taking into account (7), as well as $R^*=R$,

$$\frac{\delta_y^*}{R}=\left(0.8\pi\frac{k_0}{E_r^*}\right)^2. \tag{10}$$

In the context of spherical indentation, the approach δ is given by

$$\delta=h-h_r, \tag{11}$$

where h denotes the indentation depth and h_r represents the residual depth resulting from an inelastic deformation of the specimen. Using (10) and (11), the onset of plastic yield ($h_r \equiv 0$) occurs for the first time at the yield depth

$$\frac{h_y^*}{R}=\left(0.8\pi\frac{k_0}{E_r^*}\right)^2, \quad h_y^* := h_y|_{\nu=0.3}. \tag{12}$$

By inserting (11) into (4), it follows that

$$P=\frac{4}{3}E_r^*\sqrt{R^*(h-h_r)^3} \tag{13}$$

and further

$$P^{2/3}=\left(\frac{4}{3}E_r^*\sqrt{R^*}\right)^{2/3}h-\left(\frac{4}{3}E_r^*\sqrt{R^*}\right)^{2/3}h_r. \tag{14}$$

This last equation may be rewritten in the general form

$$P^{2/3}=Mh-B, \tag{15}$$

where

$$M=\begin{cases} m & \text{for elastic loading } (0\leqslant h\leqslant h_y) \\ m^* & \text{after plastic loading} \end{cases}, \tag{16}$$

$$B=\begin{cases} 0 & \text{for elastic loading } (0\leqslant h\leqslant h_y) \\ b^* & \text{after plastic loading} \end{cases}, \tag{17}$$

$$m := \left(\frac{4}{3}E_r^*\sqrt{R}\right)^{2/3}, \tag{18}$$

$$m^* := \left(\frac{4}{3}E_r^*\sqrt{R^*}\right)^{2/3}, \tag{19}$$

$$b^* := m^* h_r. \tag{20}$$

Next, consider a loading history with loading until a point $P_t(h_t)$, $h_t > h_y$ and unloading to $P=0$. By plotting the data in the form $P^{2/3}$ against h, two linear regimes can be considered (see Fig. 2). The first one holds during loading for $0\leqslant h\leqslant h_y$ and the second linear regime is valid during unloading after plastic loading ($h_r\leqslant h\leqslant h_t$). Accordingly, the value of m can be determined from the elastic loading ($0\leqslant h\leqslant h_y$), while the values of m^* and b^* can be obtained from a regression analysis of the unloading data.

The variables prescribed by geometry, material and loading process are

$$(R,h_t,\nu,E_r^*,k_0,E_T). \tag{21}$$

Now, suppose the quantities

$$(m,h_y,P_t,m^*,b^*) \tag{22}$$

to be functions of (21):

$$m=\hat{m}(R,h_t,\nu,E_r^*,k_0,E_T), \tag{23}$$

$$h_y=\hat{h}_y(R,h_t,\nu,E_r^*,k_0,E_T), \tag{24}$$

$$P_t=\hat{P}_t(R,h_t,\nu,E_r^*,k_0,E_T), \tag{25}$$

$$m^*=\hat{m}^*(R,h_t,\nu,E_r^*,k_0,E_T), \tag{26}$$

$$b^*=\hat{b}^*(R,h_t,\nu,E_r^*,k_0,E_T). \tag{27}$$

Table 1 Dimension analysis of elastic-plastic spherical indentation with loading and unloading

	R	h_t	ν	E_r^*	k_0	E_T	m	h_y	P_t	m^*
[N]	0	0	0	1	1	1	2/3	0	1	2/3
[m]	1	1	0	−2	−2	−2	−1	1	0	−1

It is worthy mentioning that ν is assumed to enter explicitly in (21) and thus on the right-hand sides of (23)–(27). The quantities (22) can be calculated as the result of a so-called direct problem using the finite element method as explained below (see Section 3).

Relation (23) is already given by definition (18) so that for known m, the reduced modulus E_r^* is known as well. For $\nu = 0.3$, the analytical solution for (24) is (12). The function $\hat{b}^*$ can be obtained by a linear combination of $\hat{P}_t$ and $\hat{m}^*$ using (15)–(17), (19), (20) provided the point $(h_t, P_t^{2/3})$ is known. The Eqs. (24), (25), and (26), which indicate an explicit dependency on ν, can be written in dimensionless form. The dimension analysis given in Table 1 implies a rank $r=2$.

Note that $n=7$ quantities are related in any of the Eqs. (24)–(26). Hence, when writing these equations in dimensionless form, each of them will involve $m=n-r=5$ dimensionless quantities, e.g., of the form

$$\Pi_1 := \frac{h_t}{h_y} = \hat{\Pi}_1\left(\frac{h_t}{R}, \nu, \frac{k_0}{E_r^*}, \frac{E_T}{E_r^*}\right), \tag{28}$$

$$\Pi_2 := \frac{P_t^{2/3}}{m h_t} = \hat{\Pi}_2\left(\frac{h_t}{R}, \nu, \frac{k_0}{E_r^*}, \frac{E_T}{E_r^*}\right), \tag{29}$$

$$\Pi_3 := \frac{m^*}{m} = \hat{\Pi}_3\left(\frac{h_t}{R}, \nu, \frac{k_0}{E_r^*}, \frac{E_T}{E_r^*}\right). \tag{30}$$

Here Π_1 denotes the depth related to that depth at which plastic flow occurs for the first time. For $\nu=0.3$ we have, by virtue of (12),

$$\Pi_1^* = \hat{\Pi}_1\left(\frac{h_t}{R}, 0.3, \frac{k_0}{E_r^*}, \frac{E_T}{E_r^*}\right), \tag{31}$$

$$= \frac{h_t}{h_y^*} = \frac{h_t}{R}\frac{R}{h_y^*} = \frac{h_t}{R}\left(0.8\pi\frac{k_0}{E_r^*}\right)^{-2} \tag{32}$$

or

$$\Pi_1^* = \hat{\Pi}_1^*\left(\frac{h_t}{R}, \frac{k_0}{E_r^*}\right). \tag{33}$$

Now, we assume Π_2 and Π_3 in (29),(30) to be dependent on h_t/R and k_0/E_r^* over the function $\hat{\Pi}_1^*(h_t/R, k_0/E_r^*)$, so

$$\Pi_2 = \tilde{\Pi}_2\left(\Pi_1^*, \nu, \frac{E_T}{E_r^*}\right), \tag{34}$$

$$\Pi_3 = \tilde{\Pi}_3\left(\Pi_1^*, \nu, \frac{E_T}{E_r^*}\right). \tag{35}$$

A motivation for this assumption, based on finite element simulations, will be given in Section 3.

In order to interpret the dimensionless quantity Π_2, we introduce the total load $P_t^{(e)}$ for a fictitious elastic loading until the depth h_t. According to (15)–(18),

$$P_t^{(e)} := \frac{4}{3}E_r^*\sqrt{R h_t^3} = \sqrt{m^3 h_t^3}. \tag{36}$$

Thus, Π_2 can be interpreted as the ratio

$$\Pi_2 = \left(\frac{P_t}{P_t^{(e)}}\right)^{2/3}. \tag{37}$$

Using (19), (18), and (6), it can be shown that the ratio m^*/m and therefore Π_3 may be expressed in terms of R and R_s,

$$\Pi_3 = \frac{m^*}{m} = \left(1 - \frac{R}{R_s}\right)^{-1/3}, \tag{38}$$

which represents the residual geometry of the specimen surface.

3 Finite Element Simulations

The finite element mesh is displayed in Fig. 3. It consists of eight-node axisymmetric elements, which turned out to be optimal for solving contact problems with spherical indentation (see [18]). The part of the mesh, which is concerned with the contact problem is displayed enlarged on the left-hand side. In order to achieve accurate results for very different sets of material parameters, this refined region should be proportional to the contact radius being achieved. The contact radius can be estimated for given Π_1^* by using (12) and the assumption of zero piling up.

To minimize the amount of finite element meshs, three different sizes of maximum contact areas were defined, as shown in Table 2. The edges of the mesh are 1000 mm and the radius of the indenter is $R=200$ mm. However, the results can be scaled to any indenter radius by using the PI-theorem.

The constitutive law assumed is an elastic-plastic constitutive model available in ABAQUS (see Hibbit [19], Section 4.3.2-1), based on a von Mises yield function with linear isotropic hardening and a flow rule which represents an associated normality rule. Other type of constitutive equations like those derived on the basis of the scale invariance approach ([20,21]) can also be employed, but such issues may be taken up in a future study. As already mentioned in Section 2, the yield stress is denoted by k_0 and the tangent modulus is E_T.

For a set of material parameters, which is typical for steel, a finite element simulation has been carried out for each mesh. The indenter has been pressed eight times of the yield depth into the material and was unloaded to the half of the total depth. For this material, mesh M2 has the optimum contact discretization. The resulting depth-load trajectories for the three finite element meshs are displayed in Fig. 4.

It can be seen that the loading response for M1 deviates slightly from the curves of M2 and M3 which are nearly coincident. These

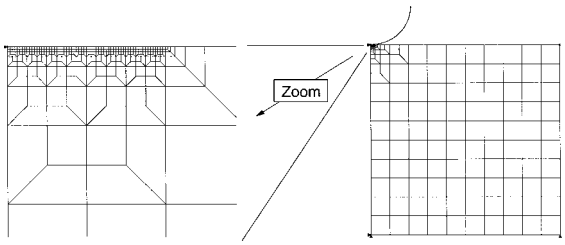


Fig. 3 Finite element mesh for spherical indentation with low loads

Table 2 The three groups of finite element meshs (all lengths in [mm])

Mesh	Element Size	Contact Radius	Depth Range
M1	0.3906	$a<25.00$	$6.25\cdot10^{-2}\leq h_t$
M2	0.0977	$a<6.25$	$2.5\cdot10^{-3}\leq h_t<6.25\cdot10^{-2}$
M3	0.0244	$a<1.5625$	$h_t<2.5\cdot10^{-3}$

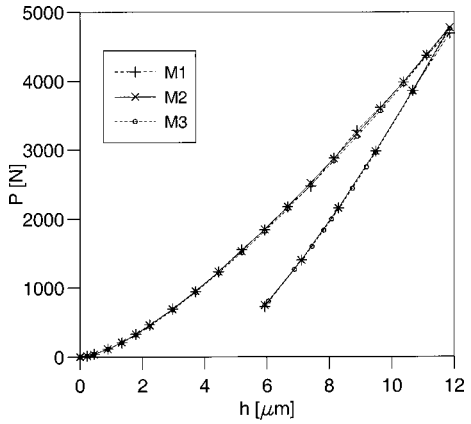


Fig. 4 Depth-load trajectories for the the finite element meshes M1–M3: $E=200$ GPa, $\nu=013$, $k_0=250$ MPa, $E_T=10$ GPa ($a<2.2$ mm)

calculations show on the one hand the mesh independency and on the other hand the higher accuracy due to the three discretization categories.

Using variable h_t, ν, k_0, and E_T each argument of the functions $\hat{\Pi}_i$ in (28)–(30) can be chosen independently. Although it is theoretically not necessary, the reduced modulus E_r^* is chosen randomly between 50 GPa and 600 GPa for each finite element calculation in order to demonstrate that quantities Π_i, and therefore ν, depend on k_0/E_r^* and E_T/E_r^* only. Note that (28)–(30) are functions of E_r^* only, and therefore, without loss of generality, the spherical indenter is modeled in the finite element simulations as a rigid surface.

Figure 5 shows three finite element simulations, with ν being the only dimensionless quantity that has been changed (E_r^* = const). The values of m, h_y^*, and P_y^*, indicated in the plot, have been calculated using (18), (12), and (7), respectively. The good agreement of these analytical values with the numerical results shows that the finite element mesh is suitable for modeling the elastic-plastic half-space accurately.

Next, we use finite element calculations to get values for Π_2 and Π_3 for given values (21), which represent pointwise the functions (29) and (30). As outlined in Huber and Tsakmakis [16], for keeping the number of necessary simulations small when training neural networks, the parameters ν, k_0/E_r^* and E_T/E_r^* have to be chosen randomly from the given intervals in order to form the training and test patterns. For the purpose of the present study,

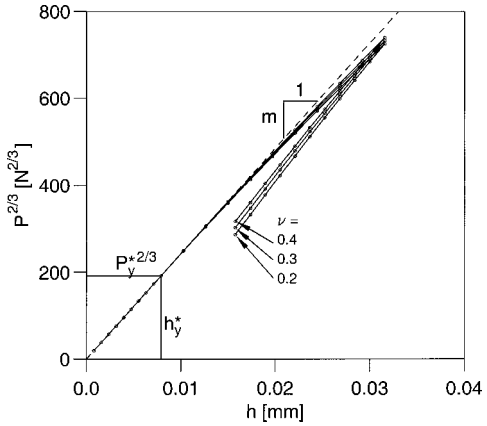


Fig. 5 Examples of finite element simulations for $E_r^*=200$ GPa, $k_0=500$ MPa, $E_T=20$ GPa, $h_t=4h_y^*$

Table 3 Ranges of the dimensionless parameters for the formation of training and test patterns

Parameter		Range
Π_1^*	[-]	1.5–8
ν	[-]	0.1–0.45
k_0/E_r^*	[-]	10^{-5}–10^{-2}
E_T/E_r^*	[-]	10^{-3}–10^{-1}

these intervals are defined as shown in Table 3, from which 100 sets were generated. The values of Π_1^* at which unloadings are inserted are chosen in equi-distant steps of 0.5 so that 14 unloadings are available for each simulation.

Before going any further we will demonstrate that ν effects the depth-load trajectory explicitly and not only by means of E_r. This will justify the assumption made that ν appears in (21) explicitly. To this end the values of Π_2 and Π_3 have been numerically calculated and plotted against ν and E_T/E_r^* for Π_1^* = const. The meshed surfaces in Figs. 6 and 7, which are parametrized with Π_1^*, indicate that Π_2 and Π_3 are affected by ν. Moreover, these

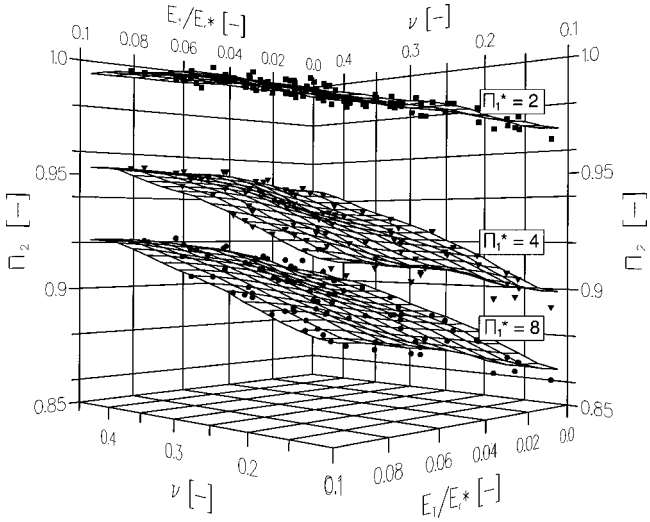


Fig. 6 The effect of ν and E_T/E_r^* on Π_2 for Π_1^* = const

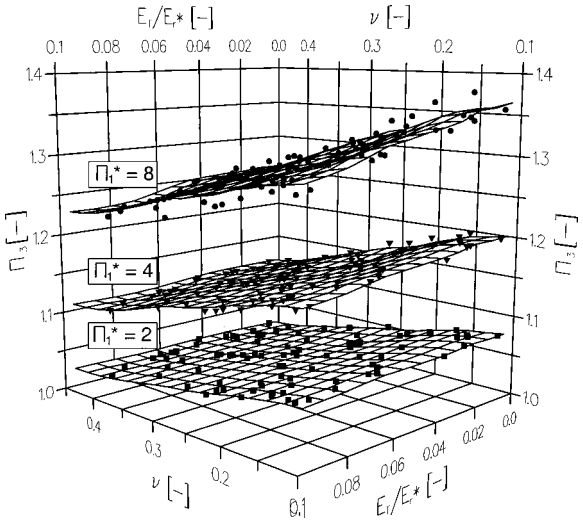


Fig. 7 The effect of ν and E_T/E_r^* on Π_3 for Π_1^* = const

results suggest to assume Π_2 and Π_3 as a function of Π_1^* which is the motivation for Eqs. (34), (35). In what follows, a neural network will be developed on the basis of (34) and (35), which will represent the solution of the inverse problem, i.e., ν as a function of Π_1^*, Π_2, and Π_3.

4 Training of the Neural Networks

Artificial neural networks represent a qualified tool for solving complex inverse problems in computational mechanics. An overview about some relevant applications is given by Yagawa and Okuda [22] and Sumpter and Noid [23]. A neural network consists of neurons connected with links to a highly parallel structure. Each neuron possesses a local memory and can carry out localized information processing operations. In general, each neuron has multiple inputs and a single output value to mimic the biological brain neuron.

According to the neural network method so-called training patterns have to be presented to the network. These consist of the input values x_i and the related desired output values d_l. The error

$$e_l = d_l - y_l \tag{39}$$

for each pattern can be computed from the actual output y_l (see Fig. 8).

Using a backpropagation algorithm, the synaptic weights w_{ij}, which represent the links between the neurons are adjusted appropriately. This way, the error of the output values is minimized and the network has been taught the relation between input and output values. The mean error for all patterns and outputs is given by

$$\mathrm{MSE} := \frac{1}{N}\frac{1}{L}\sum_n \sum_l (e_l^{(n)})^2, \tag{40}$$

where n is the pattern number, N is the number of patterns, and L is the number of output neurons. The neural network simulations are carried out using the SNNS code [24]. The relevant theory of backpropagation and preparation of the data in a form appropriate for our purposes is described, e.g., in [15,16].

4.1 One Unloading (Set 1). Towards determining the Poisson ratio ν, we regard in (34),(35), the quantities Π_1^*, Π_2, and Π_3 to be known. More specifically, Π_1^* is given by material parameters, while Π_2 and Π_3 are regarded to be determined by exploiting a loading and an unloading. Thus, we have two equations with two unknowns $(\nu, E_T/E_r^*)$. However, these equations can only be inverted in a satisfactory manner, if Π_2 and Π_3 are linear independent. By plotting all finite element results of Π_2 against Π_3 one can observe a strong correlation between these two quantities (see Fig. 9), which has the point (1.0,1.0) as origin for $\nu \to 0.5$ and $\Pi_1^* \to 1$ (cf. Figs. 6,7).

From the correlation between Π_2 and Π_3 arises a lack of information which makes difficult to invert the problem numerically. This can be shown by training a neural network, denoted as Set 1, using the input and output definition

$$(x_1, x_2, x_3) := (\Pi_1^*, \Pi_2, \Pi_3), \tag{41}$$

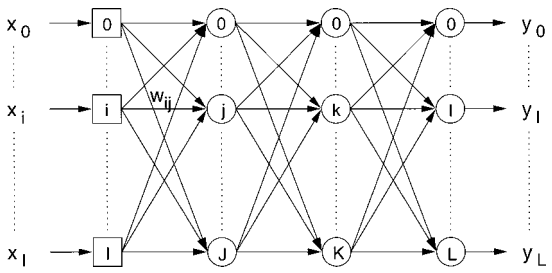


Fig. 8 Sketch of a multilayer feed forward neural net

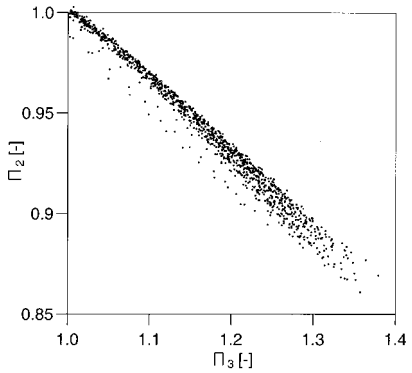


Fig. 9 Correlation of Π_2 and Π_3

$$y := \nu. \tag{42}$$

The neural network consists of three input neurons, two hidden layers with three and two neurons, respectively, and one output neuron. For $\Pi_1^* \leqslant 4$ data have been established, from which 375 training patterns and 38 test patterns were generated, the test patterns are chosen randomly from the database. These test patterns are not presented to the neural network during training and can be used to check the capability of the trained network for unseen patterns.

After 3000 training epochs the MSE value reached $9.6 \cdot 10^{-4}$ and $9.0 \cdot 10^{-4}$ for training and test patterns, respectively. The values of $\tilde{\nu}$ identified from the finite element data are plotted against the exact values ν in Fig. 10 for training and test patterns. It can be seen, that it is possible to determine ν only with a comparably large scatter.

4.2 Two Unloadings (Set 2). The identification can be improved by adding independent information to the input data of the neural network. This can be done, e.g., by using an additional unloading. To this end, to each unloading at a depth $h_{t,1}$ a further unloading $h_{t,2}$ is considered having the depth $\Pi_1^*|_{h_{t,2}} := 2\Pi_1^*|_{h_{t,1}}$. Thus a new neural network, referred to as Set 2, has been trained, for which the input data, for given material data, are

$$(x_1, \ldots, x_5) := (\Pi_1^*|_{h_{t,1}}, \Pi_2^*|_{h_{t,1}}, \Pi_3^*|_{h_{t,1}}, \Pi_2^*|_{h_{t,2}}, \Pi_3^*|_{h_{t,2}}). \tag{43}$$

Note that the depths $h_{t,1}, h_{t,2}$ together with the material data form the basis for a pattern. This neural network consists of five input neurons, two hidden layers with four and three neurons, respectively, and one output neuron. Again, 375 training patterns and 38 test patterns were generated with $1.5 \leqslant \Pi_1^*|_{h_{t,1}} \leqslant 4$. After the same duration of 3000 epochs an MSE value of $9.0 \cdot 10^{-5}$ for training and $1.2 \cdot 10^{-4}$ for test patterns was achieved. The reidentification

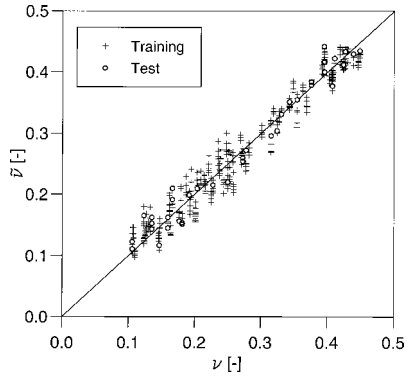


Fig. 10 Identification of ν on the basis of one unloading

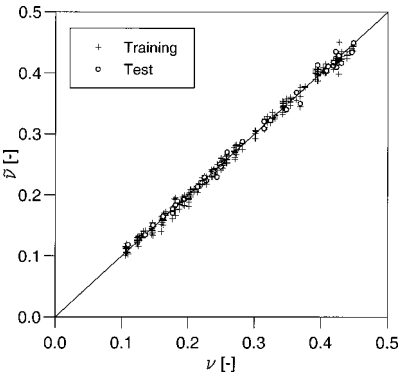


Fig. 11 Identification of ν on the basis of two unloadings

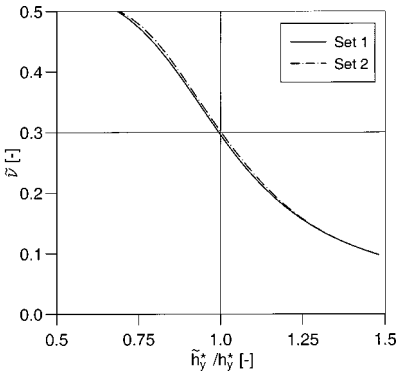


Fig. 12 Dependency of identified Poisson's ratio $\tilde{\nu}(\tilde{h}_y^*)$ using Set 1 and Set 2

results of all finite element simulations are displayed in Fig. 11. It can be seen that the identification error is reduced now to an amount which can considered to be caused by the numerical error of the finite element data.

These results illustrate, that two unloadings are able to provide sufficiently independent information in order to solve the problem of determining ν with high accuracy when the value of $\Pi_1^*|_{h_{t,1}}$ is known.

Now, assuming for an unloading at $h_t \equiv h_{t,1}$, that onset of plastic yield may occur at every $\tilde{h}_y^* < h_t$ a relation $\tilde{\nu}(\tilde{h}_y^*)$ can be determined using Set 1 or Set 2. This is plotted in Fig. 12, where the horizontal line ($\nu = 0.3$) and vertical line ($\tilde{h}_y^*/h_y^* = 1$) indicate the exact values of ν and h_y^*, respectively. From Fig. 12 it can be seen that the correct value $\nu = 0.3$ is determined by Set 1 and Set 2 with good accuracy, if h_y^* is known. Note that the curves $\tilde{\nu}(\tilde{h}_y^*)$ for both, Set 1 and Set 2, indicate the same slope at the point $\tilde{h}_y^*/h_y^* = 1$.

5 Conclusions

In the present work some basic issues towards determining Poisson's ratio from spherical indentation have been discussed. It is shown that the solution for elastic-plastic deformation is sufficient accurate, provided the onset of plastic yield is known. However, as shown in Section 4, it must be noted that a slight deviation from h_y^* affects the resulting value of ν considerably.

The determination of the exact value h_y^* where plastic deformation occurs for the first time, is not trivial. This problem cannot be solved with conventional methods to achieve the accuracy, needed. Consequently, the value of ν should be determined without knowledge of h_y^*. To this end, it is necessary to modify the inverse function appropriately in order to decrease the sensitivity corresponding to the onset of plastic yield and thus to make the method more suitable to experiments. The development of such an inverse function is addressed in [25].

Acknowledgments

The first author gratefully acknowledges financial support from the Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren. Also, the encouragement and support of Prof. E. C. Aifantis involving in particular the second author through the EU projects TMR #ERBFMRXCT960062 and REVISA #FI4SCT960024 is gratefully acknowledged.

References

[1] Rossi, R. E., and Laura, P. A. A., 1996, ''On the Effect of Poisson Ratio and Certain Approximation Schemes on Transverse Vibration of Thin Rectangular Plates With a Free Edge,'' J. Sound Vib., **194**, pp. 439–444.
[2] Rossi, R. E., and Laura, P. A. A., 1996, ''Symmetric and Antisymmetric Normal Modes of a Cantilever Rectangular Plate: Effect of Poisson's Ratio and a Concentrated Mass,'' J. Sound Vib., **195**, pp. 142–148.
[3] Laura, P. A. A., Sonzogni, V., and Romanelli, E., 1996, ''Effect of Poisson's Ratio on the Fundamental Frequency of Transverse Vibration and Buckling Load of Circular Plates With Variable Profile,'' Appl. Acoust., Oxford, **74**, pp. 263–274.
[4] Evans, A. G., and Hutchinson, J. W., 1984, ''On the Mechanics of Delamination and Spalling in Compressed Films,'' Int. J. Solids Struct., **20**, pp. 455–466.
[5] Hardwick, D. A., 1987, ''The Mechanical Properties of Thin Films: A Review,'' Thin Solid Films, **154**, pp. 109–124.
[6] Hertz, H., 1882, ''Über die Berührung fester elastischer Körper,'' J. Reine Angew. Math., **92**, pp. 156–171.
[7] Love, A. E. H., 1939, ''Boussinesq's Problem for a Rigid Cone,'' Quart. J. Math., **10**, pp. 161–175.
[8] Harding, J. W., and Sneddon, I. N., 1945, ''The Elastic Stresses Produced by the Indentation of the Plane Surface of a Semi-Infinite Elastic Solid by a Rigid Punch,'' Proc. Cambridge Philos. Soc., **43**, pp. 16–26.
[9] Doerner, M. F., and Nix, W. D., 1986, ''A Method for Interpreting the Data From Depth-Sensing Indentation,'' J. Mater. Res., **1**, pp. 601–609.
[10] Oliver, W. C., and Pharr, G. M., 1992, ''An Improved Technique for Determining Hardness and Elastic Modulus Using Load and Displacement Sensing Indentation Experiments,'' J. Mater. Res., **7**, pp. 1564–1583.
[11] Field, J. S., and Swain, M. V., 1993, ''A Simple Predictive Model for Spherical Indentation,'' J. Mater. Res., **8**, pp. 297–306.
[12] Taljat, B., Haggag, F. M., and Zacharia, T., 1997, ''Analysis of Ball-Indentation Load-Depth Data: Part I. Determining Elastic Modulus,'' J. Mater. Res., **12**, pp. 965–974.
[13] Huber, N., Munz, D., and Tsakmakis, Ch., 1997, ''Determination of Young's Modulus by Spherical Indentation,'' J. Mater. Res., **12**, pp. 2459–2469.
[14] Olaf, 1992, ''Ein Verfahren zui Bersertung des Mechanischen Verholtens von Ramischichten,'' Ph.D. thesis, Albert-Ludwigs-Universität Freiburg im Breisgau, Germany.
[15] Huber, N., and Tsakmakis, Ch., 1999, ''Determination of Constitutive Properties From Spherical Indentation Data Using Neural Networks, Part I: The Case of Pure Kinematic Hardening in Plasticity Laws,'' J. Mech. Phys. Solids, **47**, pp. 1569–1588.
[16] Huber, N., and Tsakmakis, Ch., 1999, ''Determination of Constitutive Properties From Spherical Indentation Data Using Neural Networks, Part II: Plasticity With Nonlinear Isotropic and Kinematic Hardening,'' J. Mech. Phys. Solids, **47**, pp. 1589–1607.
[17] Johnson, K. L., 1985, *Contact Mechanics*, Oxford University Press, Oxford, UK.
[18] Huber, N., and Tsakmakis, Ch., 1998, ''A Finite Element Analysis of the Effect of Hardening Rules on the Indentation Test,'' J. Eng. Mater. Technol., **120**, pp. 143–148.
[19] Hibbit, Karlsson, & Sorensen, 1996, ABAQUS theory, Version 5.6.
[20] Aifantis, E. C., 1987, ''The physics of plastic deformation,'' Int. J. Plasticity, **3**, pp. 211–247.
[21] Aifantis, E. C., 1995, ''From Micro- to Macro-Plasticity: The Scale Invariance Approach,'' J. Eng. Mater. Technol., **117**, pp. 352–355.
[22] Yagawa, G., and Okuda, H., 1996, ''Neural Networks in Computational Mechanics,'' Arch. Comput. Meth. Eng., **3**, pp. 435–512.
[23] Sumpter, B. G., and Noid, D. W., 1996, ''On the Design, Analysis and Characterization of Materials Using Computational Neural Networks,'' Annu. Rev. Mater. Sci., **26**, pp. 223–278.
[24] SNNS, 1995, ''SNNS Stuttgart Neural Network Simulator,'' User Manual, Version 4.1. University of Stuttgart, Institute for Parallel and Distributed High Performance Systems (IPVR).
[25] Huber, N., and Tsakmakis, Ch., 2001, ''Determination of Poisson's Ratio by Spherical Indentation Using Neural Networks. Part II: Identification method,'' ASME J. Appl. Mech., **68**, pp. 224–229.

Determination of Poisson's Ratio by Spherical Indentation Using Neural Networks—Part II: Identification Method

N. Huber
Forschungszentrum Karlsruhe,
Institut für Materialforschung II (IMF II),
Postfach 3640,
D-76021 Karlsruhe, Germany
e-mail: norbert.huber@imf.fzk.de

Ch. Tsakmakis
Technische Universität Darmstadt,
Institut für Mechanik 1,
Hochschulstrasse 1,
D-64289 Darmstadt, Germany
e-mail: tsakmakis@mechanik.tu-darmstadt.de

In a previous paper it has been shown that the load and the unloading stiffness are, among others, explicit functions of the Poisson's ratio, if a spherical indenter is pressed into a metal material. These functions can be inverted by using neural networks in order to determine the Poisson's ratio as a function of the load and the unloading stiffness measured at different depths. Also, the inverse function possesses as an argument the ratio of the penetration depth and that depth, at which plastic yield occurs for the first time. The latter quantity cannot be measured easily. In the present paper some neural networks are developed in order to identify the value of Poisson's ratio. After preparing the input data appropriately, two neural networks are trained, the first one being related to Set 2 of the previous paper. In order to avoid an explicit measurement of the yield depth, the second neural network has to be trained in such a way, that its solution intersects with that of Set 2 at the correct value of Poisson's ratio. This allows to identify Poisson's ratio with high accuracy within the domain of finite element data.
[DOI: 10.1115/1.1355032]

1 Introduction

In a previous paper (see Huber and Tsakmakis, [1]), it has been shown, that the Poisson's ratio can be identified from spherical indentation, provided the depth is known where plastic yield occurs for first time. To this end, the dimensionless quantities Π_2 and Π_3,

$$\Pi_2 := \frac{P_t^{2/3}}{m h_t}, \tag{1}$$

$$\Pi_3 := \frac{m^*}{m}, \tag{2}$$

have been introduced (cf. Fig. 1). Here, P_t is the maximum load at the depth h_t, where the indenter is unloaded. m is defined as the slope on the h-P-graph for elastic loading ($0 \leqslant h \leqslant h_y$), while m^* denotes the slope on the h-P-graph during elastic unloading after loading until (h_t, P_t). Note that in the elastic regime $\Pi_2 = 1$ and $\Pi_3 = 1$.

It has been shown in [1], that Π_2 and Π_3 are functions of geometry, Poisson's ratio ν, loading history, and the hardening properties

$$\Pi_2 = \tilde{\Pi}_2\left(\Pi_1^*, \nu, \frac{E_T}{E_r^*}\right), \tag{3}$$

$$\Pi_3 = \tilde{\Pi}_3\left(\Pi_1^*, \nu, \frac{E_T}{E_r^*}\right), \tag{4}$$

where

$$\Pi_1^* := \frac{h_t}{h_y^*} = \frac{h_t}{R}\left(0.8\pi \frac{k_0}{E_r^*}\right)^{-2}, \tag{5}$$

$$h_y^* := h_y|_{\nu=0.3} = R\left(0.8\pi \frac{k_0}{E_r^*}\right)^2. \tag{6}$$

(Throughout the paper use is made of the notation introduced in [1]). The hardening properties are represented by two dimensionless quantities Π_1^* and E_T/E_r^*. The parameter E_T is the tangent modulus related to a linear isotropic hardening rule and E_r^* is the reduced modulus, which is composed of the elastic properties of the specimen and the indenter

$$\frac{1}{E_r^*} = \frac{1-\nu^2}{E} + \frac{1-\nu_i^2}{E_i}. \tag{7}$$

Also, the radius of the indenter is denoted by R and k_0 represents the yield stress at the beginning of plastic flow. In the case of nonlinear hardening, the functions $\tilde{\Pi}_2$ and $\tilde{\Pi}_3$ have to be extended by additional dimensionless hardening parameters. However, as it will be shown later, for the purpose of our paper it suffices to model the hardening response by a linear hardening rule.

We recall that some kind of correlation between Π_2 and Π_3 exists, which does not allow to invert, with sufficiently accuracy, the problem with respect to ν by using neural networks and only one unloading, even if h_y^* and m are supposed to be given. However, ν can be determined by using neural networks with high accuracy under the same suppositions, when at least two unloadings are used at different depths. The inversion of this problem has been realized in [1] by employing a neural network, denoted as Set 2. Such a neural network can be trained by using finite element simulations and is able to approximate a pointwise given function. During training, the neural network learns the relation between input data and the desired output data. In our case, the input data are represented by the quantities Π_1^*, Π_2 and Π_3, while ν is the desired output value. For the numerical simulations, ν and E_T/E_r^* have been chosen randomly from the intervals $[0.1, 0.45]$ and $[10^{-3}, 10^{-1}]$, respectively. Note, that the location of the unloadings has been chosen in equidistant steps, so that

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 26, 1999; final revision, Nov. 1, 2000. Associate Editor: K. T. Ramesh. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

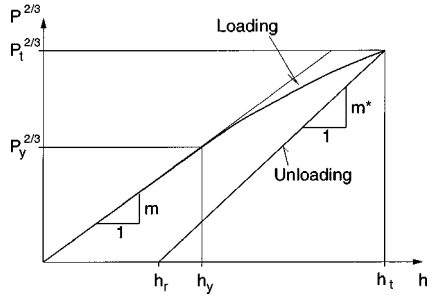


Fig. 1 Sketch of a spherical indentation depth-load response for elastic-plastic deformation

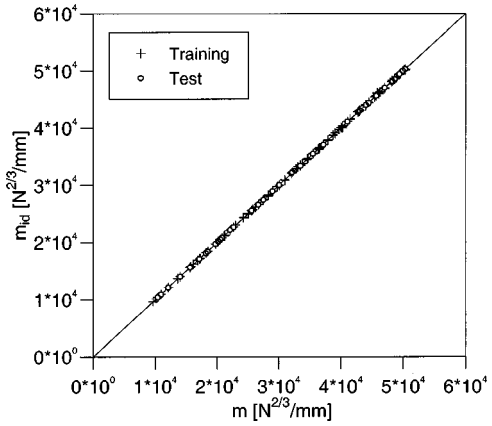


Fig. 2 Accuracy of the identified slope m_{id}

$\Pi_1^* \in \{1.5,2.0,\ldots,8.0\}$. For more details about the finite element simulations and the training of the neural networks see [1].

The present paper deals with the extention of the method for identifying ν for values of m and h_y^* which are not explicitly given and so to make the method more suitable to experiments. Following [1], we consider unloadings between the depths $h_{t,1} > h_y^*$ and $h_{t,2} > h_{t,1}$. The loading response $P_l(h)$ is regarded to be known for $0 \leqslant h \leqslant h_{t,2}$ and the unloading response $m^*(h_t)$ is regarded to be known for $h_{t,1} \leqslant h_t \leqslant h_{t,2}$, both in form of discrete finite element results.

In Section 2, the value of m will be determined by a neural network, which uses data of the loading response $P_l(h)$ only. In Section 3 two neural networks will be developed, which allow to assign to the discrete loading and unloading data $P_l(h)$ and $m^*(h_t)$ continuous distributions $\tilde{P}_l(h)$ and $\tilde{m}^*(h_t)$. Neural networks represent smooth functions and are able to ignore the scatter in the input data. Thus, the approximated values are physically meaningful and free of numerical or experimental scatter. As the transition from pure elastic to elastic plastic deformation is smooth the determination of the yield depth h_y or, which is the same, of $h_y^* := h_y|_{\nu=0.3}$ is by no means an easy task. To solve this problem we will utilize two different neural networks. The first one is an extended version of Set 2 introduced in [1] and has a high sensitivity to Π_1^* in the vicinity of h_y^* (cf. Fig. 12 in [1]). In order to avoid the difficulty to determine ν on the basis of h_y^*, a new concept will be followed here. The hypothesis is that it should be possible to train a second neural network, referred to as Set 3, which is almost insensitive to Π_1^* in the vicinity of h_y^*. Then, the identified $\tilde{h}_y^*$-$\tilde{\nu}$-distributions (cf. Section 4.2 in [1]) of both neural networks will intersect in the point (h_y^*, ν) where simultaneously the identification error vanishes. Finally, Section 4 deals with the verification of the method developed. It will be shown that the results are insensitive to nonlinear hardening as well as to the location of the unloadings.

2 Determination of m

In order to determine m, without knowing h_y, a neural network will be used, were the input data are restricted to values of the loading response at different depths.

The input definition consists of a sequence of dimensionless load ratios

$$x_i := \left(\frac{\overline{P}_l(h_i)}{P_l(h_{t,1})} \right)^{2/3}, \quad h_i := \frac{i}{10} h_{t,1}, \quad i = 1 \ldots 9, \tag{8}$$

where $h_{t,1} > h_y^*$ is the depth at which the first unloading is available. When identifying the value of ν from experimental data the yield depth h_y^* is not supposed to be known. Therefore, when training the network, the depth $h_{t,1}$ is chosen randomly in the interval $h_{t,1} \in [1.5h_y^*, 4.0h_y^*]$ for all patterns. For the purposes of this section it suffices to determine the input loading data $\overline{P}_l(h_i)$ by linear interpolating the data $P_l^{2/3}(h)$, which are given in discrete form, since the data $P_l(h)$ are discrete. The reason for using linear interpolated loading data $\overline{P}_l(h_i)$ in (8) arises from the fact that the neural network for determining m is not sensitive to these data.

The output value is defined by the dimensionless quantity

$$y := m \frac{h_{t,1}}{P_{t,1}^{2/3}}, \tag{9}$$

where the slope $P_{t,1}^{2/3}/h_{t,1}$ acts as an estimate of m in order to minimize the range of values for y. For all training patterns $y \in [0.99617, 1.28209]$. Note, that values $y < 1$ are theoretically not possible and reflect the numerical error of the finite element simulation. The exact value of the slope m is calculated analytically by

$$m = \left(\frac{4}{3} E_r^* \sqrt{R} \right)^{2/3}. \tag{10}$$

The creation of all patterns (see [1]) is based on (10), so that the neural network learns to ignore scatter due to numerical errors. In addition, the fit of the data is avoided, which would be a further source of uncertainties.

The neural network consists of nine input neurons, one output neuron as well as six and three neurons in the two hidden layers. After 2000 epochs, the 407 training and 43 test patterns reached an MSE value of $9.7 \cdot 10^{-6}$ and $6.5 \cdot 10^{-6}$, respectively. The quick training and the very low MSE values indicate an excellent determination of m by the neural network. This is proved by plotting the identified values m_{id} against the analytical values m given by (10) in Fig. 2.

3 Neural Networks for the Loading Data and the Slopes During Elastic Unloadings

In this section, two neural networks are provided, which will be used to prepare the input data for identifying ν. When determining the value of ν in Section 4, the used neural networks will be sensitive with respect to the input loading data as well as the input data concerning m^*. In this case it has been proved to be useful to work with $\tilde{P}_l(h)$ and $\tilde{m}^*(h_t)$ as input data. The distributions $\tilde{P}_l(h)$ and $\tilde{m}^*(h_t)$ are smooth and are obtained by approximating (in mathematical sense) via neural networks the discrete data $P_l(h)$ and $m^*(h_t)$.

3.1 Determination of $\tilde{P}_l^{2/3}$. In establishing the distribution $\tilde{P}_l^{2/3}(h)$ the input definitions in the neural network, in analogy to the last section, read

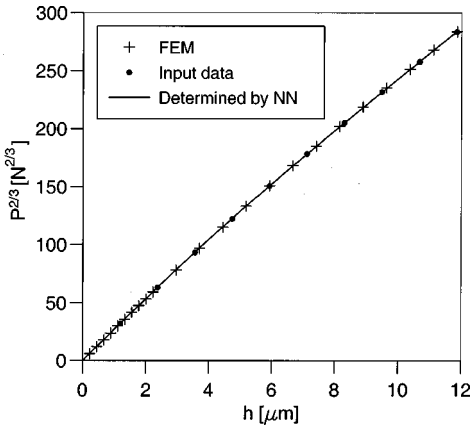


Fig. 3 Comparison of the $\tilde{P}_l(h)$ distribution with the $P_l(h)$ finite element method (FEM) data

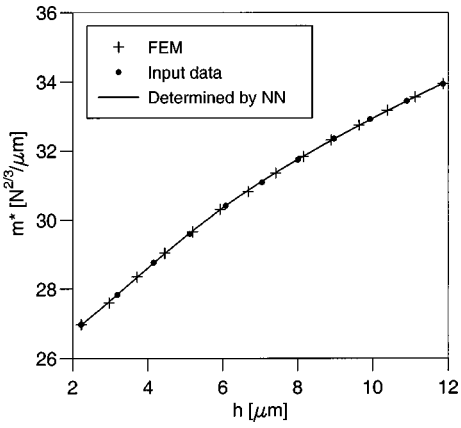


Fig. 4 Comparison of the smooth $\tilde{m}^*(h_t)$-distribution with the $m^*(h_t)$ finite element method (FEM) data

$$x_i := \left(\frac{\overline{P}_l(h_i)}{P_l(h_t)} \right)^{2/3}, \quad h_i := \frac{i}{10} h_t, \quad i = 1 \ldots 9, \tag{11}$$

$$x_{10} := \frac{h}{h_t}, \tag{12}$$

where $h \in [0, h_t]$ denotes that depth at which the load value is needed. The maximum depth h_t is allowed to be in the interval $h_t \in [1.5h_y^*, 8.0h_y^*]$. The output definition is

$$y := \left(\frac{\tilde{P}_l(h)}{P_l(h_t)} \right)^{2/3} \frac{h_t}{h}. \tag{13}$$

Again, due to the choice of y, the range of possible values is in a small interval $y \in [0.984, 1.287]$. This allows an approximation with a very high accuracy. The neural network consists of ten input neurons, one output neuron as well as eight and four neurons in the first and second hidden layer, respectively. The MSE value after 2000 epochs was $1.0 \cdot 10^{-4}$ and $4.8 \cdot 10^{-5}$ for 892 training and 92 test patterns, respectively. An example of the very good approximation is shown in Fig. 3.

3.2 Determination of $\tilde{m}^*$. For determining the unloading data $\tilde{m}^*(h_t)$, we define $h_{t,1}$ and $h_{t,2}$ to be the depths at which the first and the last unloading is carried out, respectively. The corresponding unloading slopes are denoted by $m_1^* := m^*(h_{t,1})$ and $m_2^* := m^*(h_{t,2})$. Again, the input values are given by equidistant linear interpolation of the numerical data

$$x_i := \frac{\overline{m}^*(h_i) - m_1^*}{m_2^* - m_1^*}, \quad h_i := \frac{i}{10}(h_{t,2} - h_{t,1}) + h_{t,1}, \quad i = 1 \ldots 9, \tag{14}$$

$$x_{10} := \frac{h_t - h_{t,1}}{h_{t,2} - h_{t,1}}. \tag{15}$$

Here, $\overline{m}^*(h_i)$ represent the slopes obtained by linear interpolating the discrete values $m^*(h_t)$, and $h_t \in [h_{t,1}, h_{t,2}]$ denotes that depth at which the value of m^* is needed. The maximum depth $h_{t,2}$ for the training patterns is in the interval $h_{t,2} \in [4h_y^*, 8h_y^*]$ while the minimum depth $h_{t,1}$ is always given by the location of the first unloading. The output is defined by

$$y := \frac{\tilde{m}^*(h_t) - m_1^*}{m_2^* - m_1^*}. \tag{16}$$

The neural network consists of ten input neurons, one output neuron as well as eight and four neurons in the first and second hidden layer, respectively. The MSE value after 2000 epochs was $3.8 \cdot 10^{-6}$ and $4.5 \cdot 10^{-6}$ for 2805 training and 258 test patterns, respectively. Figure 4 demonstrates the capabilities of the neural network used. The minimization of the numerical scatter by the filtering of the neural network is visible at those points where the finite element results deviate from the continuous line.

4 Identification of ν and h_y^*

The neural networks Set 1 and Set 2 in [1] are based on the assumption that the yield depth is known explicitly. However, the onset of plastic yield cannot be determined with the necessary accuracy from the loading data since the transition from pure elastic to elastic plastic deformation is not marked. In order to avoid such difficulties, a further neural network will be created. Here, in contrast to Set 1 and Set 2, the input data of this neural network, denoted as Set 3, will not make use explicitly from the value of Π_1^*.

4.1 Retraining of Set 2. The patterns of Set 2 in [1] were restricted to the exact locations of the unloadings available from the finite element simulations. By using the networks of Section 3, the unloading data are available at any depth between the first and last unloading. Thus, it is convenient to retrain the network Set 2 with a random depth $h_{t,1}$ in order to achieve best possible generalization. Also, in order to make the neural network more robust for practical problems, an additional third unloading is used between the two original unloadings. The so obtained neural network for determining the value of ν represents the first neural network, mentioned in the Introduction 1, which indicates a high sensitivity to Π_1^* in the vicinity of h_y^*.

The modified input definition is then

$$x_1 := \Pi_1^*|_{h_{t,1}} \tag{17}$$

$$x_2 \ldots x_4 := \Pi_2|_{h_t}, \quad h_t = 1.0h_{t,1},\ 1.5h_{t,1},\ 2.0h_{t,1}, \tag{18}$$

$$x_5 \ldots x_7 := \Pi_3|_{h_t}, \quad h_t = 1.0h_{t,1},\ 1.5h_{t,1},\ 2.0h_{t,1}, \tag{19}$$

where $h_{t,1}$ is allowed to be in the interval $h_{t,1} \in [1.5h_y^*, 4.0h_y^*]$.

The neural network consists now of seven input neurons, two hidden layers with five and three neurons, respectively, and one output neuron. Now, 449 training patterns and 46 test patterns were generated. After a duration of 3000 epochs a MSE value of $1.0 \cdot 10^{-4}$ and $6.8 \cdot 10^{-5}$ has been achieved for training and test patterns, respectively. Here, the test patterns display the double precision compared to [1] indicating the minimization of the scatter by the filtering and the increased reliability due to the third unloading.

4.2 Training of Set 3. We recall (see end of Section 3 in [1]) that ν can be thought of as a function of Π_1^*, Π_2, Π_3, i.e,

$$\nu = g(\Pi_1^*, \Pi_2, \Pi_3) \tag{20}$$

or

$$\nu - g(\Pi_1^*, \Pi_2, \Pi_3) =: f(\Pi_1^*, \Pi_2, \Pi_3, \nu) = 0. \tag{21}$$

In view of (3)–(5), the latter becomes

$$f\left(\left.\frac{h_t}{h_y}\right|_{\nu=0.3}, \tilde{\Pi}_2\left(\left.\frac{h_t}{h_y}\right|_{\nu=0.3}, \nu, \frac{E_T}{E_r^*}\right), \tilde{\Pi}_3\left(\left.\frac{h_t}{h_y}\right|_{\nu=0.3}, \nu, \frac{E_T}{E_r^*}\right), \nu\right) = 0 \tag{22}$$

or

$$f\left(\frac{h_t}{h_y^*}, \tilde{\Pi}_2\left(\frac{h_t}{h_y^*}, \nu, \frac{E_T}{E_r^*}\right), \tilde{\Pi}_3\left(\frac{h_t}{h_y^*}, \nu, \frac{E_T}{E_r^*}\right), \nu\right) = 0. \tag{23}$$

Now, our aim is to obtain from Eq. (23) a relation of the form

$$\nu = \tilde{\tilde{\nu}}(\tilde{h}_y^*), \tag{24}$$

where $\tilde{h}_y^*$ is defined below. To this end we first consider a sequence of indentation depths

$$h_{t,j} := \frac{j+1}{2} h_{t,1}, \quad j = 1, \ldots, n, \tag{25}$$

$$h_{t,1} := 2h_y^*. \tag{26}$$

Theoretically, it suffices to consider Eq. (23), e.g., at $h_t = h_{t,1}$, $h_t = h_{t,2} = \frac{3}{2} h_{t,1}$ and $h_t = h_{t,3} = 2h_{t,1}$ in order to eliminate E_T/E_r^* and h_y^* and so, by using a neural network, to obtain the value of ν for the given value $h_{t,1}$. However, in order to increase the accuracy of the neural network developed, Eq. (23) has been exploited at $n = 5$ indentation depths $h_{t,1}$, $h_{t,2} = 3/2 h_{t,1}, \ldots, h_{t,5} = 3h_{t,1}$.

Evidently, one may use this approach for $h_{t,1} \in [2h_y^* - \epsilon, 2h_y^* + \epsilon]$, where ϵ is a small positive real number, and so to establish, by using an appropriately trained neural network, referred to as Set 3, a function of the form

$$\nu = \overline{\nu}(h_{t,1}). \tag{27}$$

On the other hand, one may interpret, on the basis of Eq. (23), a variation of $h_{t,1}$ as a variation of h_y^*. Thus, by setting $\tilde{h}_y^* = h_{t,1}/2$ in (27), we have

$$\nu = \overline{\nu}(2\tilde{h}_y^*) =: \tilde{\tilde{\nu}}(\tilde{h}_y^*), \tag{28}$$

with $\tilde{\tilde{\nu}}$ in (24) and (28) being understood as the same function.

In training Set 3, one has to evaluate, for given material data E_T/E_r^*, the values of Π_2 and Π_3 at prescribed h_t. These values are obtained from finite element results by using Eqs. (1), (2) and represent the input data for Set 3, the output quantity being ν. Note that in a training pattern the number of Π_2-inputs must not necessarily be equal to the number of the Π_3-inputs. Thus the definitions, which produced the best results, are for given E_T/E_r^*

$$x_1 \ldots x_3 := \Pi_2|_{h_t = h}, \quad h = 0.5\tilde{h}_y^*, 1.0\tilde{h}_y^*, 1.5\tilde{h}_y^*, \tag{29}$$

$$x_4 \ldots x_8 := \Pi_2|_{h_t = h}, \quad h = 2\tilde{h}_y^*, 3\tilde{h}_y^*, \ldots, 6\tilde{h}_y^*, \tag{30}$$

$$x_9 \ldots x_{13} := \Pi_3|_{h_t = h}, \quad h = 2\tilde{h}_y^*, 3\tilde{h}_y^*, \ldots, 6\tilde{h}_y^*. \tag{31}$$

The 443 training and 48 test patterns were generated by using five randomly chosen values $\tilde{h}_y^*$ for each finite element simulation. After 45,000 epochs an MSE value of $9.5 \cdot 10^{-5}$ for training and $8.9 \cdot 10^{-5}$ for test patterns has been achieved.

5 Results and Discussion

The characteristic of Set 3 is shown in Fig. 5 for the same example as depicted in Fig. 12 in [1]. Actually, the neural network

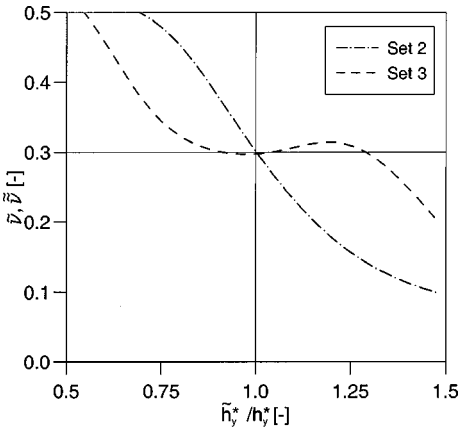


Fig. 5 Simultaneous identification of Poisson's ratio ν and h_y^* as intersection of Set 2 and Set 3

Set 3 displays an insensitivity on h in the vicinity of h_y^*, which is visible as a plateau. From the intersection of the curves $\tilde{\nu}(\tilde{h}_y^*)$ and $\tilde{\tilde{\nu}}(\tilde{h}_y^*)$, the values of ν and the value of h_y^* can be determined simultaneously.

In the following, this identification method will be applied to all patterns in order to give an impression about the theoretical approach. Next, the sensitivity of the results will be checked with respect to different nonlinear hardening rules as well as different unloading locations.

5.1 Linear Hardening. Using the identification method described above, the values of ν and h_y^* are determined for all available finite element simulations including the test patterns (see Figs. 6 and 7). Since the test patterns do not coincide for the different neural networks, training and test patterns are not distinguished in the graphs. From the 94 examples displayed in Fig. 6 a high accuracy can be observed, where 95 percent of the results possess an error less than 5 percent. The successful identification is confirmed as well by the excellent accuracy of the identified h_y^* values, displayed in Fig. 7.

5.2 Nonlinear Hardening. So far the identification method is developed for the case of a linear hardening response. However, since the indentation depths required by the method are of the order of the yield depth it is natural to expect that, when determining the value of ν for a material with nonlinear hardening, only the slope of the hardening response at the beginning of plas-

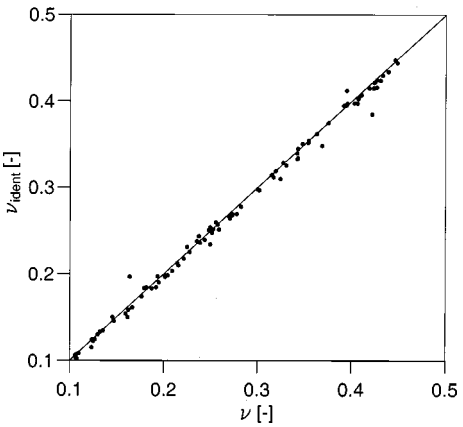


Fig. 6 Accuracy of the identified Poisson's ration ν

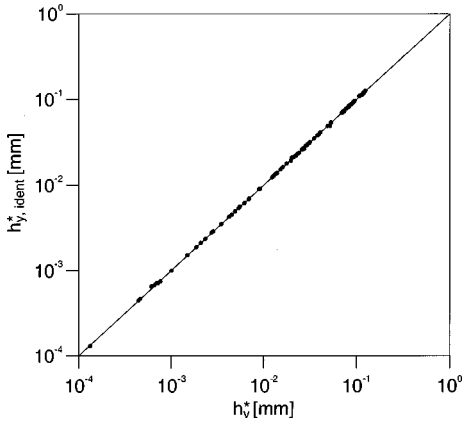


Fig. 7 Accuracy of the identified yield depth h_y^*

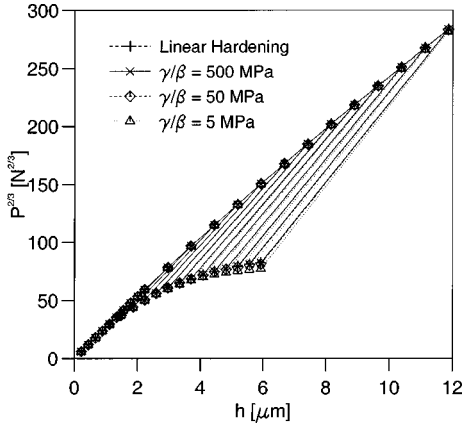


Fig. 9 Depth-load trajectories according to the hardening rules displayed in Fig. 8

tic yield will be important. To demonstrate this, some finite element simulations have been carried out, where a hardening rule

$$\dot{k}=(\gamma-\beta(k-k_0))\dot{s} \tag{32}$$

is assumed. In Eq. (32) k denotes the isotropic hardening variable. For more details on the constitutive equations see Huber and Tsakmakis [2,3] and the literature cited there. Note, that (32) reduces to linear hardening for $\beta=0$ and $\gamma=E_T$. For $\beta>0$, the initial slope of the hardening rule is given by γ while $k_0+\gamma/\beta$ is the limit of isotropic hardening.

In Fig. 8 one-dimensional strain-stress distributions are displayed for $E=210$ GPa, $\nu=0.3$, $k_0=250$ MPa, and $\gamma=10$ GPa. Four different hardening rules are considered, where the hardening limit is successively decreased from infinite (linear hardening) to 5 MPa. The resulting depth-load trajectories are displayed in Fig. 9.

In the depth range of $h_t/h_y^*\leqslant 8$, considered, only the depth-load trajectory for $\gamma/\beta=5$ MPa deviates slightly from the other trajectories with higher hardening amount. The displayed depth-load trajectories are generated by using the same unloading locations as for all training patterns, i.e., $h_t/h_y^*=1.5,2.0,\ldots,8.0$ and are denoted by Type 1. Another type of depth-load trajectories, denoted as Type 2 has been generated with the locations of the unloadings at $h_t/h_y^*=1.75,2.25,\ldots,7.75$. By using the Type 2 samples, the approximation capabilities of the neural networks can be checked with respect to the location of the unloadings.

First it can be seen from Table 1, that the value of m is identified with an excellent accuracy and is not affected by the nonlinear hardening or the location of the load increments. This robustness is necessary for the following identification since m is needed for making other quantities non-dimensional.

Next, the eight verification examples will be used to discuss the sensitivity of the interpolation networks as it has been indicated in Section 3. To this end, the neural network Set 3 has been trained by three different data sets. In the first case (LL), the loading data as well as the unloading data are created by linear interpolation. In the second case (LN), the loading data have been prepared by linear interpolation while the unloading data have been determined by using a neural network (see Section 3.2). In the third case (NN), the loading data as well as the unloading data are determined by neural networks (see Section 3.1, 3.2). From each database a different neural network Set 3 has been trained under the same conditions, where the training of Set 3 for case (NN) is described in Section 4.2.

The identified values of h_y^* and ν for all eight verification examples are given in Table 1. These results for case (NN) show that there is only a very weak but systematic effect for increasing nonlinearity on the identified value of ν. From linear hardening to almost ideal plasticity, the determined Poisson's ratio increases by approximately 0.5 percent. No effect can be observed in relation to the unloading locations.

If the loading response is interpolated linear (LN), the values of ν for Type 1 are slightly higher while for Type 2 they are slightly smaller. Here we find a remarkable difference between Type 1 and Type 2 examples (i.e., an effect to the location of the unloadings), but no significant effect with respect to nonlinear hardening.

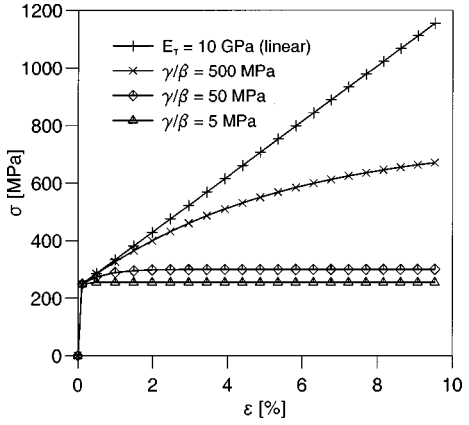


Fig. 8 Stress-strain curves for linear hardening and different nonlinear hardening rules

Table 1 Identification of additional verification patterns. Four different model-materials according to Fig. 8 have been considered, which differ with respect to the hardening response. The four model-materials exhibit common elasticity data, onset of plastic yield and slope of the hardening response at onset of plastic yielding: $m=26.65\,\mathrm{N}^{2/3}\,\mu\mathrm{m}^{-1}$, $h_y^*=1.48\,\mu$m, $\nu=0.3$.

			Set 3(LL)		Set 3(LN)		Set 3(NN)	
Type	β [−]	m [$\mathrm{N}^{2/3}\,\mu\mathrm{m}^{-1}$]	h_y^* [μm]	ν [−]	h_y^* [μm]	ν [−]	h_y^* [μm]	ν [−]
1	0	26.61	1.77	0.175	1.49	0.303	1.49	0.297
1	20	26.60	1.64	0.235	1.49	0.302	1.49	0.301
1	200	26.60	1.49	0.311	1.48	0.312	1.48	0.307
1	2000	26.61	1.46	0.320	1.46	0.319	1.47	0.315
2	0	26.61	1.62	0.241	1.53	0.284	1.50	0.297
2	20	26.61	1.63	0.238	1.53	0.286	1.49	0.300
2	200	26.61	1.50	0.302	1.52	0.292	1.49	0.306
2	2000	26.62	1.43	0.334	1.50	0.303	1.49	0.311

For linear interpolation of the loading as well as the unloading data (LL) a strong effect of the nonlinear hardening can be observed, so that the determined value of ν increases for increasing nonlinearity.

6 Conclusions

In the present paper a method for identifying the Poisson's ratio is proposed, where no information about the yield depth is supposed to be given. This has been realized by training two neural networks which possess a different sensitivity in the point where the identification error vanishes.

An exceptional point of the method developed is that neither any property of the spherical indenter nor one of the specimen has to be known. The only suppositions are: the mechanical behavior of the material can be modeled by metal plasticity and the maximum indentation depth is not higher than eight times of the yield depth. Note that the latter condition is attributed to the limits of finite element simulations, carried out, and is not a restriction in general.

The neural networks for preparing of the input data are also useful to reduce the scatter of experimental data. However, this work is concerned with the theoretical problems when identifying Poisson's ratio. It cannot be precluded that additional experimental errors, like, e.g., uncertainty of zero depth, roughness, machine compliance, or a nonideal spherical indenter may occur. Therefore, further difficulties may appear, when dealing with experimental data. In this sense the present paper should be regarded as the first step towards determining the Poisson's ratio for real materials.

Acknowledgments

The first author gratefully acknowledges financial support from the Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren.

References

[1] Huber, N., and Tsakmakis, Ch., 2001, ''Determination of Poisson's Ratio by Spherical Indentation Using Neural Networks. Part I: Theory,'' ASME J. Appl. Mech., **68**, pp. 218–223.

[2] Huber, N., and Tsakmakis, Ch., 1999, ''Determination of Constitutive Properties From Spherical Indentation Data Using Neural Networks, Part I: The Case of Pure Kinematic Hardening in Plasticity Laws,'' J. Mech. Phys. Solids, **47**, pp. 1569–1588.

[3] Huber, N., and Tsakmakis, Ch., 1999, ''Determination of Constitutive Properties From Spherical Indentation Data Using Neural Networks, Part II: Plasticity With Nonlinear Isotropic and Kinematic Hardening,'' J. Mech. Phys. Solids, **47**, pp. 1589–1607.

C. W. Bert
Perkinson Chair Professor,
Life Fellow ASME

H. Zeng
Graduate Research Assistant

School of Aerospace and Mechanical Engineering,
The University of Oklahoma,
Norman, OK 73019-1052

Generalized Bending of a Large, Shear Deformable Isotropic Plate Containing a Circular Hole or Rigid Inclusion

The problem of a large isotropic plate with a circular hole or rigid circular inclusion is considered here. The plate experiences transverse shear deformation and is subjected to an arbitrary bending field. By using Reissner's plate theory, a general solution, in terms of Poisson's ratio ν, a geometric ratio, and bending moment ratio B, is obtained to satisfy both the boundary conditions along the edge and at great distances from the edge. The stress couple concentration factors are calculated and compared with classical plate theory, three-dimensional elasticity theory, higher-order plate theory, and an experimental result. [DOI: 10.1115/1.1348014]

Introduction

The problem of the bending stress concentration for a circular hole in a large, thin isotropic plate subjected to a bending stress field was discussed by Bickley [1]. He considered the uniaxial bending and cylindrical bending cases. Without reference to the work of Bickley, the uniaxial bending problem was reconsidered by Lekhnitskii [2] and the cylindrical bending problem was reworked by Goodier [3]. They obtained explicit expressions for the stress concentration factor *K*. Goodier also first solved the pure twisting problem. The research was generalized by Bert [4] to obtain a general equation for *K* as

$$K=\frac{5+3\nu}{3+\nu}+\frac{1-\nu}{3+\nu}B \tag{1}$$

where ν is Poisson's ratio and $B\equiv M_y/M_x$ is defined to be the ratio of principal bending moments. For uniaxial bending, cylindrical bending, pure twisting, and balanced biaxial bending, $B=0,\ \nu,\ -1,\ 1$, respectively.

Goland [5] was the first to consider the effect of a rigid circular inclusion on plate bending. He considered the cases of uniaxial bending, cylindrical bending, and pure twisting.

The importance of transverse shear deformation effects on stress concentration factor was first recognized by Reissner [6]. He considered the problem of an infinite plate with a circular hole and subjected to uniaxial bending and pure twisting by allowing three boundary conditions along an edge, taking into account the effect of transverse shear deformation, and assuming that bending stresses are distributed linearly over the thickness of the plate. The case of an infinite shear deformable plate with a circular hole and a rigid circular inclusion subjected to cylindrical bending was treated in his later paper ([7]). Cheng [8] considered the problem of a large plate with a circular hole subjected to uniaxial bending by means of his refined plate theory, which was deduced from three-dimensional elasticity and took into account the shear deformation. Chen and Archer [9] reconsidered this problem by using their 12-order thick plate theory.

The case of an infinite plate with a rigid circular inclusion subjected to uniaxial bending was complimented by Hirsch [10] by using Reissner's plate theory.

The purpose of the present work is to generalize to an *arbitrary bending field* the previous research on the stress concentration factor for large, shear deformable isotropic plates containing a circular hole or rigid inclusion, by using Reissner's plate theory. Comparisons are made with classical plate theory, three-dimensional elasticity theory by Alblas [11], Cheng's ([8]) results, Chen and Archer's ([9]) results, and Dumont's ([12]) experimental result.

Governing Differential Equations

Reissner [13] derived the expressions for stress couples *M* and transverse shear stress resultants *Q* for a uniform isotropic shear deformable plate (with absence of surface load) in polar coordinates (r,θ) as

$$M_r=-D\left(\varpi_{,rr}+\nu\frac{\varpi_{,r}}{r}+\nu\frac{\varpi_{,\theta\theta}}{r^2}\right)+\frac{2}{\lambda^2}\left(\frac{\chi_{,\theta r}}{r}-\frac{\chi_{,\theta}}{r^2}\right) \tag{2a}$$

$$M_\theta=-D\left(\nu\varpi_{,rr}+\frac{\varpi_{,r}}{r}+\frac{\varpi_{,\theta\theta}}{r^2}\right)-\frac{2}{\lambda^2}\left(\frac{\chi_{,\theta r}}{r}-\frac{\chi_{,\theta}}{r^2}\right) \tag{2b}$$

$$M_{r\theta}=-(1-\nu)D\left(\frac{\varpi_{,\theta r}}{r}-\frac{\varpi_{,\theta}}{r^2}\right)-\chi+\frac{2}{\lambda^2}\left(\frac{\chi_{,r}}{r}-\frac{\chi_{,\theta\theta}}{r^2}\right) \tag{2c}$$

$$Q_r=-D(\nabla^2\varpi)_{,r}+\frac{\chi_{,\theta}}{r}, \tag{2d}$$

$$Q_\theta=-D\frac{(\nabla^2\varpi)_{,\theta}}{r}-\chi_{,r}, \tag{2e}$$

$$\phi_r=-\varpi_{,r}+\frac{2\lambda^{-2}}{D(1-\nu)}\frac{\chi_{,\theta}}{r}, \tag{2f}$$

$$\phi_\theta=-\frac{\varpi_{,\theta}}{r}-\frac{2\lambda^{-2}}{D(1-\nu)}\chi_{,r}, \tag{2g}$$

where ν is Poisson's ratio, ϕ_r and ϕ_θ are bending slopes,

$$\varpi=w+\frac{2}{\lambda^2}\frac{\nabla^2 w}{1-\nu}, \tag{3a}$$

$$\frac{1}{\lambda^2}=\frac{1-\nu}{2}SD \tag{3b}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 6, 1999; final revision, June 27, 2000. Associate Editor: A. K. Mal. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

with D and S being bending stiffness and transverse shear flexibility, w the normal deflection, and ϖ and χ being solutions of differential equations

$$D\nabla^4 w=0, \tag{4}$$

$$\nabla^2\chi-\lambda^2\chi=0. \tag{5}$$

For three-dimensionally homogeneous plates, the factors D, S, and λ are

$$D=\frac{Eh^3}{12(1-\nu^2)}, \tag{6a}$$

$$S=\frac{12(1+\nu)}{5Eh}, \tag{6b}$$

$$\lambda=\frac{\sqrt{10}}{h} \tag{6c}$$

where E is the Young's modulus.

Boundary Conditions

There are two kinds of boundaries involved in this problem. The conditions at infinity are arbitrary biaxial bending moments, i.e., finite M_x and $Q_x=0$ at $\theta=0$ deg and finite $M_y=BM_x$ and $Q_y=0$ at $\theta=\pi/2$. Then converting from rectangular Cartesian coordinates to plane polar coordinates yields

$$M_r=\frac{M_x}{2}[(1+B)+(1-B)\cos(2\theta)];$$
$$M_{r\theta}=-\frac{M_x}{2}(1-B)\sin(2\theta); \quad Q_r=0. \tag{7}$$

The boundary conditions at $r=a$ for a hole are

$$M_r=0, \quad M_{r\theta}=0, \quad Q_r=0 \tag{8}$$

and for a rigid inclusion are

$$w=0, \quad \phi_r=0, \quad \phi_\theta=0. \tag{9}$$

Solution

The solution of Eq. (4) for an infinite plate with a circular hole or inclusion and subjected to generalized bending at great distances from the hole or inclusion has the same form as for the analogous problem of a thin plate with a hole (Bert [4]):

$$w(r,\theta)=c_1+c_2\ln\left(\frac{r}{a}\right)+c_3\frac{r^2}{a^2}+c_4\frac{r^2}{a^2}\ln\left(\frac{r}{a}\right)+\left(d_1\frac{r^4}{a^4}+d_2\frac{a^2}{r^2}+d_3\frac{r^2}{a^2}+d_4\right)\cos(2\theta). \tag{10}$$

The general solution for χ can be written as (Abramovich and Stegun [14])

$$\chi(r,\theta)=\sum_{m=0}^{\infty}\{[A_mI_m(\lambda r)+C_mK_m(\lambda r)]\cos(m\theta)+[B_mI_m(\lambda r)+D_mK_m(\lambda r)]\sin(m\theta)\} \tag{11}$$

where I_m and K_m are the first and second kinds of the modified Bessel functions of order m.

In conjunction with the far-distance boundary conditions, Eq. (7), solutions for $w(r,\theta)$ and $\chi(r,\theta)$ have the form of

$$w(r,\theta)=c_1+c_2\ln\left(\frac{r}{a}\right)+c_3\frac{r^2}{a^2}+\left(d_2\frac{a^2}{r^2}+d_3\frac{r^2}{a^2}+d_4\right)\cos(2\theta) \tag{12}$$

$$\chi(r,\theta)=A_0I_0(\lambda r)+C_0K_0(\lambda r)+\left[\begin{Bmatrix}A_2\\B_2\end{Bmatrix}I_2(\lambda r)+\begin{Bmatrix}C_2\\D_2\end{Bmatrix}K_2(\lambda r)\right]\times\begin{Bmatrix}\cos(2\theta)\\\sin(2\theta)\end{Bmatrix}. \tag{13}$$

It is noted that the coefficients c_4, d_1, A_1, $A_m(m>2)$, $B_m(m\neq 2)$, C_1, $C_m(m>2)$, and $D_m(m\neq 2)$ must vanish in order to satisfy the far-distance boundary conditions.

Substitution of Eqs. (12) and (13) into boundary condition, Eqs. (7), yields $A_0=A_2=B_2=0$, and

$$c_3=-\frac{M_xa^2}{4D}\frac{(1+B)}{(1+\nu)}; \tag{14a}$$

$$d_3=-\frac{M_xa^2}{4D}\frac{(1-B)}{(1-\nu)}. \tag{14b}$$

Solution for Plate With a Circular Hole

At boundary $r=a$, the conditions are listed as follows:

$\cos(2\theta)$	$\sin(2\theta)$	1	Boundary Conditions	Eq.
d_2, d_4, B_2, D_2	A_2, C_2	c_2	$M_r=0$	(15a)
A_2, C_2	d_2, d_4, B_2, D_2	A_0, C_0	$M_{r\theta}=0$	(15b)
d_4, B_2, D_2	A_2, C_2	⋯	$Q_r=0$	(15c)

c_2, d_2, d_4, and D_2 can be solved as

$$N_p\equiv\nu+p+\frac{8}{\mu^2}+\frac{4}{\mu}\frac{K_2'(\mu)}{K_2(\mu)} \tag{16}$$

$$c_2=-\frac{M_xa^2}{2D}\frac{(1+B)}{(1-\nu)}; \tag{17a}$$

$$d_2=-\frac{M_xa^2}{4D}\frac{N_{-1}(1+B)}{N_3(1-\nu)} \tag{17b}$$

$$d_4=-\frac{M_xa^2}{2D}\frac{1-B}{N_3}; \tag{17c}$$

$$D_2=-\frac{2M_x}{K_2(\mu)}\frac{(1-B)}{N_3} \tag{17d}$$

where $\mu=\lambda a=\sqrt{10}a/h$.

The nonzero coefficients in solutions, Eqs. (12) and (13), for a shear deformable plate with a circular hole are c_2, c_3, d_2, d_3, d_4, D_2, as in Eqs. (14) and (17), and c_1, which can be any constant. Therefore the solution for $\chi(r,\theta)$ can be further simplified as

$$\chi(r,\theta)=D_2K_2(\lambda r)\sin(2\theta). \tag{18}$$

Solution for Plate With a Rigid Inclusion

$\cos(2\theta)$	$\sin(2\theta)$	1	Boundary Conditions	Eq.
d_2, d_4	⋯	c_1	$w=0$	(19a)
d_2, d_4, B_2, D_2	A_2, C_2	C_2	$\phi_r=0$	(19b)
A_2, C_2	d_2, d_4, B_2, D_2	A_0, C_0	$\phi_\theta=0$	(19c)

c_1, c_2, d_2, d_4, and D_2 can be solved as

$$P_\alpha=\nu-1+\alpha\left(\frac{8}{\mu^2}+\frac{16}{\mu^3}\frac{K_2(\mu)}{K_2'(\mu)}\right) \tag{20}$$

$$c_1=\frac{M_xa^2}{4D}\frac{(1+B)}{(1+\nu)}; \tag{21a}$$

$$c_3=\frac{M_x a^2}{2D}\frac{(1+B)}{(1+\nu)}; \tag{21b}$$

$$d_2=-\frac{M_x a^2}{4D}\frac{(1-B)P_1}{(1-\nu)P_{-1}}; \tag{21c}$$

$$d_4=-\frac{M_x a^2}{2D}\frac{(1-B)}{P_{-1}}; \tag{21d}$$

$$D_2=\frac{4M_x(1-B)}{\mu K_2'(\mu)P_{-1}}. \tag{21e}$$

The nonzero coefficients in solutions, Eqs. (12) and (13), for a shear deformable plate with a rigid inclusion are c_1, c_2, c_3, d_2, d_3, d_4, and D_2, as in Eqs. (17) and (21). The solution for $\chi(r,\theta)$ can be simplified to the same form as for the circular case in Eq. (18), but with a different value for coefficient D_2 in Eq. (21*e*).

Stress Concentration Factor *K*

To determine the stress concentration factor, defined by $K=M_{\max}/M_x$, Eqs. (12) and (13) are applied to obtain

$$K_H=\frac{M_\theta\left(a,\frac{\pi}{2}\right)}{M_x}=\frac{4+2\nu(1-B)+(1+B)N_1}{N_3} \tag{22a}$$

$$K_R=\frac{M_r(a,0)}{M_x}=\frac{-2+2(1+2\nu)B-(1+B)N_1}{(1+\nu)P_{-1}} \tag{22b}$$

with subscripts H and R standing for hole and rigid inclusion, respectively.

By taking $B=\nu$, which corresponds to cylindrical bending, the stress concentration factors for plates with circular hole or rigid inclusion are

$$K_H=(1+B)\left[1+\frac{2-2\nu}{N_3}\right]; \tag{23a}$$

$$K_R=1+\frac{2\nu-2}{P_{-1}} \tag{23b}$$

which agree with Eqs. (20) and (28) in Reissner's paper ([7]).

By taking $B=0$ and $K_2'(\mu)=-K_1(\mu)-2/\mu K_2(\mu)$ in Eq. (22*b*), one can obtain Hirsch's results ([10]).

For $\mu\to\infty$, we obtain

$$K_H(\infty)=\frac{5+3\nu}{\nu+3}+\frac{1-\nu}{\nu+3}B; \tag{24a}$$

$$K_R(\infty)=\frac{3+\nu-B(1+3\nu)}{(1-\nu^2)}. \tag{24b}$$

By taking $B=0$, ν, -1 in Eq. (24*b*), one can get Goland's results ([5]).

For $\mu\to 0$,

$$\frac{8}{\mu^2}+\frac{16}{\mu^3}\frac{K_2(\mu)}{K_2'(\mu)}\to 2; \tag{25a}$$

$$\frac{8}{\mu^2}+\frac{4}{\mu}\frac{K_2'(\mu)}{K_2(\mu)}\to -2. \tag{25b}$$

The introduction of Eqs. (25) into Eqs. (22) gives

$$K_H(0)=3-B; \tag{26a}$$

$$K_R(0)=\frac{5+\nu+B(1-3\nu)}{(1+\nu)(3-\nu)}. \tag{26b}$$

Now we are ready to make further comparisons. In Table 1, the stress concentration factors for a hole are compared with the exact three-dimensional elastic analysis by Alblas [11], classical thin plate theory, Cheng's ([8]) results, and Chen and Archer's ([9]) results. By no means do the present results deviate from Reissner's ([7]) results since the expressions have been compared previously to show the consistency. Slight differences between the results appear in Table 1. Often in practice, a simple method of solution, although only approximate, is preferable. Reissner's sixth-order theory is relatively practical, comparing with Chen and Archer's 12-order theory, Cheng's plate theory combining fourth-order, second-order, and transcendental equations, or the most complicated one, three-dimensional elasticity.

Table 1 Stress concentration factors for an infinite plate ($\nu=1/4$) subjected to uniaxial bending

a/h	With a circular hole ($B=0$)			
	Exact	Present	Chen & Archer	Cheng
0	-	3.0	3.0	3.0
0.5	2.268	2.242	2.264	2.245
1	2.045	2.038	2.050	2.040
1.5	1.960	1.956	1.964	1.957
2	1.914	1.912	1.918	1.913
2.5	1.896	1.884	1.890	1.885
∞	1.769	1.769	-	1.769
Classical thin plate theory				1.769

Dumont [12] measured the stress concentration factor for an aluminum-alloy plate (with $\nu=1/3$), 55 inches square, 1.062 inches thick, containing a circular hole 8 inches in diameter, and subjected to uniaxial bending loading, with a value of 1.85. By means of the present method, a stress concentration factor of 1.878 is obtained, which is very close to the experimental result.

As a consequence, Table 2 lists stress concentration factors for various values of ν and B, for plates with a circular hole or a rigid inclusion.

Table 2 Stress concentration factors for an infinite plate with a circular hole or rigid inclusion

a/h	With a circular hole				With a circular rigid inclusion			
ν = 0	**B = 0**	**ν**	**-1**	**1**	**B = 0**	**ν**	**-1**	**1**
0	3.0	3.0	4.0	2.0	1.667	1.667	1.333	2.0
0.5	2.135	2.135	2.270	2.0	2.057	2.057	2.114	2.0
1	1.927	1.927	1.853	2.0	2.418	2.418	2.837	2.0
1.5	1.845	1.845	1.691	2.0	2.629	2.629	3.258	2.0
2	1.803	1.803	1.605	2.0	2.749	2.749	3.498	2.0
2.5	1.776	1.776	1.552	2.0	2.821	2.821	3.642	2.0
∞	1.667	1.667	1.333	2.0	3.0	3.0	4.0	2.0
ν = 1/4	**B = 0**	**ν**	**-1**	**1**	**B = 0**	**ν**	**-1**	**1**
0	3.0	2.75	4.0	2.0	1.527	1.546	1.455	1.6
0.5	2.242	2.182	2.485	2.0	2.018	1.913	2.435	1.6
1	2.038	2.029	2.076	2.0	2.524	2.293	3.448	1.6
1.5	1.956	1.967	1.911	2.0	2.845	2.534	4.091	1.6
2	1.912	1.934	1.823	2.0	3.038	2.679	4.477	1.6
2.5	1.884	1.913	1.769	2.0	3.158	2.768	4.715	1.6
∞	1.769	1.827	1.539	2.0	3.467	3.0	5.333	1.6
ν = 1/3	**B = 0**	**ν**	**-1**	**1**	**B = 0**	**ν**	**-1**	**1**
0	3.0	2.667	4.0	2.0	1.5	1.5	1.5	1.5
0.5	2.272	2.182	2.545	2.0	2.033	1.855	2.566	1.5
1	2.070	2.047	2.140	2.0	2.607	2.238	3.715	1.5
1.5	1.988	1.992	1.976	2.0	2.986	2.491	4.472	1.5
2	1.944	1.963	1.888	2.0	3.219	2.646	4.937	1.5
2.5	1.916	1.944	1.833	2.0	3.365	2.7435	5.229	1.5
∞	1.8	1.8667	1.6	2.0	3.75	3.0	6.0	1.5
ν = 1/2	**B = 0**	**ν**	**-1**	**1**	**B = 0**	**ν**	**-1**	**1**
0	3.0	2.5	4.0	2.0	1.467	1.4	1.6	1.333
0.5	2.326	2.163	2.652	2.0	2.103	1.718	2.873	1.333
1	2.128	2.064	2.257	2.0	2.864	2.099	4.395	1.333
1.5	2.047	2.023	2.094	2.0	3.415	2.374	5.496	1.333
2	2.003	2.001	2.005	2.0	3.775	2.554	6.216	1.333
2.5	1.975	1.988	1.950	2.0	4.010	2.671	6.686	1.333
∞	1.857	1.929	1.714	2.0	4.667	3.0	8.0	1.333
ν = -1/2	**B = 0**	**ν**	**-1**	**1**	**B = 0**	**ν**	**-1**	**1**
0	3.0	3.5	4.0	2.0	2.571	1.857	1.143	4.0
0.5	1.792	1.688	1.584	2.0	2.836	2.254	1.672	4.0
1	1.603	1.404	1.206	2.0	3.047	2.571	2.094	4.0
1.5	1.536	1.304	1.072	2.0	3.158	2.736	2.315	4.0
2	1.502	1.253	1.004	2.0	3.217	2.825	2.434	4.0
2.5	1.482	1.222	0.963	2.0	3.251	2.877	2.503	4.0
∞	1.4	1.1	0.8	2.0	3.333	3.0	2.667	4.0

Discussion and Conclusion

It is noted that the stress concentration for a thick plate with a circular hole for large values of μ, Eq. (24*a*), converges to the result, Eq. (1), by means of classical thin plate theory. So does the stress concentration factor for a thick plate with a circular rigid inclusion for large value of μ, Eq. (24*b*). It is worth noting that result (26a) coincides with the stress concentration factor for a large, *thin*, isotropic plate with a circular hole subjected to an *in-plane* stress field. Actually, ''in the limit of a very thick plate (large plate thickness/hole radius), K (stress concentration factor) for bending approaches that for in-plane loading,'' Bert [4]. The authors claim that Eq. (26*b*) describes the stress concentration factor for a very thick plate with a circular rigid inclusion subjected to an in-plane stress field.

Recall from Bert [4], in connection with a large thin plate with a circular hole, ''for generalized plane stress, K is independent of ν, while for generalized bending it depends upon ν (except for the case of $B=1$).'' It is interesting to note that in fact there is a similar case for a thick plate. For a large thick plate with a circular hole subjected to balanced bending stress ($B=1$), Eq. (22*a*) yields $K_H=2$, which is independent of Poisson's ratio ν and geometry ratio a/h. Actually its limiting cases $K_H(\infty)$, Eq. (24*a*), and $K_H(0)$, Eq. (26*a*), also yield a constant K_H with value 2, independent of ν and a/h. Other than for balanced bending, K_H is a function of B, ν and a/h.

However, for a large plate with a circular rigid inclusion, K_R depends on ν, B, and a/h, except for the special case ($B=1$), $K_R=2/(1+\nu)$, which is independent of a/h.

Acknowledgments

This paper is dedicated to the memory of Prof. Eric Reissner, late of the University of California, San Diego (UCSD). The first author had the opportunity to discuss this problem prior to and during his sabbatical leave at UCSD with Professor Reissner.

References

[1] Bickley, W. G., 1924, ''The Effect of a Hole in a Bent Plate,'' Philos. Mag., 6th Ser., **48**, pp. 1014–1024.
[2] Lekhnitskii, S. G., 1936, according to G. Savin, *Stress Concentration Around Holes*, English translation of 1951 Russian edition, Pergamon, New York, 1961, p. 322.
[3] Goodier, J. N., 1936, ''The Influence of Circular and Elliptical Holes on the Transverse Flexure of Elastic Plates,'' Philos. Mag., 7th Ser., **22**, pp. 69–80.
[4] Bert, C. W., 1988, ''Generalized Bending of a Large Plate Containing a Circular Hole,'' Mech. Res. Commun., **15**, No. 1, pp. 55–60.
[5] Goland, M., 1943, ''The Influence of the Shape and Rigidity of an Elastic Inclusion on the Transverse Flexure of Thin Plates,'' ASME J. Appl. Mech., **10**, pp. A69–A75.
[6] Reissner, E., 1945, ''The Effect of Transverse Shear Deformation on the Bending of Elastic Plates,'' ASME J. Appl. Mech., **12**, pp. A69–A77.
[7] Reissner, E., 1983, ''Stress Couple Concentrations for Cylindrically Bent Plates With Circular Holes or Rigid Inclusions,'' ASME J. Appl. Mech., **50**, pp. 85–87.
[8] Cheng, S., 1979, ''Elasticity Theory of Plates and a Refined Theory,'' ASME J. Appl. Mech., **46**, pp. 644–650.
[9] Chen, P. S., and Archer, R. R., 1989, ''Stress Concentration Factors due to the Bending of a Thick Plate with Circular Hole,'' Ing.-Arch., **59**, pp. 401–411.
[10] Hirsch, R. A., 1952, ''Effect of Rigid Circular Inclusion on Bending of Thick Elastic Plate,'' ASME J. Appl. Mech., **19**, pp. 28–32.
[11] Alblas, J. B., 1957, ''Theorie van de Driedimensionale Spanningstoestad in een Doorboorde Plaat,'' dissertation, Delft. See Chen and Archer ([9]).
[12] Dumont, C., 1939, ''Stress Concentration Around an Open Circular Hole in a Plate Subjected to Bending Normal to the Plane of the Plate,'' National Advisory Committee for Aeronautics, Technical Note No. 740.
[13] Reissner, E., 1980, ''On the Analysis of First and Second-Order Shear Deformation Effects for Isotropic Elastic Plates,'' ASME J. Appl. Mech., **47**, pp. 959–961.
[14] Abramovich, M., and Stegun, I. A., 1964, *Handbook of Mathematical Functions*, Applied Mathematics Series, No. 55, National Bureau of Standards, U.S. Department of Commerce.

Three-Dimensional Solutions of Smart Functionally Graded Plates

J. N. Reddy
Fellow ASME
e-mail: jnreddy@shakti.tamu.edu

Z.-Q. Cheng
Mem. ASME

Department of Mechanical Engineering,
Texas A&M University,
College Station, TX 77843-3123

A smart functionally graded plate consists of a plate made of a functionally gradient material and actuators made of an active material. The active material, a layer or set of patches, is bonded on the metal-rich surface of the functionally graded plate. When the ceramic-rich surface of the substrate is subjected to thermomechanical loadings, displacements, and stresses may be controlled, and vibration amplitudes may be suppressed by the actuators with supplied electric power. In the attempt towards a basic understanding of the new type of smart structural system, this study considers a benchmark problem, namely, the bending of a functionally graded rectangular plate with an attached piezoelectric actuator. The transfer matrix and asymptotic expansion techniques are employed to obtain a three-dimensional asymptotic solution. In numerical computations, the locally effective material properties of the functionally gradient material are estimated by the Mori-Tanaka scheme. The three-dimensional distributions of displacements and stresses for different volume fractions of the ceramic and metallic constituents could serve as benchmark results to assess approximate theories and numerical methods.
[DOI: 10.1115/1.1347994]

1 Introduction

Functionally gradient materials (FGMs) are spatial composites within which material properties vary continuously and inhomogeneously. This is achieved by gradually changing the composition of the constituent materials along one direction, usually the thickness direction from one surface to another, to obtain smooth variation of material properties and optimum response to externally applied thermomechanical loads. Typically, functionally graded materials are made from mixture of two or more materials that are appropriate to achieve the desired objective. For example, thermal barrier structures are made of ceramic and metal to withstand high-temperature gradient environments while maintaining the structural strength and fracture toughness. The ceramic constituent of the material provides the high temperature resistance due to its low thermal conductivity. The ductile metal constituent, on the other hand, is placed where greater toughness is needed. There have been several studies of microstructural (see [1,2]) and macrostructural (see [3–13] and references therein) aspects of FGMs. In addition, numerous symposia that are organized in recent years attest to the increased interest in the topic.

Structures with surface-mounted or embedded sensors and actuators are referred to as smart structures. The type of structural system is capable of adapting or taking corrective action to changing operating conditions. The passive structure in a smart system is the load bearing part, whereas the active material part is to perform the operations of sensing and actuation. The actuators work to induce a counteractive static deformation or vibration suppression of the passive structure (e.g., [14–18]).

By integrating active materials onto the structures made of FGMs, a smart functionally graded (FG) structural system is naturally manifested. It makes possible that the FG structures can be actuated by properly applied electric voltage to the actuators to achieve desired shapes and suppress the amplitudes of vibration. With the advent of the smart FG structural system, the operating environments pose serious problem to the design and maintenance. The FG structures are subjected to thermomechanical loads and the integrated actuators are subjected to appropriate electric voltages, featuring functionality and thermoelectromechanical coupling.

There are two approaches to the bending solution of the benchmark problems of a laminated rectangular plate. Pagano [19,20] has developed exact solutions of simply supported laminated elastic plates by using the three-dimensional elasticity theory. Pagano's method simply treats the laminate layer by layer and then enforces the interface continuity conditions. An alternative approach is the transfer matrix technique, in which the interfacial continuity conditions are utilized. Both approaches have been extended widely to the analyses of composite laminated plates and shells for different materials and loads ([21–27], among others). However, the two approaches are only valid for laminated plates and shells, where the material properties are piecewise constant.

Pagano's method and the transfer matrix method are not valid for finding solutions of plate and shell problems with continuous inhomogeneity. A structure made of a FGM is a typical case. Asymptotic expansion is, instead, an efficient method for this kind of problems. This method has seen its applications in single-layer piezoelectric plates ([28,29]) for a leading-order solution, laminated elastic, and piezoelectric plates ([30–39]) for higher-order solutions.

In this paper, the transfer matrix formulation is presented in combination with the asymptotic expansion and is used for obtaining an asymptotic solution to a desired degree of numerical accuracy. The Mori-Tanaka scheme is used to give locally effective material properties of the FGM ([2,36]). Numerical results are presented for smart FG rectangular plates, and the results may serve as a reference for checking the validity of approximate theories and numerical methods.

2 Transfer Matrix Formulation

Figure 1 shows a FG plate attached on its bottom surface by a piezoelectric actuator. A Cartesian coordinate system $Ox_1x_2x_3$ is used and the reference plane $x_3=0$ is located at the bottom surface of the actuator. The top surfaces of the undeformed FG plate and the actuator lie at $x_3=h$ and h_E. Two thin-film conducting electrodes are placed, respectively, on the upper and lower surfaces of the actuator to carry an alternating forcing electric potential. For simplicity, negligible thickness of an electrode is assumed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 6, 1999; final revision, May 2, 2000. Associate Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Editor, Prof. Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

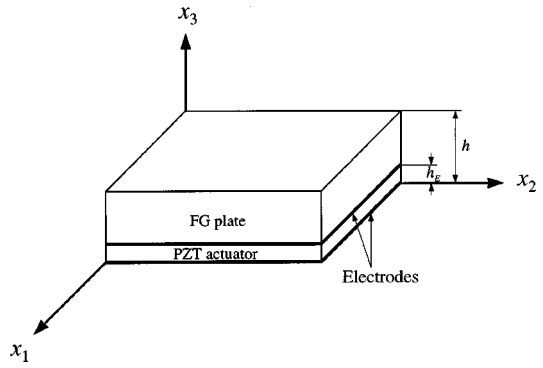


Fig. 1 Geometry of a smart FG plate

Let τ and **S** denote the symmetric stress and strain tensors, **D** and **E** the electric displacement and the electric field vectors, **u** the mechanical displacement vector, φ the electric potential, and T the increment in temperature from a stress-free reference configuration. The governing equations of linear piezothermoelasticity, in the absence of body forces and electric charge density, are ([40])

$$\nabla^T\cdot\tau=\mathbf{0},\quad \nabla\cdot\mathbf{D}=0; \tag{1}$$

$$\mathbf{S}=\frac{1}{2}[\nabla\mathbf{u}+(\nabla\mathbf{u})^T],\quad \mathbf{E}=-\nabla\varphi, \tag{2}$$

where ∇ denotes the gradient operator in three dimensions. For a piezoceramic material of class 6mm with poling in the x_3-direction, the constitutive relations can be written in the matrix form ([41]) as

$$\begin{bmatrix}\tau_{11}\\ \tau_{22}\\ \tau_{33}\\ \tau_{23}\\ \tau_{31}\\ \tau_{12}\\ D_1\\ D_2\\ D_3\end{bmatrix}=\begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 & 0 & 0 & e_{31}\\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 & 0 & 0 & e_{31}\\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 & 0 & 0 & e_{33}\\ 0 & 0 & 0 & c_{44} & 0 & 0 & 0 & e_{15} & 0\\ 0 & 0 & 0 & 0 & c_{44} & 0 & e_{15} & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & c_{66} & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & e_{15} & 0 & -\epsilon_{11} & 0 & 0\\ 0 & 0 & 0 & e_{15} & 0 & 0 & 0 & -\epsilon_{11} & 0\\ e_{31} & e_{31} & e_{33} & 0 & 0 & 0 & 0 & 0 & -\epsilon_{33}\end{bmatrix}\times\begin{bmatrix}S_{11}\\ S_{22}\\ S_{33}\\ 2S_{23}\\ 2S_{31}\\ 2S_{12}\\ -E_1\\ -E_2\\ -E_3\end{bmatrix}+\begin{bmatrix}-\lambda_{11}\\ -\lambda_{22}\\ -\lambda_{33}\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\end{bmatrix}T, \tag{3}$$

with $c_{66}=(c_{11}-c_{12})/2$. Deformation of the actuator made of lead zirconate titanate (PZT) is governed by the three sets of Eqs. (1)–(3). They can also be used for the FG plate made of an inhomogeneous isotropic material by setting

$$c_{11}=c_{33},\quad c_{12}=c_{13},\quad c_{44}=c_{66},\quad \epsilon_{11}=\epsilon_{33},$$
$$e_{31}=e_{33}=e_{15}=0,\quad p_3=0. \tag{4}$$

The nonzero elastic and dielectric moduli in (4) for a FGM are functions of x_3. The dielectricity equation for the FGM, which is decoupled from the elastostatics, is not of interest in this paper. Thus the dielectric moduli of the FGM are trivial to the deformation of the actuated plate. This observation will be utilized later.

Equations (1) through (3) may be reformulated as the following state-space equation:

$$\partial_3\begin{bmatrix}\mathbf{F}\\ \mathbf{G}\end{bmatrix}=\begin{bmatrix}\mathbf{0} & \mathbf{A}\\ \mathbf{B} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{F}\\ \mathbf{G}\end{bmatrix}+\begin{bmatrix}\mathbf{0}\\ \mathbf{C}\end{bmatrix}T, \tag{5}$$

where $\partial_i\equiv\partial/\partial x_i\ (i=1,2,3)$, and

$$\mathbf{F}=[u_1\ \ u_2\ \ \tau_{33}\ \ D_3]^T,\quad \mathbf{G}=[\tau_{13}\ \ \tau_{23}\ \ u_3\ \ \varphi]^T. \tag{6}$$

The 4×4 operator matrices **A** and **B** and 4×1 operator matrix **C** contain the in-plane differential operators ∂_1 and ∂_2, and depend on x_3 only through the material moduli:

$$\mathbf{A}=\begin{bmatrix}1/c_{44} & 0 & -\partial_1 & -k_1\partial_1\\ 0 & 1/c_{44} & -\partial_2 & -k_1\partial_2\\ -\partial_1 & -\partial_2 & 0 & 0\\ -k_1\partial_1 & -k_1\partial_2 & 0 & k_2\Delta\end{bmatrix},$$

$$\mathbf{B}=\begin{bmatrix}-k_3\partial_1^2-c_{66}\partial_2^2 & (c_{66}-k_3)\partial_1\partial_2 & -k_4\partial_1 & -k_5\partial_1\\ (c_{66}-k_3)\partial_1\partial_2 & -c_{66}\partial_1^2-k_3\partial_2^2 & -k_4\partial_2 & -k_5\partial_2\\ -k_4\partial_1 & -k_4\partial_2 & \epsilon_{33}k_0 & e_{33}k_0\\ -k_5\partial_1 & -k_5\partial_2 & e_{33}k_0 & -c_{33}k_0\end{bmatrix}, \tag{7}$$

$$\mathbf{C}=\begin{bmatrix}k_6\partial_1\\ k_6\partial_2\\ (\epsilon_{33}\lambda_{33}-e_{33}p_3)k_0\\ (e_{33}\lambda_{33}+c_{33}p_3)k_0\end{bmatrix},$$

where $\Delta\equiv\partial_1^2+\partial_2^2$ is the Laplace operator in two dimensions. The in-plane stresses and in-plane electric displacements, which are discontinuous across the internal electrode at $x_3=h_E$, are given by

$$\begin{aligned}\tau_{11}&=k_3\partial_1u_1+(k_3-2c_{66})\partial_2u_2+k_4\tau_{33}+k_5D_3-k_6T,\\ \tau_{22}&=(k_3-2c_{66})\partial_1u_1+k_3\partial_2u_2+k_4\tau_{33}+k_5D_3-k_6T,\\ \tau_{12}&=c_{66}(\partial_2u_1+\partial_1u_2),\\ D_1&=k_1\tau_{13}-k_2\partial_1\varphi,\\ D_2&=k_1\tau_{23}-k_2\partial_2\varphi\end{aligned} \tag{8}$$

where

$$\begin{aligned}k_0&=\frac{1}{c_{33}\epsilon_{33}+e_{33}^2},\quad k_1=\frac{e_{15}}{c_{44}},\quad k_2=\frac{e_{15}^2}{c_{44}}+\varepsilon_{11},\\ k_3&=c_{11}-(c_{13}^2\epsilon_{33}+2c_{13}e_{31}e_{33}-c_{33}e_{31}^2)k_0,\\ k_4&=(c_{13}\epsilon_{33}+e_{31}e_{33})k_0,\quad k_5=(c_{13}e_{33}-c_{33}e_{31})k_0,\\ k_6&=\lambda_{11}-(c_{13}\epsilon_{33}+e_{31}e_{33})\lambda_{33}k_0+(c_{13}e_{33}-c_{33}e_{31})p_3k_0.\end{aligned} \tag{9}$$

3 Asymptotic Scheme

The mechanical loading is specified by the tangential tractions $q_1^{\pm}$, $q_2^{\pm}$ and the normal pressures $q_3^{\pm}$ at the plate surfaces $x_3=h,0$. The electric loading is specified by applied electric potentials V_E and V^- at the electrodes $x_3=h_E,0$. The plate is also subjected to a three-dimensional thermal load T, to be separately solved from a heat conduction problem. The thickness coordinate is scaled as $z=x_3/\chi$ by the small parameter $\chi=h/a$, where a is a typical in-plane dimension and z varies from 0 to a as x_3 goes

from 0 to h. The interface between the FG plate and the actuator is at $z=a_E\equiv h_E/\chi$. The mechanical and thermal loads are then scaled as

$$\tau_{13}|_{z=0}=\chi^2 q_1^-, \quad \tau_{23}|_{z=0}=\chi^2 q_2^-, \quad \tau_{33}|_{z=0}=-\chi^3 q_3^-,$$
$$\varphi|_{z=0}=\chi^2 V^-, \tag{10}$$

$$\tau_{13}|_{z=a}=\chi^2 q_1^+, \quad \tau_{23}|_{z=a}=\chi^2 q_2^+, \quad \tau_{33}|_{z=a}=-\chi^3 q_3^+,$$
$$\varphi|_{z=a_E}=\chi^2 V_E. \tag{11}$$

The state space functions **F** and **G** are expanded in terms of the small thickness parameter χ as

$$\begin{bmatrix}\mathbf{F}\\ \mathbf{G}\end{bmatrix}=\sum_{n=0}^{\infty}\chi^{2n}\begin{bmatrix}\chi\mathbf{f}^{(n)}\\ \mathbf{g}^{(n)}\end{bmatrix}. \tag{12}$$

Denoting the integral operators

$$Q(\cdots)\equiv\int_0^z(\cdots)dz, \quad \overline{Q}(\cdots)\equiv\int_0^a(\cdots)dz,$$
$$Q_E(\cdots)\equiv\int_0^{a_E}(\cdots)dz, \tag{13}$$

and substituting the expansion (12) into Eq. (5), with the help of Eq. (10), leads to

$$\mathbf{g}^{(0)}=\begin{bmatrix}0\\0\\U_3^{(0)}\\0\end{bmatrix}, \quad \mathbf{g}^{(n+1)}=\begin{bmatrix}q_1^-\delta_{n0}\\ q_2^-\delta_{n0}\\ U_3^{(n+1)}\\ V^-\delta_{n0}\end{bmatrix}+Q\mathbf{B}\mathbf{f}^{(n)}+\delta_{n0}Q\mathbf{C}\theta,$$
$$\mathbf{f}^{(n)}=\mathbf{X}^{(n)}+\mathbf{H}^{(n)}, \quad (n\geqslant 0), \tag{14}$$

where

$$\mathbf{X}^{(n)}=\begin{bmatrix}U_1^{(n)}-z\partial_1 U_3^{(n)}\\ U_2^{(n)}-z\partial_2 U_3^{(n)}\\ 0\\ U_4^{(n)}+\langle D_3^{(n)}\rangle H(z-a_E)\end{bmatrix}, \tag{15}$$

$$\mathbf{H}^{(n+1)}=\delta_{n0}\left\{Q\mathbf{A}\begin{bmatrix}q_1^-\\ q_2^-\\ 0\\ V^-\end{bmatrix}-\begin{bmatrix}0\\0\\q_3^-\\0\end{bmatrix}+Q\mathbf{A}Q\mathbf{C}\theta\right\}$$
$$+Q\mathbf{A}Q\mathbf{B}(\mathbf{X}^{(n)}+\mathbf{H}^{(n)}),$$

with $\mathbf{H}^{(0)}=\mathbf{0}$, $\theta=T/\chi$ and $\langle D_3^{(n)}\rangle\equiv D_3^{(n)}|_{z=a_E^+}-D_3^{(n)}|_{z=a_E^-}$. Because of the internal conducting electrode at $z=a_E$, the transverse electric displacement must be different on the two sides of the electrode. In this case, the discontinuity in the transverse electric displacement must be accommodated. In addition to the unknown $\langle D_3^{(n)}\rangle$, the basic unknowns are the components of the three mechanical displacements and the electric displacement at the bottom surface $z=0$ of the actuator:

$$U_1^{(n)}=u_1^{(n)}|_{z=0}, \quad U_2^{(n)}=u_2^{(n)}|_{z=0}, \quad U_3^{(n)}=u_3^{(n)}|_{z=0},$$
$$U_4^{(n)}\equiv D_3^{(n)}|_{z=0^+}. \tag{16}$$

These unknowns are to be determined such that the conditions (11) are satisfied through Eq. (14). After some manipulation, this gives

$$\mathbf{R}\mathbf{f}^{(n)}=\delta_{n0}\mathbf{Y}, \tag{17}$$

where

$$\mathbf{R}=\begin{bmatrix}-\overline{Q}k_3\partial_1^2-\overline{Q}c_{66}\partial_2^2 & \overline{Q}(c_{66}-k_3)\partial_1\partial_2 & -\overline{Q}k_4\partial_1 & -\overline{Q}k_5\partial_1\\ \overline{Q}(c_{66}-k_3)\partial_1\partial_2 & -\overline{Q}c_{66}\partial_1^2-\overline{Q}k_3\partial_2^2 & -\overline{Q}k_4\partial_2 & -\overline{Q}k_5\partial_2\\ -\overline{Q}zk_3\partial_1\Delta & -\overline{Q}zk_3\partial_2\Delta & -\overline{Q}zk_4\Delta & -\overline{Q}zk_5\Delta\\ -Q_Ek_5\partial_1 & -Q_Ek_5\partial_2 & Q_Ee_{33}k_0 & -Q_Ec_{33}k_0\end{bmatrix}, \tag{18}$$

$$\mathbf{Y}=\begin{bmatrix}q_1^+-q_1^--\overline{Q}k_6\partial_1\theta\\ q_2^+-q_2^--\overline{Q}k_6\partial_2\theta\\ -q_3^++q_3^-+a(\partial_1 q_1^++\partial_2 q_2^+)-\overline{Q}zk_6\Delta\theta\\ V_E-V^--Q_E(e_{33}\lambda_{33}+c_{33}p_3)k_0\theta\end{bmatrix}. \tag{19}$$

The set of Eqs. (17) of each order have four equations for five unknowns $U_i^{(n)}(i=1,2,3,4)$ and $\langle D_3^{(n)}\rangle$. The jump in the transverse electric displacement across the internal electrode at $z=a_E$ can be evaluated from a rigid dielectricity equation for the FGM, which is decoupled from the thermoelastic deformation problem under consideration. Because it is a problem of no interest in practice, an important procedure is proposed to simplify Eq. (17).

Physically, it is clear that the dielectric properties and the thermoelastic properties of the FGM are irrelevant to each other. Thus the dielectric moduli of the FGM have trivial contributions to the thermoelastic deformation of the FG plate. As mentioned earlier, this property may be utilized to advantage. By setting the transverse dielectric modulus $\epsilon_{33}^{\mathrm{FGM}}$ of the FGM to be infinite, it will only change the dielectricity solution for the material. The thermoelasticity solution for the FG plate will not change and nor will the piezothermoelasticity solution for the actuator. It is clear from $(7)_2$ and $(9)_1$, when $a_E<z\leqslant a$, that $B_{44}=-c_{33}k_0=0$ as $\epsilon_{33}^{\mathrm{FGM}}\to\infty$. In addition, $B_{14}=B_{24}=B_{34}=0$ when $a_E<z\leqslant a$. It follows that the term $\langle D_3^{(n)}\rangle H(z-a_E)$ involved in $\mathbf{X}^{(n)}$ has a trivial contribution to $Q\mathbf{B}\mathbf{X}^{(n)}$, and hence to the recurrence relation $(15)_2$. The same is that $\langle D_3^{(n)}\rangle H(z-a_E)$ has a trivial contribution to Eq. (17) as a result of $R_{14}=-Q_Ek_5\partial_1$, $R_{24}=-Q_Ek_5\partial_2$, $R_{34}=-Q_Ezk_5\Delta$ and $R_{44}=-Q_Ec_{33}k_0$. Consequently, $\langle D_3^{(n)}\rangle H(z-a_E)$ may be dropped in Eq. $(15)_1$ when taking $\epsilon_{33}^{\mathrm{FGM}}\to\infty$. This reduces Eq. (17) to

$$\tilde{\mathbf{R}}\tilde{\mathbf{X}}^{(n)}=\delta_{n0}\mathbf{Y}-\mathbf{R}\mathbf{H}^{(n)}, \tag{20}$$

where

$$\tilde{\mathbf{X}}^{(n)}=[U_1^{(n)} \;\; U_2^{(n)} \;\; -U_3^{(n)} \;\; U_4^{(n)}]^T, \tag{21}$$

$$\tilde{\mathbf{R}} = -\begin{bmatrix} \overline{Q}k_3\partial_1^2 + \overline{Q}c_{66}\partial_2^2 & \overline{Q}(k_3 - c_{66})\partial_1\partial_2 & \overline{Q}zk_3\partial_1\Delta & Q_E k_5\partial_1 \\ \overline{Q}(k_3 - c_{66})\partial_1\partial_2 & \overline{Q}c_{66}\partial_1^2 + \overline{Q}k_3\partial_2^2 & \overline{Q}zk_3\partial_2\Delta & Q_E k_5\partial_2 \\ \overline{Q}zk_3\partial_1\Delta & \overline{Q}zk_3\partial_2\Delta & \overline{Q}z^2k_3\Delta^2 & Q_E zk_5\Delta \\ Q_E k_5\partial_1 & Q_E k_5\partial_2 & Q_E zk_5\Delta & Q_E c_{33}k_0 \end{bmatrix}. \tag{22}$$

Note that the symmetric matrix operator $\tilde{\mathbf{R}}$ is a generalization of that for the bending of inhomogeneous elastic plates derived under the Kirchhoff thin plate assumptions ([42]). Details for the numerical asymptotic procedure may be found in [37].

4 Numerical Results

The locally effective material properties of an FGM can be predicted by micromechanical models such as the Mori-Tanaka estimates, the self-consistent, generalized self-consistent or differential schemes, or the cell method proposed by Aboudi [43]. It should be noted that these average field schemes, except the cell method, are originally developed for statistically homogeneous aggregates based on a representative volume element (RVE). In general, however, such a representative volume element may not be defined in a FGM. Assessment of these existing micromechanical schemes has been given in, for example, [44–47].

Consider a two-phase composite plate consisting of a matrix phase denoted by subscript 1 and a particulate phase denoted by the subscript 2. The composite is reinforced by spherical particles, randomly distributed in the plane of the plate. The locally effective bulk modulus K and shear modulus μ of the FGM are given by the Mori-Tanaka estimates ([48,49]) as

$$\frac{K-K_1}{K_2-K_1} = \frac{V_2}{1+(1-V_2)\dfrac{K_2-K_1}{K_1+\frac{4}{3}\mu_1}},$$

$$\frac{\mu-\mu_1}{\mu_2-\mu_1} = \frac{V_2}{1+(1-V_2)\dfrac{\mu_2-\mu_1}{\mu_1+\pi_1}}, \quad \pi_1 = \frac{\mu_1(9K_1+8\mu_1)}{6(K_1+2\mu_1)}. \tag{23}$$

The locally effective heat conductivity coefficient κ is given by ([50])

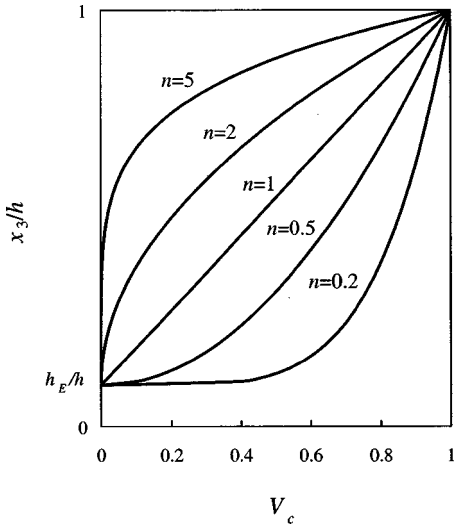


Fig. 2 Through-the-thickness distribution of the ceramic phase in the FG plate

$$\frac{\kappa-\kappa_1}{\kappa_2-\kappa_1} = \frac{V_2}{1+(1-V_2)\dfrac{\kappa_2-\kappa_1}{3\kappa_1}}. \tag{24}$$

The coefficient of thermal expansion α is determined in terms of the exact correspondence relation ([51])

$$\frac{\alpha-\alpha_1}{\alpha_2-\alpha_1} = \frac{\dfrac{1}{K}-\dfrac{1}{K_1}}{\dfrac{1}{K_2}-\dfrac{1}{K_1}}. \tag{25}$$

Here V_2 denotes the volume fraction of the particulate phase. The Mori-Tanaka estimates on statistically homogeneous composites with spherical reinforcements coincide with the Hashin-Shtrikman upper and lower bounds on elastic moduli ([52]), when the stiffer phase serves as a matrix or reinforcement of well-ordered composites.

It is assumed that the volume fraction of the ceramic phase is of the power-law type $V_c = [(x_3 - h_E)/(h - h_E)]^n$. Figure 2 shows the through-thickness variation for $n = 0.2, 0.5, 1, 2, 5$. Note that the bottom surface of the FG plate is metal-rich and the top surface is ceramic-rich. In actual service conditions, top zirconia surface provides a thermal barrier on Ni-based structural components in aircraft engines. The constituent materials of the FG plate are taken to be nickel-based alloy, Monel (70Ni-30Cu), and zirconia with their material properties ([3,53,54])

$$\begin{aligned} &K_m = 227.24\times10^9\ \text{N/m}^2, && K_c = 125.83\times10^9\ \text{N/m}^2, \\ &\mu_m = 65.55\times10^9\ \text{N/m}^2, && \mu_c = 58.077\times10^9\ \text{N/m}^2, \\ &\alpha_m = 15\times10^{-6}/\text{K}, && \alpha_c = 10\times10^{-6}/\text{K}, \\ &\kappa_m = 25\ \text{W/mK}, && \kappa_c = 2.09\ \text{W/mK}, \end{aligned} \tag{26}$$

where the subscripts m and c stand for the metal and ceramic. When Monel and zirconia serve as a matrix phase, respectively, two sets of effective material moduli are obtained. It can be proved that if the shear moduli of two constituents are identical, the two sets of estimates on the effective material moduli K, μ, and α will be identical. Because the value of the shear modulus of Monel is close to that of zirconia, the two sets of estimates on K, μ and α are very close to each other. In the following numerical results Monel is chosen to serve as a matrix phase. The material of the actuator is taken as PZT-5A with the following material properties ([25]):

$$\begin{gathered} c_{11} = 99.201\times10^9\ \text{N/m}^2, \quad c_{12} = 54.016\times10^9\ \text{N/m}^2, \\ c_{13} = 50.778\times10^9\ \text{N/m}^2, \\ c_{33} = 86.856\times10^9\ \text{N/m}^2, \quad c_{44} = 21.1\times10^9\ \text{N/m}^2, \\ e_{31} = -7.209\ \text{C/m}^2, \quad e_{33} = 15.118\ \text{C/m}^2, \quad e_{15} = 12.322\ \text{C/m}^2, \\ \epsilon_{11} = 1.53\times10^{-8}\ \text{F/m}, \quad \epsilon_{33} = 1.5\times10^{-8}\ \text{F/m}, \\ \lambda_{11} = 0.3314\times10^6\ \text{N/m}^2\text{K}, \quad \lambda_{33} = 0.326\times10^6\ \text{N/m}^2\text{K}, \\ p_3 = 7\times10^{-4}\ \text{C/m}^2\text{K}, \\ \kappa_{11} = 1.8\ \text{W/mK}, \quad \kappa_{33} = 1.8\ \text{W/mK}. \end{gathered} \tag{27}$$

Because the bottom surface of the FG plate is a metal surface, the applied electric potential at the internal electrode is $V_E=0$. A rectangular plate is considered in the following example. The only nonzero loads are specified as

$$[T^+ \quad q_3^+ \quad V^-]=[\hat{T}^+ \quad \hat{q}_3^+ \quad \hat{V}^-] \sin l_1 x_1 \sin l_2 x_2, \tag{28}$$

$$l_1=\frac{m_1\pi}{a}, \quad l_2=\frac{m_2\pi}{b}, \tag{29}$$

where a quantity with a superimposed hat denotes the amplitude of the corresponding physical quantity. The solution procedure for the temperature field is taken similarly to that given in the last section, i.e., an asymptotic scheme for the steady-state heat conduction problem, but details are omitted.

In general, specifying the edge boundary conditions in the sense of the Kirchhoff plate theory only yields the accurate leading-order interior solution. An accurate and consistent description of the boundary conditions for solving higher-order interior solutions should account for the specified edge distribution to achieve a decaying state, i.e., asymptotic to the exact solution away from edges ([55,56]). However, for the special case of the edge boundary conditions:

$$u_2=u_3=\tau_{11}=\varphi=T=0, \quad \text{at } x_1=0,a, \tag{30}$$

$$u_1=u_3=\tau_{22}=\varphi=T=0, \quad \text{at } x_2=0,b,$$

there are no boundary layer effects for the specific problem. A fully three-dimensional solution may be generated to any desired degree of numerical accuracy in terms of

$$\tilde{\mathbf{X}}^{(n)}=\begin{bmatrix} U_1^{(n)} \\ U_2^{(n)} \\ -U_3^{(n)} \\ U_4^{(n)} \end{bmatrix}=\begin{bmatrix} \hat{U}_1^{(n)} \cos l_1 x_1 \sin l_2 x_2 \\ \hat{U}_2^{(n)} \sin l_1 x_1 \cos l_2 x_2 \\ -\hat{U}_3^{(n)} \sin l_1 x_1 \sin l_2 x_2 \\ \hat{U}_4^{(n)} \sin l_1 x_1 \sin l_2 x_2 \end{bmatrix},$$

$$T=\hat{T} \sin l_1 x_1 \sin l_2 x_2. \tag{31}$$

The peak values of the physical quantities are nondimensionalized by

$$\overline{u}_i=\frac{\hat{u}_i}{Pa}, \quad \overline{\tau}_{ij}=\frac{\hat{\tau}_{ij}}{Pc^*}, \quad \overline{T}=\frac{\alpha^*\hat{T}}{P}, \tag{32}$$

where $P=\alpha^*\hat{T}^+$ for applied thermal load T^+, $P=\hat{q}_3^+/c^*$ for applied mechanical load q_3^+, and $P=\hat{V}^-(e^*/ac^*)$ for applied electric load $-V^-$. Results for complex loadings can be obtained by a linear superposition of respective results caused by simple loadings. $a/b=1$, $h_E/h=0.1$, $m_1=m_2=1$, $c^*=10^{10}\,\mathrm{Nm}^{-2}$, e^*

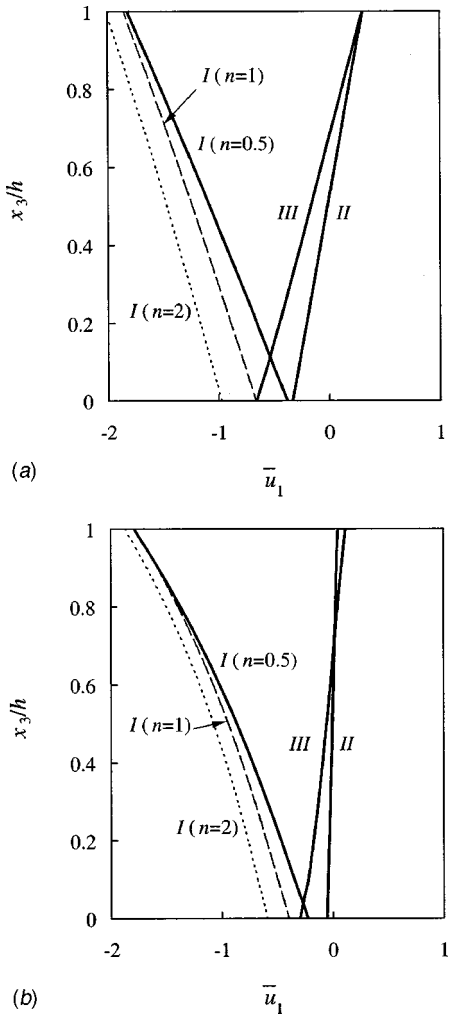


Fig. 3 Through-the-thickness distribution of the dimensionless displacement $\overline{u}_1$ of the plate ($a/b=1$, $n=0.5, 1, 2$) under (I) the thermal load T^+, (II) the mechanical load q_3^+, and (III) the electric load $-V^-$: (*a*) $a/h=10$; (*b*) $a/h=4$

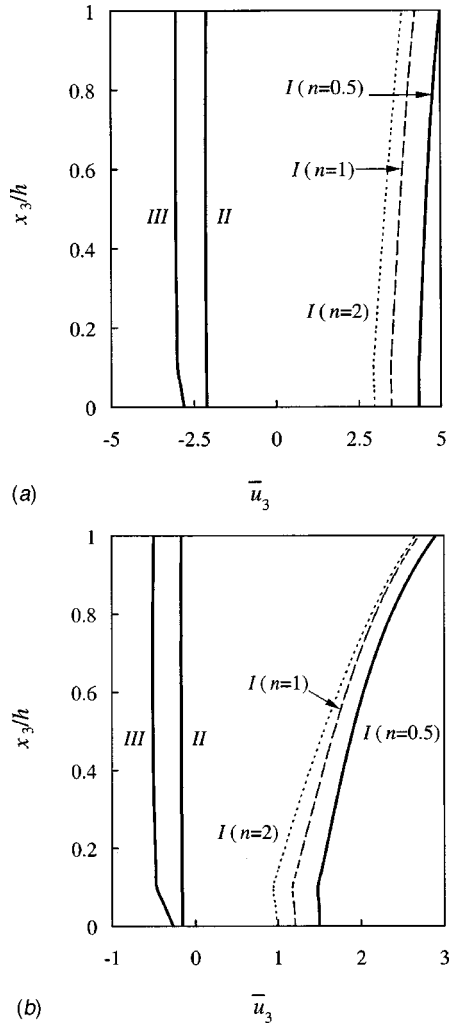


Fig. 4 Through-the-thickness distribution of the dimensionless deflection $\overline{u}_3$ of the plate ($a/b=1$, $n=0.5, 1, 2$) under (I) the thermal load T^+, (II) the mechanical load q_3^+, and (III) the electric load $-V^-$: (*a*) $a/h=10$; (*b*) $a/h=4$

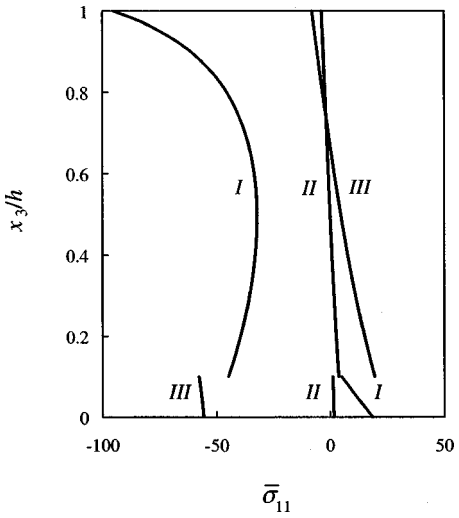


Fig. 5 Through-the-thickness distribution of the dimensionless longitudinal stress $\bar{\sigma}_{11}$ of the plate under (I) the thermal load T^+, (II) the mechanical load q_3^+, and (III) the electric load $-V^-$ ($a/b=1$, $a/h=4$, $n=1$)

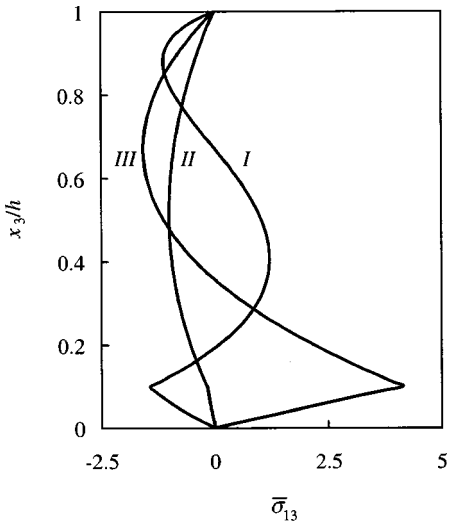


Fig. 7 Through-the-thickness distribution of the dimensionless transverse shear stress $\bar{\sigma}_{13}$ of the plate under (I) the thermal load T^+, (II) the mechanical load q_3^+, and (III) the electric load $-V^-$ ($a/b=1$, $a/h=4$, $n=1$)

$=10\,\mathrm{Cm}^{-2}$, and $\alpha^*=10^{-6}\,\mathrm{K}^{-1}$ are used. The 30th-order solution is given to ensure numerical convergence to five significant digits. Information on the numerical convergence can be found in [37].

The through-thickness distributions of the displacements $\bar{u}_1$ and $\bar{u}_3$ ($a/b=1$, $n=0.5$, 1, 2) under (I) the thermal load T^+, (II) the mechanical load q_3^+, (III) the electric load $-V^-$ are plotted in Figs. 3 and 4, respectively, for (*a*) $a/h=10$ and (*b*) $a/h=4$. Under mechanical and electric loads, differences of the distributions of the displacements $\bar{u}_1$ and $\bar{u}_3$ between the three volume fractions $n=0.5$, 1, 2 are not visibly noticeable. The volume fraction distribution is significant only with respect to the applied temperature field. The in-plane displacement $\bar{u}_1$ is linearly distributed through the plate thickness for the case of a moderate thick plate ($a/h=10$) and is nearly linearly distributed for the case of a thick plate ($a/h=4$). The through-thickness distribution of the transverse displacement $\bar{u}_3$ is constant for the case of the mechanical load, even for a thick plate. The through-thickness distribution of the deflection $\bar{u}_3$ of the plate under the electric load is piecewise linear. The through-thickness distribution of the deflection $\bar{u}_3$ for the case of the thermal load is nonlinear for a thick plate ($a/h=4$). The information provided is useful for constructing a displacement-based approximate theory.

Through-the-thickness distributions of the stresses $\bar{\sigma}_{11}$, $\bar{\sigma}_{12}$, $\bar{\sigma}_{13}$, and $\bar{\sigma}_{33}$ are, respectively, depicted in Figs. 5–8 for $a/h=4$ and $n=1$, when the plate is loaded by (I) the thermal load T^+, (II) the mechanical load q_3^+, and (III) the electric load $-V^-$. The longitudinal stresses $\bar{\sigma}_{11}$ and $\bar{\sigma}_{12}$ are discontinuous across the interface between the FG plate and the actuator. In the case of thermal load, their magnitudes are much bigger than those of the transverse stresses $\bar{\sigma}_{13}$ and $\bar{\sigma}_{33}$. It is seen in Fig. 7 that the interfacial stress $\bar{\sigma}_{13}$ at $z=a_E$ is significant in the case of thermal and electric loads. This adverse effect may possibly generate delamination between the actuator and the substrate when using the active material to actuate the substrate structure. Although results

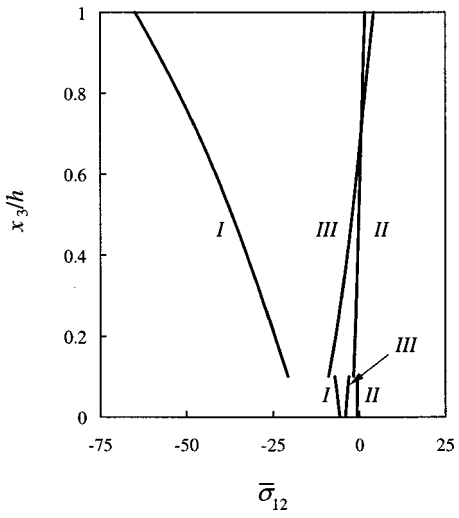


Fig. 6 Through-the-thickness distribution of the dimensionless longitudinal stress $\bar{\sigma}_{12}$ of the plate under (I) the thermal load T^+, (II) the mechanical load q_3^+, and (III) the electric load $-V^-$ ($a/b=1$, $a/h=4$, $n=1$)

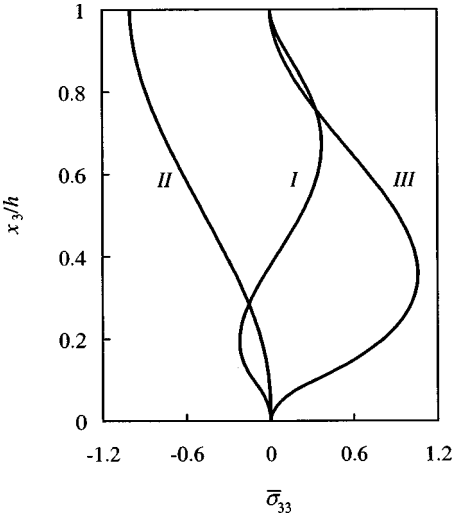


Fig. 8 Through-the-thickness distribution of the dimensionless transverse normal stress $\bar{\sigma}_{33}$ of the plate under (I) the thermal load T^+, (II) the mechanical load q_3^+, and (III) the electric load $-V^-$ ($a/b=1$, $a/h=4$, $n=1$)

Table 1 Results for the smart FG plate under the thermal load ($a/b=1$, $n=2$)

	x_3/h	$a/h=4$	$a/h=10$	$a/h=50$
$\bar{T}$	0.55	0.37792	0.51626	0.55365
	0.1	0.24759	0.37127	0.40524
$\bar{u}_1$	1	-1.8717	-2.0160	-2.0560
	0.55	-1.1348	-1.5130	-1.6141
	0.1	-0.6838	-1.0684	-1.1746
	0	-0.5957	-0.9724	-1.0770
$\bar{u}_3$	1	2.6555	3.8281	15.688
	0.55	1.6408	3.3455	15.587
	0.1	0.9383	2.9365	15.499
	0	0.9854	2.9670	15.505
$\bar{\sigma}_{11}$	1	-88.873	-79.094	-76.381
	0.55	-45.216	-64.787	-70.126
	0.1^+	-44.578	-62.876	-67.852
	0.1^-	8.423	13.893	15.414
	0	25.703	41.019	45.261
$\bar{\sigma}_{12}$	1	-68.304	-73.570	-75.031
	0.55	-45.354	-60.466	-64.507
	0.1^+	-28.163	-44.002	-48.376
	0.1^-	-9.710	-15.171	-16.679
	0	-8.458	-13.808	-15.294
$\bar{\sigma}_{13}$	0.55	1.0015	0.5451	0.11705
	0.1	-2.0493	-1.3173	-0.29106
$\bar{\sigma}_{33}$	0.55	-0.19845	-0.10537	-0.0050519
	0.1	-0.17742	-0.04562	-0.0020160

for $n=0.5$ and 2 are not given herein, significant differences of the stress distributions of $\bar{\sigma}_{11}$, $\bar{\sigma}_{12}$, $\bar{\sigma}_{13}$, and $\bar{\sigma}_{33}$ between the three different volume fractions are only noticed for the case of thermal load. The volume fraction does not change much the distributions of these stresses in the case of mechanical and electric loads.

Table 2 Results for the smart FG plate under the mechanical load ($a/b=1$, $n=2$)

	x_3/h	$a/h=4$	$a/h=10$	$a/h=50$
$\bar{u}_1$	1	0.045615	0.29601	7.4617
	0.55	0.000846	0.01913	0.5374
	0.1	-0.041700	-0.25642	-6.3858
	0	-0.052462	-0.31901	-7.9257
$\bar{u}_3$	1	-0.15724	-2.0287	-245.24
	0.55	-0.16058	-2.0481	-245.35
	0.1	-0.15176	-2.0285	-245.26
	0	-0.14744	-2.0179	-245.20
$\bar{\sigma}_{11}$	1	-3.5199	-20.489	-506.11
	0.55	-0.3672	-1.893	-45.11
	0.1^+	3.7135	22.875	569.84
	0.1^-	1.1703	7.263	181.19
	0	1.5989	9.695	240.74
$\bar{\sigma}_{12}$	1	1.6646	10.802	272.30
	0.55	0.0338	0.765	21.48
	0.1^+	-1.7175	-10.561	-263.01
	0.1^-	-0.5921	-3.641	-90.68
	0	-0.7450	-4.530	-112.54
$\bar{\sigma}_{13}$	0.55	-1.0038	-2.5283	-12.659
	0.1	-0.1609	-0.3946	-1.964
$\bar{\sigma}_{33}$	0.55	-0.54592	-0.54688	-0.54682
	0.1	-0.01323	-0.01294	-0.01288

Table 3 Results for the smart FG plate under the electric load ($a/b=1$, $n=2$)

	x_3/h	$a/h=4$	$a/h=10$	$a/h=50$
$\bar{u}_1$	1	0.11356	0.29614	1.4916
	0.55	-0.03535	-0.10602	-0.5467
	0.1	-0.21339	-0.51997	-2.5875
	0	-0.28772	-0.62597	-3.0438
$\bar{u}_3$	1	-0.47162	-2.8980	-72.147
	0.55	-0.48511	-2.9103	-72.159
	0.1	-0.44064	-2.8604	-72.108
	0	-0.24373	-2.6653	-71.914
$\bar{\sigma}_{11}$	1	-7.696	-20.07	-101.09
	0.55	3.400	9.03	45.64
	0.1	19.231	46.48	230.92
	0.1^-	-57.815	-145.46	-728.14
	0	-55.757	-141.69	-710.57
$\bar{\sigma}_{12}$	1	4.1440	10.807	54.43
	0.55	-1.4130	-4.237	-21.85
	0.1^+	-8.7887	-21.416	-106.57
	0.1^-	-3.0301	-7.384	-36.74
	0	-4.0856	-8.889	-43.22
$\bar{\sigma}_{13}$	0.55	-1.3747	-1.4284	-1.4384
	0.1	4.1714	4.2534	4.2686
$\bar{\sigma}_{33}$	0.55	0.83128	0.34944	0.070514
	0.1	0.32442	0.13276	0.026662

Tables 1, 2, and 3 provide useful results at some particular points of the plate for $a/h=4$ (thick), 10 (moderately thick), or 50 (thin), respectively, under thermal, mechanical, and electric loads. For a thin plate $a/h=50$, a nearly constant through-thickness distribution of the deflection is found in every load case. When a plate is moderately thick or thick, change in the through-thickness deflection becomes significant, especially for a thick plate ($a/h=4$) under the thermal and electric loads.

Deflection control is illustrated by taking $a/b=1$, $a/h=10$, $h_E/h=0.1$, $n=2$. According to Tables 1–3, for the applied temperature $T^+=300$ K and vanishing mechanical load $q_3^+=0$, the required electric voltage through the thickness of the actuator to reduce to $(1-\beta)$ percent of the central deflection of the FG plate ($x_3=0.55h$) is calculated as $\hat{V}^-/h_E=-34.486\beta$ (MV/m). For example, $\beta=0.4$ means that the actuator reduces 40 percent of the deflection at $x_3=0.55h$ of the smart FG plate.

5 Conclusions

Active control of deflections of a functionally graded plate to which piezoelectric actuators are bonded on the metal-rich surface of the FG plate has been developed. The benchmark problem of a rectangular FG plate with an attached piezoelectric actuator has been studied. An asymptotic scheme has been used to generate a three-dimensional solution. The solution is exact in the sense that any desired numerical accuracy may be achieved. Numerical results are presented that may serve as a reference for developing approximate theories and checking numerical solutions. It is observed that the volume fraction distribution is significant only with respect to the applied temperature field.

Acknowledgment

The support of this research by the Army Research Office through Grant DAAH 04-96-1-0080 and the Oscar S. Wyatt Chair is gratefully acknowledged.

References

[1] Aboudi, J., Pindera, M. J., and Arnold, S. M., 1997, ''Microstructural Optimization of Functionally Graded Composites Subjected to a Thermal Gradient via the Coupled Higher-Order Theory,'' Composites, Part B, **28B**, pp. 93–108.

[2] Nadeau, J. C., and Ferrari, M., 1999, ''Microstructural Optimization of a Functionally Graded Transversely Isotropic Layer,'' Mech. Mater., **31**, pp. 637–651.

[3] Praveen, G. N., and Reddy, J. N., 1998, ''Nonlinear Transient Thermoelastic Analysis of Functionally Graded Ceramic-Metal Plates,'' Int. J. Solids Struct., **35**, pp. 4457–4476.

[4] Praveen, G. N., Chin, C. D., and Reddy, J. N., 1999, ''Thermoelastic Analysis of Functionally Graded Ceramic-Metal Cylinder,'' J. Eng. Mech., **125**, pp. 1259–1267.

[5] Reddy, J. N., and Chin, C. D., 1998, ''Thermomechanical Analysis of Functionally Graded Cylinders and Plates,'' J. Therm. Stresses, **21**, pp. 593–626.

[6] Reddy, J. N., Wang, C. M., and Kitipornchai, S., 1999, ''Axisymmetric Bending of Functionally Graded Circular and Annular Plates,'' Eur. J. Mech. A/Solids, **18**, pp. 185–199.

[7] Reddy, J. N., 2000, ''Analysis of Functionally Graded Plates,'' Int. J. Numer. Methods Eng., **47**, pp. 663–684.

[8] Loy, C. T., Lam, K. Y., and Reddy, J. N., 1999, ''Vibration of Functionally Graded Cylindrical Shells,'' Int. J. Mech. Sci., **41**, pp. 309–324.

[9] Gong, S. W., Lam, K. Y., and Reddy, J. N., 1999, ''The Elastic Response of Functionally Graded Cylindrical Shells to Low Velocity Impact,'' Int. J. Impact Eng., **22**, pp. 397–417.

[10] Cheng, Z. Q., and Kitipornchai, S., 1999, ''Membrane Analogy of Buckling and Vibration of Inhomogeneous Plates,'' J. Eng. Mech., **125**, pp. 1293–1297.

[11] Cheng, Z. Q., and Kitipornchai, S., 2000, ''Exact Bending Solution of Inhomogeneous Plates from Homogeneous Thin-Plate Deflection,'' AIAA J., **38**, pp. 1289–1291.

[12] Cheng, Z. Q., and Batra, R. C., 2000, ''Exact Correspondence between Eigenvalues of Membranes and Functionally Graded Simply Supported Polygonal Plates,'' J. Sound Vib., **229**, pp. 879–895.

[13] Cheng, Z. Q., and Batra, R. C., 2000, ''Deflection Relationships between the Homogeneous Kirchhoff Plate Theory and Different Functionally Graded Plate Theories,'' Arch. Mech., **52**, pp. 143–158.

[14] Crawley, E. F., and de Luis, J., 1987, ''Use of Piezoelectric Actuators as Elements of Intelligent Structures,'' AIAA J., **25**, pp. 1373–1385.

[15] Zhou, Y. S., and Tiersten, H. F., 1994, ''Elastic Analysis of Laminated Composite Plates in Cylindrical Bending due to Piezoelectric Actuators,'' Smart Mater. Struct., **3**, pp. 255–265.

[16] Mitchell, J. A., and Reddy, J. N., 1995, ''A Refined Hybrid Plate Theory for Composite Laminates With Piezoelectric Laminae,'' Int. J. Solids Struct., **32**, pp. 2345–2367.

[17] Robbins, D. H., and Reddy, J. N., 1996, ''An Efficient Computational Model for the Stress Analysis of Smart Plate Structures,'' Smart Mater. Struct., **5**, pp. 353–360.

[18] Reddy, J. N., 1999, ''On Laminated Composite Plates with Integrated Sensors and Actuators,'' Eng. Struct., **21**, pp. 568–593.

[19] Pagano, N. J., 1969, ''Exact Solutions for Composite Laminates in Cylindrical Bending,'' J. Compos. Mater., **3**, pp. 398–411.

[20] Pagano, N. J., 1970, ''Exact Solutions for Rectangular Bi-directional Composites and Sandwich Plates,'' J. Compos. Mater., **4**, pp. 20–34.

[21] Heyliger, P., 1994, ''Static Behavior of Laminated Elastic/Piezoelectric Plates,'' AIAA J., **32**, pp. 2481–2484.

[22] Bisegna, P., and Maceri, F., 1996, ''An Exact Three-Dimensional Solution for Simply Supported Rectangular Piezoelectric Plates,'' ASME J. Appl. Mech., **63**, pp. 628–638.

[23] Kapuria, S., Dube, G. P., and Dumir, P. C., 1997, ''Exact Piezothermoelastic Solution for Simply Supported Laminated Flat Panel in Cylindrical Bending,'' Z. Angew. Math. Mech., **77**, pp. 281–293.

[24] Sosa, H. A., 1992, ''On the Modelling of Piezoelectric Laminated Structures,'' Mech. Res. Commun., **19**, pp. 541–546.

[25] Xu, K., Noor, A. K., and Tang, Y. Y., 1995, ''Three-Dimensional Solutions for Coupled Thermoelectroelastic Response of Multilayered Plates,'' Comput. Methods Appl. Mech. Eng., **126**, pp. 355–371.

[26] Xu, K., Noor, A. K., and Tang, Y. Y., 1997, ''Three-Dimensional Solutions for Free Vibrations of Initially-Stressed Thermoelectroelastic Multilayered Plates,'' Comput. Methods Appl. Mech. Eng., **141**, pp. 125–139.

[27] Lee, J. S., and Jiang, L. Z., 1996, ''Exact Electroelastic Analysis of Piezoelectric Laminae via State Space Approach,'' Int. J. Solids Struct., **33**, pp. 977–990.

[28] Maugin, G. A., and Attou, D., 1990, ''An Asymptotic Theory of Thin Piezoelectric Plates,'' Q. J. Mech. Appl. Math., **43**, pp. 347–362.

[29] Bisegna, P., and Maceri, F., 1996, ''A Consistent Theory of Thin Piezoelectric Plates,'' J. Intell. Mater. Syst. Struct., **7**, pp. 372–389.

[30] Rogers, T. G., Watson, P., and Spencer, A. J. M., 1992, ''An Exact Three-Dimensional Solution for Normal Loading of Inhomogeneous and Laminated Anisotropic Elastic Plates of Moderate Thickness,'' Proc. R. Soc. London, Ser. A, **437**, pp. 199–213.

[31] Rogers, T. G., Watson, P., and Spencer, A. J. M., 1995, ''Exact Three-Dimensional Elasticity Solutions for Bending of Moderately Thick Inhomogeneous and Laminated Strips under Normal Pressure,'' Int. J. Solids Struct., **32**, pp. 1659–1673.

[32] Wang, Y. M., and Tarn, J. Q., 1994, ''A Three-Dimensional Analysis of Anisotropic Inhomogeneous and Laminated Plates,'' Int. J. Solids Struct., **31**, pp. 497–515.

[33] Tarn, J. Q., and Wang, Y. M., 1995, ''Asymptotic Thermoelastic Analysis of Anisotropic Inhomogeneous and Laminated Plates,'' J. Therm. Stresses, **18**, pp. 35–58.

[34] Cheng, Z. Q., Lim, C. W., and Kitipornchai, S., 1999, ''Three-Dimensional Exact Solution for Inhomogeneous and Laminated Piezoelectric Plates,'' Int. J. Eng. Sci., **37**, pp. 1425–1439.

[35] Cheng, Z. Q., Lim, C. W., and Kitipornchai, S., 2000, ''Three-Dimensional Asymptotic Approach to Inhomogeneous and Laminated Piezoelectric Plates,'' Int. J. Solids Struct. **37**, pp. 3153–3175.

[36] Cheng, Z. Q., and Batra, R. C., 2000, ''Three-Dimensional Thermoelastic Deformations of a Functionally Graded Elliptic Plate,'' Composites, Part B, **31**, pp. 97–106.

[37] Cheng, Z. Q., and Batra, R. C., 2000, ''Three-Dimensional Asymptotic Scheme for Piezothermoelastic Laminates,'' J. Therm. Stresses, **23**, pp. 95–110.

[38] Cheng, Z. Q., and Batra, R. C., 2000, ''Three-Dimensional Asymptotic Analysis of Multiple-Electroded Piezoelectric Laminates,'' AIAA J., **38**, pp. 317–324.

[39] Cheng, Z. Q., and Batra, R. C., 2000, ''Generalized Plane Solution for Monoclinic Piezoelectric Laminates,'' AIAA J., **38**, pp. 335–341.

[40] Tiersten, H. F., 1969, *Linear Piezoelectric Plate Vibrations*, Plenum Press, New York.

[41] Mindlin, R. D., 1974, ''Equations of High Frequency Vibrations of Thermopiezoelectric Crystal Plates,'' Int. J. Solids Struct., **10**, pp. 625–632.

[42] Reddy, J. N., 1997, *Mechanics of Laminated Composite Plates: Theory and Analysis*, CRC Press, Boca Raton, FL.

[43] Aboudi, J., 1991, *Mechanics of Composite Materials—A Unified Micromechanical Approach*, Elsevier Amsterdam.

[44] Pindera, M. J., Aboudi, J., and Arnold, S. M., 1995, ''Limitations of the Uncoupled, RVE-Based Micromechanical Approaches in the Analysis of Functionally Graded Composites,'' Mech. Mater., **20**, pp. 77–94.

[45] Zuiker, J. R., 1995, ''Functionally Graded Materials: Choice of Micromechanics Model and Limitations in Property Variation,'' Composites Eng., **5**, pp. 807–819.

[46] Reiter, T., Dvorak, G. J., and Tvergaard, V., 1997, ''Micromechanical Models for Graded Composite Materials,'' J. Mech. Phys. Solids, **45**, pp. 1281–1302.

[47] Reiter, T., and Dvorak, G. J., 1998, ''Micromechanical Models for Graded Composite Materials: II. Thermomechanical Loading,'' J. Mech. Phys. Solids, **46**, pp. 1655–1673.

[48] Mori, T., and Tanaka, K., 1973, ''Average Stress in Matrix and Average Elastic Energy of Materials With Misfitting Inclusions,'' Acta Metall., **21**, pp. 571–574.

[49] Benveniste, Y., 1987, ''A New Approach to the Application of Mori-Tanaka's Theory in Composite Materials,'' Mech. Mater., **6**, pp. 147–157.

[50] Hatta, H., and Taya, M., 1985, ''Effective Thermal Conductivity of a Misoriented Short Fiber Composite,'' J. Appl. Phys., **58**, pp. 2478–2486.

[51] Rosen, B. W., and Hashin, Z., 1970, ''Effective Thermal Expansion Coefficients and Specific Heats of Composite Materials,'' Int. J. Eng. Sci., **8**, pp. 157–173.

[52] Hashin, Z., and Shtrikman, S., 1963, ''A Variational Approach to the Theory of the Elastic Behavior of Multiphase Materials,'' J. Mech. Phys. Solids, **13**, pp. 213–222.

[53] Murphy, G., 1957, *Properties of Engineering Materials*, 3rd Ed., International Textbook Co., Scranton, PA.

[54] Van Vlack, L. H., 1985, *Elements of Materials Science and Engineering*, 5th Ed., Addison-Wesley, Reading, MA.

[55] Gol'denveizer, A. L., 1969, ''Boundary Layer and Its Interaction with the Interior State of Stress of an Elastic Thin Shell,'' J. Appl. Math. Mech., **33**, pp. 971–1001.

[56] Dauge, M., and Gruais, I., 1998, ''Edge Layers in Thin Elastic Plates,'' Comput. Methods Appl. Mech. Eng., **157**, pp. 335–347.

Simulations of Crack Propagation in Porous Materials

T. Nakamura
Mem. ASME

Z. Wang

Department of Mechanical Engineering,
State University of New York
Stony Brook, NY 11794

Failure propagation behavior of thermally sprayed coatings containing many random pores is investigated. The porous coatings are subjected to either external mechanical loads or residual stresses generated by temperature changes. The failure growth criterion is governed by the critical energy release rate. In our finite element analysis, the cohesive model is used to separate element boundaries during crack propagation in the inhomogeneous materials. The accuracy of the cohesive elements for the quasi-static crack growth is closely evaluated by an error analysis. We have observed that the cohesive elements may artificially increase the model compliance and introduce numerical errors. In order to minimize such errors, the parameters for cohesive model must be chosen carefully. Their numerical convergence and stability conditions with an implicit time integration scheme are also examined. In the porous material analysis, crack propagation is simulated to characterize its unique failure process. It appears a crack tends to propagate along the shortest path between neighboring pores. In addition, crack/pore coalescence mechanism causes the apparent crack length to increase discontinuously. Under thermally loaded conditions, the residual stresses generated by material mismatch in multilayered coatings drive cracks to grow. Using the present crack propagation model, the critical temperature leading to the complete porous coating failure can be approximated. [DOI: 10.1115/1.1356029]

1 Introduction

The microstructure of a thermally sprayed ceramic coating is characterized by the existence of various pores, microcracks, splat boundaries, and unmelted particles. These attributes greatly influence the failure characteristics of coatings. In general, the porous microstructure reduces the overall coating strength as well as its resistance to failure. Thermally sprayed coatings are also anisotropic, adding complexity to their mechanical characterization. There have been numerous experiments to measure the coating's fracture toughness ([1,2]). For plasma sprayed coatings, the toughness is measured to be a small fraction of corresponding value for the bulk material. In this study, we attempt to identify the mechanisms that lead to lower propagation toughness in these porous materials.

In our computational analysis, unique microstructural features of sprayed coatings, described by randomly distributed pores with various sizes and shapes, are included in the finite element models. These models are similar to the ones used to determine the effective moduli of coatings ([3]). Since thermally sprayed coating contains no two pores exactly alike, the pores must be modeled nonuniformly. Thus, the modeling requires a complex procedure to represent actual coatings. To closely study the properties of actual coatings we have modeled ceramic coatings with statistical distributions of pore sizes and shapes that follow those of alumina-titania coatings ([4]).

To simulate crack propagation behavior in a highly inhomogeneous medium, we have implemented cohesive elements in the porous coating model. The cohesive type elements have been used successfully in various dynamic crack propagation and fragmentation analyses ([5–8]). While the cohesive elements were used with noniterative explicit time integration schemes in these analyses, we have implemented the elements in quasi-static analyses with an implicit time integration scheme. In the past, the cohesive elements were also employed in number of quasi-static analyses: for example by Needleman [9], Tvergaard and Hutchinson [10], and Lin et al. [11]. When the cohesive elements are used in conjunction with a common iterative scheme in static analyses, a numerical instability may arise. This aspect and numerical inaccuracy associated with the cohesive model are closely examined in the error analysis presented in this study.

2 Microstructure of Porous Coatings

2.1 Characteristics of Plasma Sprayed Coatings. The process of plasma spraying renders unique coating microstructures which are greatly different from those of corresponding bulk materials. The characteristics of coating microstructures can be summarized as porous lamella structures. The pancake-shaped splat which is about 1–5 μm thick and 10–50 μm in diameter is the basic structural unit of a coating. Inside a splat, perpendicular columnar grain structures can be observed and indicate the gradient direction of the solidification process during cool-down. Due to the nature of the thermal spraying process, various kinds of defects can be observed in coatings ([12,13]). Those defects may lie along splat boundaries and can be caused by weak adhesion between splats. The pores and delaminations and microcracks may grow under certain mechanical and/or thermal load and weaken coatings. Numerous experiments have been carried out to evaluate the porosity of the thermally sprayed coatings. Depending on the coating process and spray parameters, the porosity or the total volume fraction of pores in ceramic coatings may range from less than a few percent to about 20 percent.

The global or average properties of plasma sprayed coatings are usually very different from those of fully dense materials. The effective elastic modulus and fracture initiation toughness of ceramic coatings can be measured by various experimental methods, such as uniaxial tension, four-point bending, indentation, and ultrasonic tests. Due to their microstructure, the modulus as well as toughness of ceramic coatings can be one to two orders of magnitude smaller than the bulk or intrinsic modulus and toughness. Pores and splat boundaries constitute the major reductions. In addition, inhomogeneous phases, impurities, and residual stresses probably contribute to lowering of the material constants.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 14, 1999; final revision, July 26, 2000. Associate Editor: A. Needleman. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2.2 Distributions of Pore Sizes and Shapes. Experimentally measured pore geometry provides not only microstructural details but also the basis for modeling porous coating and studying how coating microstructures affect mechanical properties. There are several precise quantitative studies of pore sizes and shapes in coatings using various techniques, including X-ray tomography. Here our aim is not to exactly model a small section of a coating with a few pores, but is to identify underlying geometrical factors which influence the failure response. For this purpose, collective information on many pores is needed rather than very accurate descriptions of few pores. Such results are available from a stereologically measured data obtained by Leigh and Berndt [4] where a complete set of data on the pore size and shape distributions found in a plasma-sprayed alumina-titania (Al_2O_3–13 wt % TiO_2) coating were presented. According to the measurements, the sizes of most pores fall between 10 and 100 μm and sharp and cracklike pores are abundant in the alumina-titania coating. In general, many of the larger sized pores are cracklike while most small pores are close to circular shape.

3 Crack Propagation Procedure

3.1 Various Methods to Simulate Crack Growth. Computational simulation of a propagating crack is complex and difficult since the crack shape and tip location must be constantly redefined and updated. There are a few classes of approaches to simulate crack propagation. One approach is based on remodeling or remeshing technique. After crack growth rate and direction are determined by the stress intensity factors and/or the maximum opening stress, remeshing of the computational model was carried out at every increment. Continuous remeshing model is shown by Wawrzynek and Ingraffea [14] and more recently by meshfree or so-called element-free Galerkin method which was introduced by Belytschko et al. [15]. The advantage of this approach is that crack path can be chosen arbitrary but the computational cost tends to be high due to remodeling at every increment. Another approach uses a continuous array of special elements to represent the path of crack growth. The material properties of these elements are based on observed constitutive behaviors of an actual crack in propagation. One of such models is ''smeared crack model'' developed by Dagher and Kulendran [16]. The element property is governed by the observed linear strain softening in the wake of crack propagation. A similar method was proposed by Padovan and Guo [17] and was further modified by Padovan and Jae [18]. Their model implements a movable crack template which moves within a global field. The template contains special elements and as soon as the crack tip passes through such elements, their stiffness perpendicular to the path is vanished to simulate the effect of traction-free crack surfaces. The trace of these wake elements showcases the trail of crack growth.

Yet another approach involves a very little or no remodeling of original mesh. However, possible crack propagation path is restricted along element boundaries. Many traditional dynamic crack propagation simulations were carried out with the node release/separation technique and the moving singular element method [19,20]. The latter method requires a limited remeshing near the tip although the crack path remains along the predefined boundaries. In the node release method, when a certain crack growth criterion is met, the displacement constraint at the crack-tip node is removed and replaced by an equivalent force. During several subsequent increments, this force is reduced and the crack tip is advanced to the next node. A similar method, but with a different procedure to create new crack surfaces along element boundaries was proposed by Xu and Needleman [5]. Their method places a special spring element between two nodes to be separated. Essentially, element boundaries are treated as cohesive surfaces and separation of such surfaces simulates a crack growth. A major advantage of this method is that the cohesive elements can be included to an existing code without a major reprogramming. Also the cohesive model can handle simultaneous multiple crack growths and its load or time increments are not tied to the crack growth behavior. Recently, Gao and Klein [21] and Zhang et al. [22] introduced another type of cohesive model. In their virtual internal bonding method, constitutive relation was modified to include a failure criterion and a crack propagation occurs within elements instead of along element boundaries.

3.2 Derivation and Formulation of Cohesive Model. In this study, we follow the model which was originally developed by Xu and Needleman [5]. Camacho and Ortiz [6] also introduced a slightly different type of cohesive model in their successful analysis of dynamic impact of brittle material. In general, the cohesive element acts as a nonlinear spring and the crack growth criterion is embedded in its constitutive equation. The traction-displacement relation of the spring element is defined through a potential function Φ. Since the energy required to separate two nodes is equivalent to the integral of traction over displacement, this Φ directly relates to the fracture energy of growing crack. For two-dimensional cracks, the crack may grow under mixed-mode condition, and this potential function includes both the normal (Mode I) and tangential (Mode II) contributions. The form is shown as ([5]),

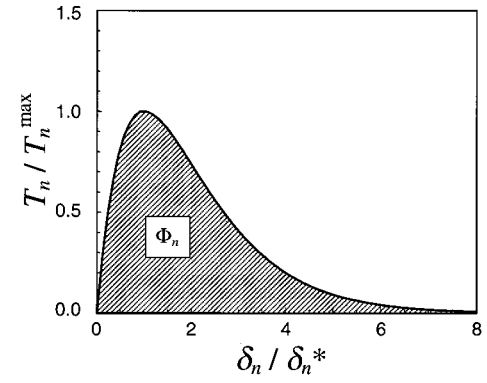
$$\Phi(\delta_n,\delta_t)=\Phi_n\left\{1+e^{-\delta_n/\delta_n^*}\left[\left(1+\frac{\delta_n}{\delta_n^*}\right)(q-1)-\left(1+\frac{\delta_n}{\delta_n^*}\right)qe^{-(\delta_t/\delta_t^*)^2}\right]\right\}. \tag{1}$$

Here δ_n and δ_t are the displacement components normal and tangential to the crack plane, respectively. Also δ_n^* and δ_t^* are the reference displacements, Φ_n is the fracture or the total separation energy (per unit advance) required under pure Mode I condition (i.e., $\delta_t=0$) and q represents the ratio of the Mode II and Mode I fracture energies (i.e., $q=\Phi_t/\Phi_n$). If fracture toughness under Mode II condition is higher than that of Mode I condition, then $q>1$. The normal and shear traction components relate to the displacement components as

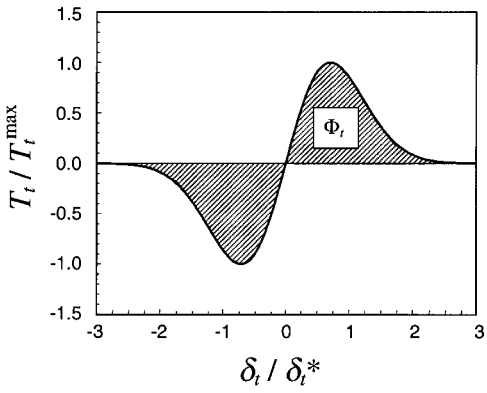
$$\begin{aligned}T_n&=\frac{\partial\Phi}{\partial\delta_n}\\&=\frac{\Phi_n}{\delta_n^*}e^{-\delta_n/\delta_n^*}[(\delta_n/\delta_n^*)e^{-(\delta_t/\delta_t^*)^2}\\&\quad+(1-q)(\delta_n/\delta_n^*)(1-e^{-(\delta_t/\delta_t^*)^2})]\end{aligned}$$

$$T_t=\frac{\partial\Phi}{\partial\delta_t}=\frac{2\Phi_n\delta_t}{(\delta_t^*)^2}q(1+\delta_n/\delta_n^*)e^{-\delta_n/\delta_n^*}e^{-(\delta_t/\delta_t^*)^2}. \tag{2}$$

Normalized displacement and traction relationships according to the above equations are shown in Fig. 1. For the normal component, the maximum traction $T_n^{\max}=\Phi_n/e\delta_n^*$ occurs at $\delta_n=\delta_n^*$ while the shear traction is maximum $T_t^{\max}$ at $\delta_t=\delta_t^*/\sqrt{2}$. Note the reference tangential displacement is given as $\delta_t^*=\sqrt{2e}\,\delta_n^*qT_n^{\max}/T_t^{\max}$. In carrying out the crack growth analysis, the cohesive elements are placed along every element boundary where fracture may occur. Essentially these nonlinear springs tie nodes of adjacent elements. A potential source of error associated with the cohesive elements is the *compliance* introduced between regular elements. If the springs have a finite stiffness, the extra compliance is present along the element boundaries. When the total amount of this compliance is large, the overall structural response is artificially softened and an error arises. Alternatively, the cohesive energy stored in these elements can overwhelm the strain energy of structure [22]. In order to quantify the possible error associated with the cohesive model, we have carried out a detailed error analysis.

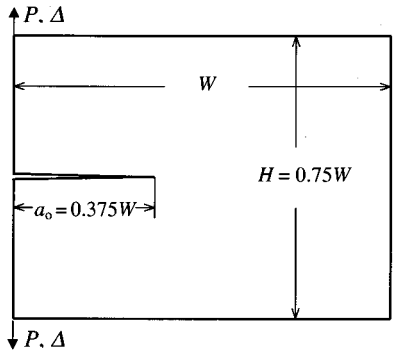


(a)

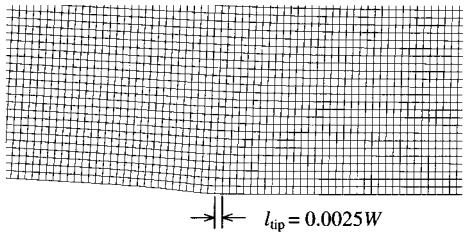


(b)

Fig. 1 Relationships between normalized displacement and traction used cohesive model. The shaded areas represent the fracture energy. (*a*) Normal component for Mode I, (*b*) tangential component for Mode II.



(a)



(b)

Fig. 2 (*a*) Schematic of edge-crack panel used in the error analysis. (*b*) Top-half of finite element mesh near crack tip. All the elements in this zone are shaped square with the side length equal to 1/400 of the panel width.

4 Error Analysis of Cohesive Model

4.1 Computational Model. If the cohesive model is designed to simulate crack growth in brittle materials, its accuracy can be measured with known solutions of growing linear elastic cracks. Here we have carried out an error analysis using a simple crack model under pure Mode I loading condition. Suppose the separation energy Φ_n in (1) is kept constant everywhere in the model, the crack must propagate under constant energy release rate. This means once the energy release rate $\mathcal{G}$ reaches the critical value $\mathcal{G}_c(=\Phi_n)$ and the initiation occurs, the crack continues to grow at $\mathcal{G}=\mathcal{G}_c$. This condition is simulated with an edge-crack specimen containing cohesive elements.

Under pure Mode I condition ($\delta_t=0$), the normal traction-displacement relation (2) reduces to

$$T_n=\Phi_n(\delta_n/\delta_n^{*2})e^{-\delta_n/\delta_n^*}. \tag{3}$$

It is clear from the above equation, complete nonlinear spring properties are defined by Φ_n and δ_n^*. Since Φ_n corresponds to the material's critical fracture toughness $\mathcal{G}_c$, the reference displacement δ_n^* is the only parameter which can be set arbitrarily. Once δ_n^* is chosen, the maximum traction T_n^{max} as well as the stiffness of the cohesive model are also defined. According to (3), a small δ_n^* translates to a large initial $T_n-\delta_n$ slope or spring stiffness. Since the stiffness of element boundaries is ideally infinite before crack growth, it suggests that δ_n^* be set as small as possible (i.e., $\delta_n^*\to 0$). In fact, any compliance ($\delta_n^*\neq 0$) introduced by the cohesive elements in a region without cracks may erroneously reduce the entire model response. However, in quasi-static analyses with a common iterative scheme (e.g., Newton-Raphson), it is difficult to achieve numerical convergence under $\delta_n^*\to 0$ condition. The numerical instability arises from the abrupt change in the tangent/slope of traction-displacement curve at $\delta_n=\delta_n^*$ as shown in Fig. 1(*a*). As $\delta_n^*\to 0$, the rate of change increases. This means, while a smaller δ_n^* is desirable in reducing the artificial compliance, it worsens the convergence rate during iterative time integration. In general, the numerical instability cannot be resolved with taking smaller load increments. Furthermore, the stability condition worsens when smaller elements are used near crack tips. We note that the linear or extrinsic traction-displacement relation (T_n^{max} at $\delta_n=0$) such as developed by Camacho and Ortiz [6] limits any unwanted compliance in the model. However, in many crack problems, such a relation cannot be used in conjunction with an implicit code where iterations are carried out to achieve the equilibrium. Regardless of models, a sudden introduction of active cohesive elements initiates a numerical instability. In order to avoid this condition, an initial stiffness of cohesive element must be set finite. Many past analyses with cohesive elements were carried out with explicit central difference scheme under dynamic conditions where no iterations were needed.

Here we tested several different values of δ_n^* to investigate the level of error generated by nonzero δ_n^*. In the error analysis, a cantilever-type edge-crack specimen as shown in Fig. 2(*a*) is considered. The specimen is loaded by the displacement Δ prescribed at the edges. When the energy release rate reaches the critical value, the crack begins to grow. During the propagation, the energy release rate remains at the constant level in an ideally brittle solid. Due to the symmetry condition, only the top half of the specimen is modeled. A total of 19,200 four-noded elements is used to construct the finite element mesh. As shown in Fig. 2(*b*), elements in the vicinity of the crack tip are kept square at a constant scale. The side-length of those elements are set at $l_{tip}=0.0025W$, where W is the width of the specimen. The model dimensions are selected carefully to minimize the boundary effect.

First, we chose the crack to propagate over 100 elements or about 25 percent of the width W. The lead to the initial crack length to be $a_0=0.375W$. The half-height is also as $0.375W$ so that the growing crack tip is at least the same distance away from outer surface. Suppose the K-field size is about 1/10 of this length, then it spans at least 15 elements in any directions. This means the solutions should be nearly independent of geometry and the near-tip element size is the only *length scale* that influences the accuracy of the cohesive model.

4.2 Dimensional Analysis. Under Mode I loading condition, the parameters that can influence the computed results are the Young's modulus E, the Poisson's ratio ν, the normal separation energy Φ_n, the reference displacement δ_n^*, and the near-tip element length l_{tip}. Based on dimensional consideration, the error function $\mathcal{E}$ can be expressed as

$$\mathcal{E}=\mathcal{E}(\delta_n^*/l_{\text{tip}}, E/\Phi_n l_{\text{tip}}, \nu). \tag{4}$$

The primary error arises in the cohesive model due to the presence of extra compliance introduced along element boundaries. This means the magnitude of the error scales with the relative stiffness of the cohesive element $(\partial T_n/\partial \delta_n)$ with respect to the material modulus (E). The average stiffness of cohesive element before reaching the maximum traction is $T_n^{\max} l_{\text{tip}}/\delta_n^*$ or it is proportional to $\Phi_n l_{\text{tip}}/\delta_n^{*2}$. Using this expression, the error associated with the cohesive elements can be given as a function of the normalized stiffness as

$$\mathcal{E}\approx\mathcal{E}\left(\frac{\Phi_n l_{\text{tip}}}{E\delta_n^{*2}}\right). \tag{5}$$

The error in the energy release rate can be obtained by considering the cumulative stiffness S_{coh} of all cohesive elements. We can assume normalized S_{coh} to be approximated with one-term power-law expression as

$$\frac{S_{\text{coh}}}{S}\approx\alpha\left(\frac{\Phi_n l_{\text{tip}}}{E\delta_n^{*2}}\right)^{\beta}. \tag{6}$$

Here S is the overall structural stiffness without cohesive elements, α and β are the factors which depend on locations and total number of cohesive elements and assumed to be $\alpha>0$ and $\beta>0$. Any effects of the Poisson's ratio are implicitly included in these terms. Using the above approximation, the dimensionless error in $\mathcal{G}$ can be expressed as

$$\mathcal{E}_G=\frac{\mathcal{G}_{\text{ideal}}-\mathcal{G}_{\text{coh}}}{\mathcal{G}_{\text{ideal}}}=\frac{S}{S+S_{\text{coh}}}\approx\left[\alpha\left(\frac{\Phi_n l_{\text{tip}}}{E\delta_n^{*2}}\right)^{\beta}+1\right]^{-1}. \tag{7}$$

Here, $\mathcal{G}_{\text{ideal}}$ corresponds to the exact solution for ideally brittle solids and a displacement-controlled loading condition is assumed in deriving the expression. In (7), the error is directly influenced by the reference displacement δ_n^* chosen for the cohesive elements. Note as $\delta_n^*\to 0$, the error vanishes. The values of α and β can be determined from the computations of the test problem in comparison with the ideal solutions. We also note that the above expression is valid for the error evaluation of other types of traction-displacement relations for cohesive elements. For example, if the relation is given by bilinear curves as by Guebelle and Baylor [7], the displacement at the peak load can be substituted for δ_n^* in (7).

4.3 Computational Results. In placing the cohesive elements within the crack model, three separate cases are considered. They differ in domain sizes where the elements are included. The domains are, Case A—only along the crack propagation path (i.e., symmetry line), Case B—within a smaller domain surrounding the crack tip, and Case C—within a larger domain surrounding the crack tip. These domains are illustrated in Fig. 3. The cohesive spring elements are added along every element boundary within the respective domains. There are 150, 4,437, and 10,577 cohesive elements in Cases A, B, and C, respectively. It is expected that as the total number of cohesive element increases, the error also rises. During the analysis, the energy release rate is calculated at every time increment using the domain integral expression of $\mathcal{G}$ ([23]). In each model, computations are carried out with various δ_n^* to determine the error parameters α and β in (7). Also in all cases, q in (1) is set to unity.

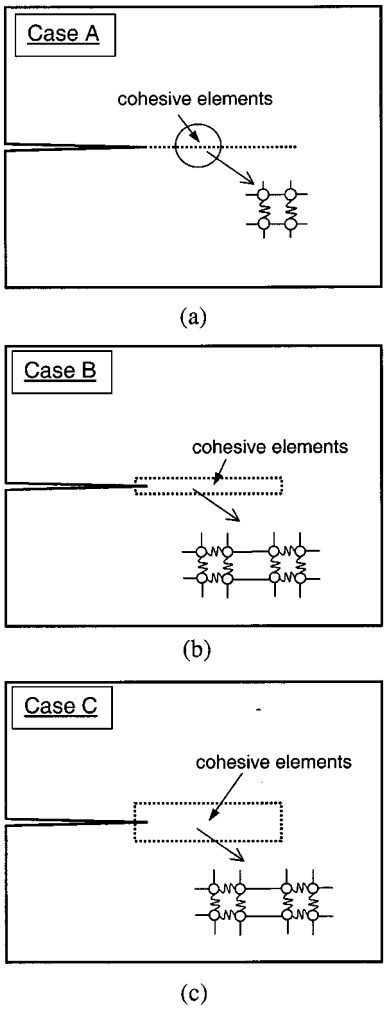


Fig. 3 Three cases with different domains where cohesive elements are placed. (*a*) Case A with cohesive elements only along the crack path; (*b*) Case B with cohesive elements in $0.04W\times 0.38W$ domain; (*c*) Case C with cohesive elements in $0.1W\times 0.38W$ domain.

The results of the energy release rate with the cohesive elements along crack path (Case A) are shown in Fig. 4(*a*). As the prescribed displacement increases, $\mathcal{G}$ increases until it reaches the critical value $\Phi_n=\mathcal{G}_c$. For the ideally brittle solid, the crack initiates and continues to grow at constant $\mathcal{G}$ as shown in the figure. Although this condition may be simulated with $\delta_n^*\to 0$ in the cohesive elements, it is difficult to obtain converged solutions for a very small δ_n^* in an implicit integration (e.g., Newton-Raphson). In the present model, the numerical instability occurs when $\delta_n^*/l_{\text{tip}}<1.4\times 10^{-3}$. When smaller values are assigned to δ_n^*, the equilibrium convergence cannot be achieved even with very small increment sizes. The maximum difference between the ideally brittle solid and the cohesive models occurs at the crack growth initiation point $(\Delta/l_{\text{tip}}=0.0498)$. Based on the results of different

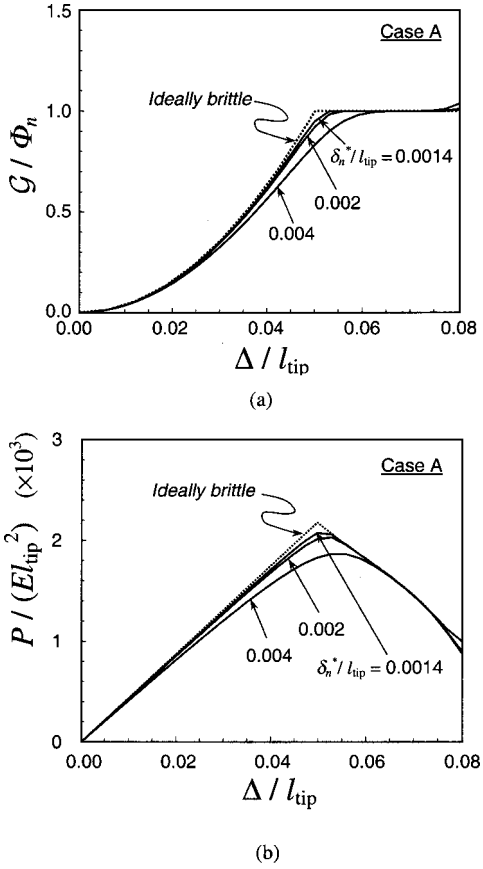


Fig. 4 Computed results for various reference displacements δ_n^* in Case A. (*a*) Normalized energy release rate shown as a function of normalized prescribed displacement, (*b*) normalized load shown as a function of normalized prescribed displacement.

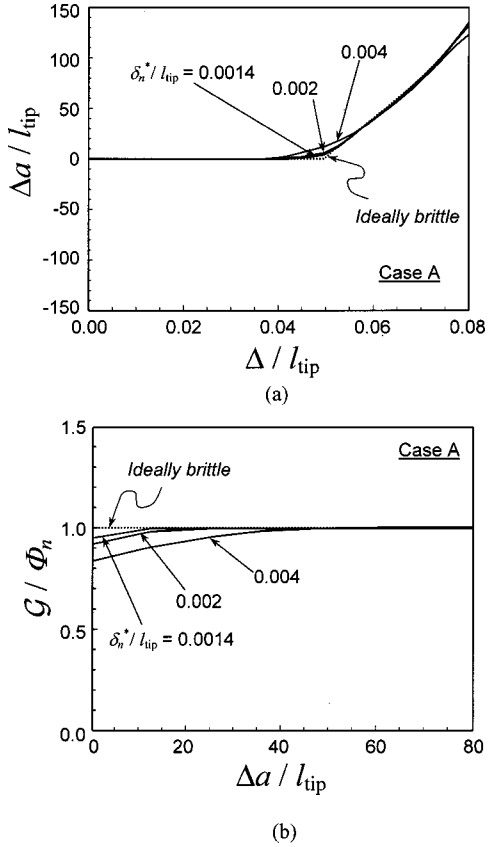


Fig. 5 Computed results for various reference displacements δ_n^* in Case A. (*a*) Normalized crack advanced distance shown as a function of normalized prescribed displacement, (*b*) normalized energy release rate shown as a function of normalized crack advance distance.

values of δ_n^*, the parameters in (7) can be approximated as $\alpha=27$ and $\beta=0.6$. With these values, the error at the initiation can be expressed as

$$\mathcal{E}_G^{\max}=\left.\frac{\mathcal{G}_{\text{ideal}}-\mathcal{G}_{\text{coh}}}{\mathcal{G}_{\text{ideal}}}\right|_{\text{initiation}}\approx\left[27\left(\frac{\Phi_n l_{\text{tip}}}{E\delta_n^{*2}}\right)^{0.6}+1\right]^{-1}. \quad (8)$$

Note that the above error is the maximum error at the crack growth initiation. In fact, as the crack propagates substantially, the energy release rate of every cohesive model converges to the ideal case. Although not shown here, we have tested other values of δ_n^* and found the expression (8) to hold for a wide range of δ_n^*. Different values of E were also shown to agree well with the error estimate. In addition, we have inspected the load-displacement relation of the growing crack model as shown in Fig. 4(*b*). Here P is the reaction force at the prescribed displacement. Without any cohesive elements, the $P-\Delta$ curve has a sharp peak at the initiation point ($\Delta/l_{\text{tip}}=0.0498$). When the cohesive elements are included, the structural stiffness is reduced and the maximum load is lowered as shown in the figure. The increased compliance of the cohesive models is also evident from the reduced slopes of $P-\Delta$ curves. The difference in the peak load can be shown with an error expression similar to (7).

The accuracy of crack growth Δa is examined in Fig. 5(*a*). During crack propagation, there are usually several active cohesive elements with $\delta_n>0$ and $T_n>0$ along the crack path and the determination of the crack-tip location is difficult. In order to clarify this problem, we have operationally defined the tip location by the following method. First we assume that Mode I K-field surrounds elements near the crack tip and the opening stress ahead of the crack to be square-root singular in radial distance. With four-noded isoparametric elements in our model, an equivalent nodal force due to the singular stress can be determined by integrating the stress with linear interpolation/weighting functions. The force can be calculated for a node located at an arbitrary distance away from the tip. The maximum force is attained when the node is located $l_{\text{tip}}/3$ ahead of the presumed crack-tip location. If the crack propagates and coincides with the node, the nodal force is reduced to 86.6 percent of the maximum value. According to Fig. 1(*a*), $0.866T_n^{\max}$ occurs at $\delta_n/\delta_n^*=1.64$ in the cohesive elements. Also at this point, the fracture energy is 48 percent of the critical energy or $\Phi=0.48\Phi_n$. Since Φ is also applicable in mixed-mode cases, this value is monitored throughout our analyses to identify the location of crack tip. We note that Xu and Needleman [5] used $\delta_n/\delta_n^*\approx 1$ and Needleman and Rosakis [24] used $\delta_n/\delta_n^*=5$ to define the crack-tip location. The crack extension is shown as a function of the prescribed displacement in Fig. 5(*a*). Better agreements with the ideally brittle solution can be observed with smaller δ_n^*. With larger δ_n^* (e.g., $\delta_n^*/l_{\text{tip}}=0.004$), the initial crack growth occurs prematurely before $\mathcal{G}$ reaches the critical value. However, all models converge to the ideal solution at larger Δ. These results confirm the suitability of the definition used for the crack-tip location. In Fig. 5(*b*), the normalized energy release rate is shown as a function of crack growth normalized by the element length. The results show $\mathcal{G}$ of the cohesive models approaches the ideal solution at larger crack growth. In fact, after growth over several element lengths, the error is significantly reduced.

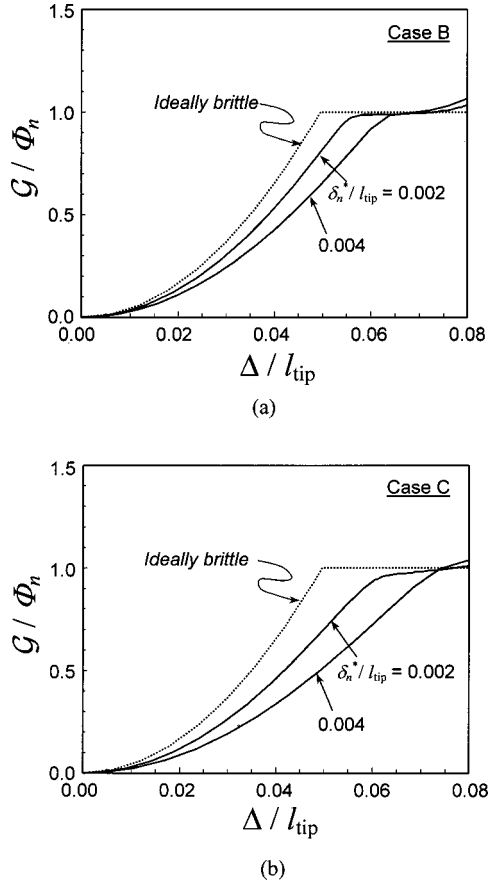


Fig. 6 Normalized energy release rate for various reference displacements δ_n^*. (*a*) Case B, (*b*) Case C.

The computations are also carried out for the models with larger cohesive element domains (Cases B and C). Since much more cohesive elements are present, the results are expected to deteriorate. Only the energy release rate as a function of the prescribed displacement is shown in Fig. 6. Unlike Case A, the calculation diverges when $\delta_n^*/l_{tip}=0.0014$ is prescribed in both cases. In Case B, the results still coincide with the ideal solution but at a larger Δ as shown in Fig. 6(*a*). The errors in $\mathcal{G}$ at the initiation point are also much greater than those of Case A. Similar results are obtained for Case C as shown in Fig. 6(*b*). Based on these results, we have again computed the error parameters in (7). The proportional factor α is 9.5 and 5.3 for Cases A and B, respectively. The power coefficient β appears to remain constant at 0.6 in both cases.

The present analysis has quantified the error due to the cohesive elements in a crack propagation analysis. The size of error is related to the total number of cohesive elements as well as the choice of reference displacement δ_n^*. In order to minimize the error, the reference displacement δ_n^* must be kept small. However, the possible size of δ_n^* is restricted by the convergence condition of the implicit time integration scheme. Prior to any crack propagation analyses, an error for a given δ_n^* can be estimated with the simple formula introduced in (7). It was observed that while the parameter α depends on the total number of cohesive elements, the other parameter β appears to be constant at 0.6. The parameter α decreases with greater number of cohesive elements in the model here. Although the error analysis is carried out with an edge-crack model, the present error estimates should be applicable to other models as long as the K-field is sufficiently larger than near-tip element sizes.

5 Crack Propgation in Porous Materials

5.1 Computational Models for Porous Coatings. In order to construct the unique microstructures of plasma sprayed coating, it is necessary to establish a model which allows for random distributions of numerous pores with various sizes and aspect ratios. Within the model, pores are placed at random locations, and every pore is assumed to possess three additional geometrical attributes: pore size/area, pore shape, or aspect ratio, and pore orientation with respect to the spray direction. A computational program has been developed to generate finite element models which simulate coatings with such attributes. This procedure is a modification of the one developed by Nakamura et al. [3] to investigate the effective properties of porous materials. This procedure is briefly discussed below.

First, all the pore shapes are idealized to be hexagonal. The size/area is assumed to range from A_{min} to $10A_{min}$ where A_{min} is chosen to be about 80 μm^2. Any pores which have less area have a very small influence on the overall response. Also the aspect ratios of pores are set to vary from $a/b=1$ to 10 (a and b are major and minor axes). Furthermore, instead of allowing continuous variations in the area and the aspect ratio, we have used five distinct pore areas, $A/A_{min}=1$, 2, 4, 6, and 10 and five aspect ratios $a/b=1$, 2, 4, 6, and 10. Therefore, there are 25 possible types of pores varying in area and shape. Categorizing pores into 25 types enables us to assign various weight factors to the pore types. The weights are set to the pore size and shape distributions

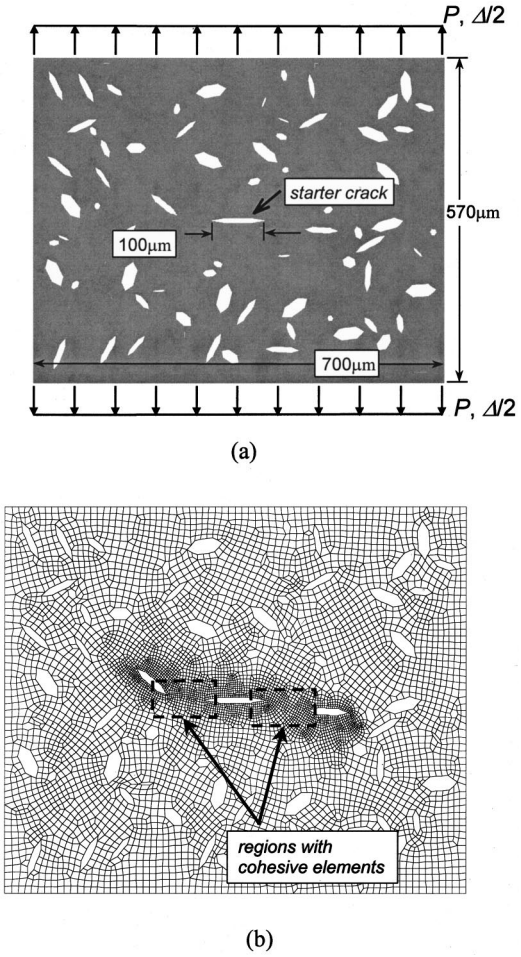


Fig. 7 (*a*) Schematic of panel with porous material under tensile load. The starter crack is placed in the center of the panel. (*b*) Finite element mesh. Regions with porous elements are indicated.

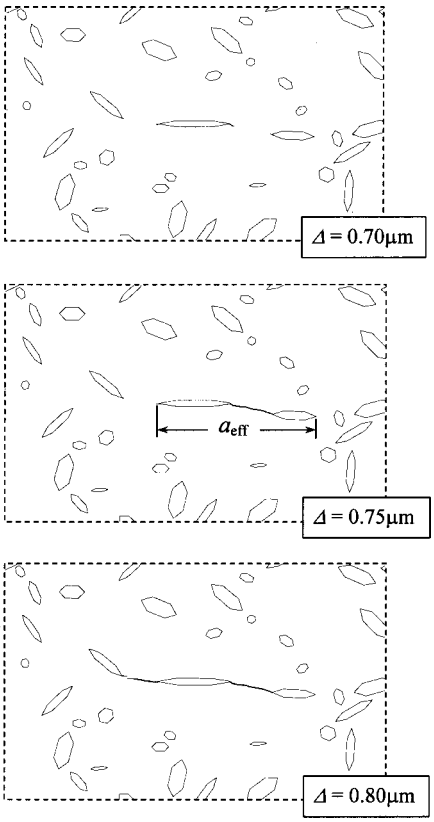


Fig. 8 Sequences of crack growth at three levels of prescribed displacements. Only the region near the starter crack is shown for clarity. The starter crack grows toward neighboring pores.

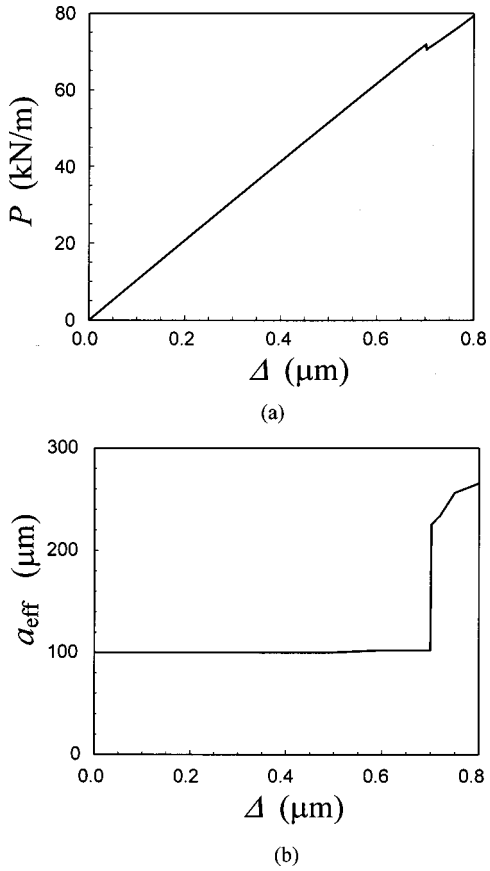


Fig. 9 Computed results of porous material under tensile load. (*a*) Load versus displacement. A small drop in the load is due to a large jump in crack length. (*b*) Effective crack length (crack length profile) versus displacement.

of alumina-titania coatings given by Leigh and Berndt [4]. In addition to the variations in shapes and sizes, every pore is rotated by an arbitrary angle. In carrying out the analysis, the mesh design has shown to possess a significant effect on the computational accuracy. The elements near the crack path must be sufficiently small to capture accurate crack propagation behavior. We have tried several mesh designs to minimize mesh dependence. However, detailed mesh convergence and error analyses are not carried out with the porous models due to the computational limitation. The inclusions of cohesive elements significantly increase the model size and the computations would require beyond our available computational resources.

All the solid elements are chosen to be four-noded plane-strain elements. The matrix or pore-free modulus is chosen from the nano-indentation test data and set as 150 GPa ([25]). Although the bulk elastic modulus of alumina-titania is much higher, splat boundaries, impurities, and other factors unique to thermally sprayed coatings contribute to this lower modulus. The matrix Poisson's ratio is chosen to be $\nu=0.25$. In the first case of the coating model analysis, the crack propagation under far-field uniaxial tensile condition is examined. In the second case, the crack is propagated by the residual stresses generated by thermal mismatch in a multilayered model.

5.2 Crack Propagation Under Tensile Load. For the simulation of crack propagation under tensile load, a rectangular plate as shown in Fig. 7(*a*) is considered. The shapes and sizes of pores follow the measured distribution of sprayed alumina-titania. There is a total of 72 pores in the model. The location of each pore is chosen arbitrarily using a random generator program except for the horizontal ''starter pore/crack'' located in the center of the model. The starter pore is placed to initiate crack propagation and it has the length $a=100\,\mu$m and aspect ratio $a/b=10$. The porosity of the model is 7.5 percent and the effective elastic modulus and Poisson's ratio are separately calculated as $E_{\text{eff}}=108$ GPa and $\nu_{\text{eff}}=0.21$, respectively. These values are obtained from the extension and lateral contraction of the model under uniaxially loaded condition. The 28 percent drop in the Young's modulus is attributed to the pores in the material. A finite element mesh is constructed with about 10,000 elements as shown in Fig. 7(*b*). The average element size near the starter pore is 3 μm. Following the results of the error analysis, cohesive elements are placed within selected regions instead of the entire plate. Since the crack propagation is designed to occur at the center, two separate domains at both sides of the starter crack are chosen for the cohesive elements as shown in Fig. 7(*b*). The critical fracture energy is assumed to be $\Phi_n=10\,\text{J/m}^2$ and the minimum reference displacement before the instability is found to be $\delta_n^*=0.1\,\mu$m. Unlike the error analysis, the element sizes are not small enough compared to the pore/crack sizes and the assumed K-field size. This means the error approximation formula is not applicable here. However, if respective values are assigned, the maximum error is about 30 percent in (8).

The porous material is loaded gradually by increasing the displacement along the top and bottom boundaries of the model. The sequences of the crack propagation are illustrated in Fig. 8. The starter crack begins to grow when the prescribed displacement reaches $\Delta=0.6\,\mu$m. Initially the right tip propagates toward the neighboring pore on the right side. At about $\Delta=0.7\,\mu$m, the crack and the pore coalesce to form a longer crack. The effective crack length which measures the crack length projected onto the horizontal axis is denoted by a_{eff}. At higher load, the crack grows toward the nearest pore on the left side. However, the coalescence

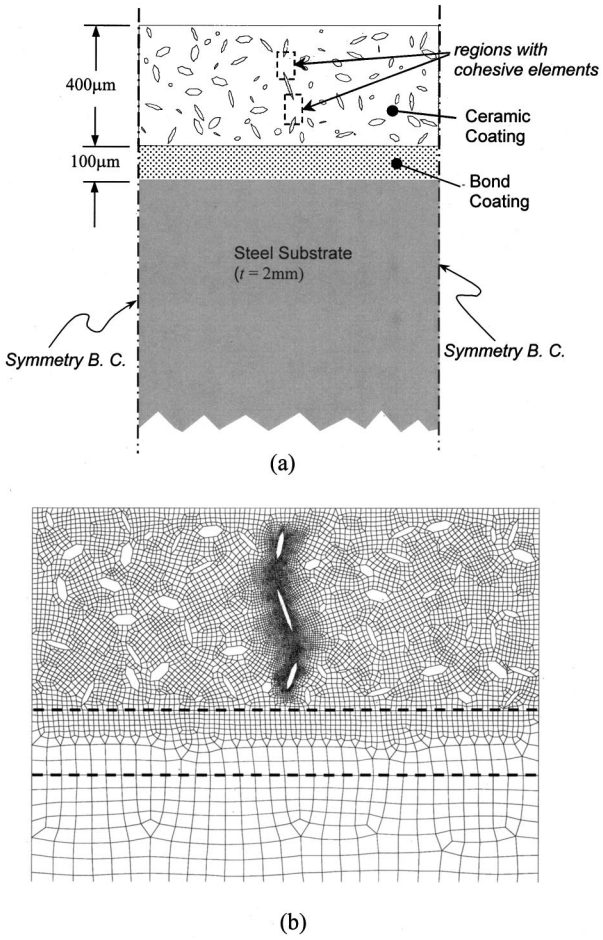


Fig. 10 (*a*) Schematic of multilayered model with porous coatings. Regions with cohesive elements are indicated. (*b*) Top part of finite element mesh for the multilayered model.

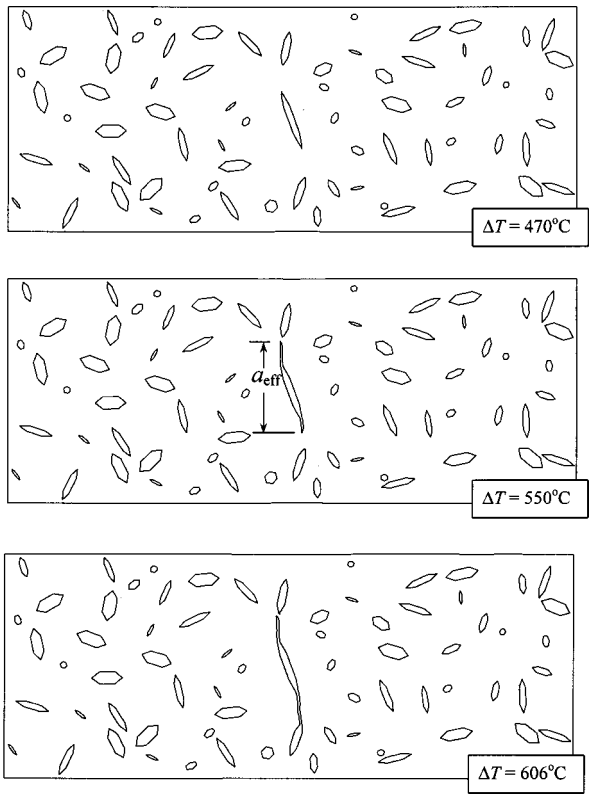


Fig. 11 Sequences of crack growth at three levels of temperatures. Only the ceramic coating is shown for clarity. The starter crack grows toward neighboring pores.

with this pore does not occur during the calculation. One solid element is trapped between the cracks extending from the center crack and the left pore. The relationship between the reaction force and prescribed displacement is shown in Fig. 9(*a*). Here the relation is nearly linear except near $\Delta = 0.7\ \mu$m where a small load drop is observed. This point corresponds to the coalescence of the crack and the pore. The effective crack length is also shown as a function of prescribed displacement in Fig. 9(*b*). Due to the porous structure, the crack growth occurs nonuniformly. A large jump in a_{eff} corresponds to the crack coalescence with the neighboring pore.

5.3 Crack Propagation Under Thermal Load. Many thermally sprayed ceramic coatings are used as thermal barriers in high-temperature environments. The porous microstructure provides increased insulation for the substrate material. In general, metallic bond coat is sprayed onto the substrate followed by ceramic coating. A typical multilayered model for the thermal barrier coating is illustrated in Fig. 10(*a*). Here the width of the model is 1 mm while the steel substrate has the thickness of 2 mm. These dimensions are large enough so that the boundaries have limited effect on the crack propagation behavior. In addition, both sides of the model are constrained to remain straight to represent the symmetry condition.

Due to the thermal stresses generated by the material mismatch, a crack propagation can occur under temperature changes. The ceramic coating is assumed to be elastic while the bond and substrate materials are modeled as elastic-perfectly plastic. The matrix properties of the ceramic (Al_2O_3–TiO_2) is $E = 150$ GPa, $\nu = 0.25$, $\alpha = 8.2 \times 10^{-6}$ where α is the coefficient of thermal expansion. The properties of bond coat (NiCrAlY) and the steel substrate are $E = 80$ GPa, $\nu = 0.25$, $\alpha = 13 \times 10^{-6}$, $\sigma_o = 200$ MPa, and are $E = 80$ GPa, $\nu = 0.3$, $\alpha = 12.5 \times 10^{-6}$, $\sigma_o = 800$ MPa, respectively. Here σ_o is the yield stress. All the material properties are temperature and time independent. For the ceramic coating, the shape and size distributions of pores again follow those of alumina-titania. There are total of 72 pores in the model and the porosity is 7.5 percent. As in the tensile load case, a slightly tilted starter pore with $a = 100\ \mu$m is placed in the center of the model. The effective elastic modulus and Poisson's ratio are calculated as $E_{\text{eff}} = 108$ GPa and $v_{\text{eff}} = 0.21$, respectively. To facilitate the crack growth, the starter pore is oriented close to the vertical direction. Under thermal loading, large residual stress is expected in the direction parallel to the layer boundaries. There is total of about 13,000 four-node elements and the average element size near the starter pore is 2 μm. Finite element mesh for the top part of the multilayered model is shown in Fig. 10(*b*). Cohesive elements are placed within the two separate regions at both sides of the starter crack as shown in Fig. 10(*a*). The critical fracture energy in tension is again chosen to be $\Phi_n = 10$ J/m^2 and the reference displacement is set at $\delta_n^* = 0.35\ \mu$m, which is the minimum displacement for convergence. During the calculation, the temperature is uniformly increased to generate tensile residual stresses within the ceramic coating. No temperature variation within the model is considered in this analysis.

The evolution of crack growth within the ceramic coating at different temperatures is shown in Fig. 11. Crack growth initially starts when the temperature reaches about 460°C. Both tips of the starter crack slowly grow toward the neighboring upper and lower pores, respectively. The coalescence of the crack and the lower pore occurs at about $\Delta T = 600$°C. The residual stress within the ceramic layer for the temperature range $300°\text{C} < \Delta T < 700°\text{C}$ is

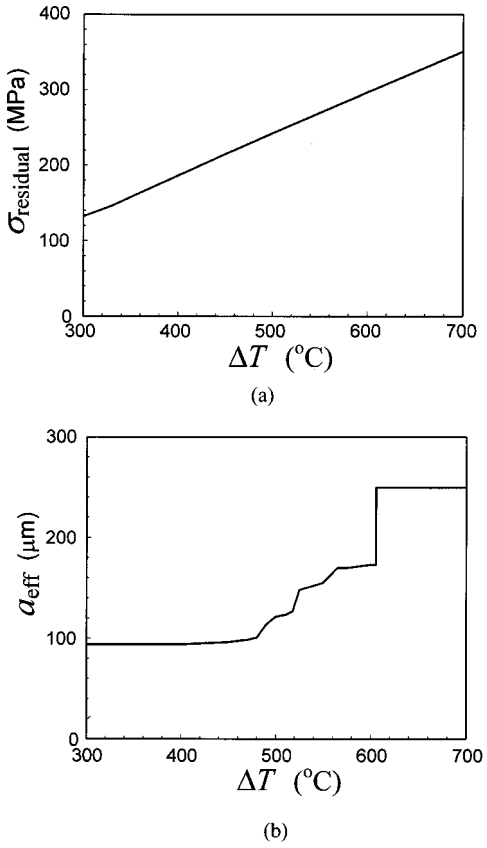


Fig. 12 Computed results of porous material under temperature increase. (*a*) average residual stress within porous ceramic coating; (*b*) effective (apparent) crack length as a function of temperature.

shown in Fig. 12(*a*). The normal stress along the horizontal direction is averaged over the thickness of the ceramic layer. As the temperature rises, the residual thermal stress increases linearly. The effective crack length is also shown as a function of the temperature in Fig. 12(*b*). As in the tensile model, the growth is highly discontinuous. A big jump occurs when the coalescence occurs near $\Delta T = 600°C$. According to the present model, the estimated crack length exceeds a half of the coating thickness (400 mm) when ΔT is about 600°C. This temperature may be regarded as the failure temperature of ceramic coating. We note that any rate-dependent effects at high temperature probably reduce the residual stresses and may slow the crack propagation rate in actual coatings. In addition, temperature variation through thickness can also alter the crack propagation behavior.

6 Discussions

Simulations of crack propagation within porous solids are carried out to explore the fracture behavior in a highly inhomogeneous material. Unlike homogeneous solids, the crack propagation is nonuniform and influenced by pore arrangements. In the analysis, cohesive elements are implemented to separate element boundaries during crack growth under quasi-static condition. The suitability and accuracy of the cohesive elements are closely examined in our error analysis. In general, nonlinear constitutive relation of cohesive element initiates instability and creates convergence problems during equilibrium iterations. The numerical difficulties can be somewhat alleviated by introduction of artificial damping as shown by Zhang et al. [22]. However, we have found a sufficient damping to overcome the instability can also deteriorate the accuracy of calculated results. Alternatively, the coefficient of damping must be kept very low to avoid the error associated with the artificial viscosity.

The cohesive model introduced by Xu and Needleman [5] offers a convenient approach to simulate crack propagation. The model does not require a major modification of existing finite element codes and only inclusion of nonlinear spring elements is needed. However, the cohesive elements may introduce unwanted compliance and reduce the accuracy of calculated results. In the present study, the accuracy of cohesive elements in brittle crack propagation is carefully investigated. A formula for the approximated error is also introduced. The formula should be applicable for other cohesive models if their near-tip elements are sufficiently small compared to the K-field sizes. Our results suggest the initial slope of traction-displacement curve to be set as steep as possible to minimize the error. However, the slope or the reference displacement is also restricted by the convergence criterion stated above. Unfortunately, other types of traction-displacement relation cannot alleviate this difficulty. In general, the traction-based fracture initiation criterion introduced by Camacho and Ortiz [6] also causes numerical instability when it is used with an implicit time integration scheme (e.g., Newton-Raphson) under quasi-static conditions.

The analysis of crack propagation in porous materials offers a means to estimate the critical failure load or temperature. Presence of various pores in thermally sprayed ceramic coatings not only lowers the material stiffness but also reduces the fracture toughness. Our numerical model has simulated crack/pore coalescence during propagation. It appears that the crack growth rate as well as the path is highly influenced by neighboring pores. The coalescing mechanism in porous materials may explain the very brittle nature of sprayed ceramic coatings. We also note that the critical failure loads found in terms of force or temperature probably represent conservative estimates. Inclusion of the cohesive elements increases the structural compliance and may underestimate the failure loads.

Acknowledgments

The authors acknowledge the support of the Center for Thermal Spray Research at Stony Brook, which is funded by NSF-MRSEC DMR9632570. The computations are carried out on HP7000/C180 and C360 workstations using the geometric modeling code I-DEAS and the finite element code ABAQUS. Both codes were made available under academic license from SDRC and Hibbitt, Karlsson and Sorensen, Inc., respectively.

References

[1] Lin, C. K., and Berndt, C. C., 1994, ''Measurement and Analysis of Adhesion Strength for Thermally Sprayed Coatings,'' J. Therm. Spray Tech., **3**, pp. 75–104.
[2] Qian, G., Nakamura, T., Berndt, C. C., and Leigh, S. H., 1997, ''Tensile Fracture Toughness Test and High Temperature Fracture Analysis of Thermal Barrier Coatings,'' Acta Mater., **45**, No. 4, pp. 1767–1784.
[3] Nakamura, T., Qian, G., and Berndt, C. C., 2000, ''Effects of Pores on Mechanical Properties of Plasma Sprayed Ceramic Coatings,'' J. Am. Ceramic Soc., **83**, pp. 578–584.
[4] Leigh, S. H., and Berndt, C. C., 2000, ''Quantitative Evaluation of Void Distribution Within a Plasma Sprayed Ceramic,'' J. Am. Ceram. Soc., **82**, pp. 17–38.
[5] Xu, X.-P., and Needleman, A., 1994, ''Numerical Simulations of Fast Crack Growth in Brittle Solids,'' J. Mech. Phys. Solids, **42**, pp. 1397–1434.
[6] Camacho, G. T., and Ortiz, M., 1996, ''Computational Modeling of Impact Damage in Brittle Materials,'' Int. J. Solids Struct., **33**, pp. 2899–2938.
[7] Guebelle, P. H., and Baylor, J. S., 1998, ''Impact-Induced Delamination of Composites: A 2D Simulation,'' Composites, Part B, **29**, No. 5, pp. 589–602.
[8] Miller, O., Freund, L. B., and Needleman, A., 1999, ''Modeling and Simulation of Dynamic Fragmentation in Brittle Materials,'' Int. J. Fract., **96**, pp. 101–125.
[9] Needleman, A., 1987, ''A Continuum Model for Void Nucleation by Inclusion Debonding,'' ASME J. Appl. Mech., **54**, pp. 525–531.
[10] Tvergaard, V., and Hutchinson, J. W., 1992, ''Relation Between Crack Growth Resistance and Fracture Process Parameters in Elastic-Plastic Solids,'' J. Mech. Phys. Solids, **40**, pp. 1377–1392.
[11] Lin, G., Kim, Y. J., Comec, A., and Schwalbe, K. H., 1997, ''Fracture Tough-

ness of a Constrained Metal Layer,'' Comput. Mater. Sci., **9**, pp. 36–47.
[12] Montavon, G., Sampath, S., Berndt, C. C., Herman, H., and Coddet, C., 1995, ''Effects of Vacuum Plasma Spray Processing Parameters on Splat Morphology,'' J. Therm. Spray Tech., **4**, pp. 67–74.
[13] Bengtsson, P., and Johannesson, T., 1995, ''Characterization of Microstructural Defects in Plasma-Sprayed Thermal Barrier Coatings,'' J. Therm. Spray Tech., **4**, pp. 245–251.
[14] Wawrzynek, P. A., and Ingraffea, A. R., 1989, ''Interactive Approach to Local Remeshing Around a Propagating Crack,'' Finite Elem. Anal. Design, **5**, pp. 87–96.
[15] Belytschko, T., Lu, Y. Y., and Tabbara, M., 1995, ''Element-Free Galerkin Methods for Static and Dynamic Fracture,'' Int. J. Solids Struct., **32**, pp. 2547–2570.
[16] Dagher, H. J., and Kulendran, S., 1992, ''Finite Element Modeling of Corrosion Damage in Concrete,'' ACI Struct. J., **89**, pp. 699–708.
[17] Padovan, J., and Guo, Y. H., 1994, ''Moving Template Analysis of Crack Growth—I. Procedure Development,'' Eng. Fract. Mech., **48**, pp. 405–425.
[18] Padovan, J., and Jae, J., 1997, ''FE Modeling of Expansive Oxide Induced Fracture of Rebar Reinforced Concrete,'' Eng. Fract. Mech., **56**, pp. 797–812.
[19] Nakamura, T., Shih, C. F., and Freund, L. B., 1985, ''Computational Methods Based on an Energy Integral in Dynamic Fracture,'' Int. J. Fract., **27**, pp. 229–243.
[20] Atluri, S. N., and Nishioka, T., 1985, ''Numerical Studies in Dynamic Fracture Mechanics,'' Int. J. Fract., **27**, pp. 245–261.
[21] Gao, H., and Klein, P., 1998, ''Numerical Simulation of Crack Growth in an Isotropic Solids With Randomized Internal Cohesive Bonds,'' J. Mech. Phys. Solids, **46**, pp. 187–218.
[22] Zhang, P., Klein, P., Huang, Y., and Gao, H., 1999, ''Numerical Simulation of Cohesive Fracture by Virtual-Internal-Bond Model,'' Report by the Department of Mechanical and Industrial Engineering, University of Illinois.
[23] Shih, C. F., Moran, B., and Nakamura, T., 1986, ''Energy Release Rate Along a Three-Dimensional Crack Front in a Thermally Stressed Body,'' Int. J. Fract., **30**, pp. 79–102.
[24] Needleman, A., and Rosakis, A. J., 1999, ''The Effect of Bond Strength and Loading Rate on the Conditions Governing the Attainment of Intersonic Crack Growth Along Interfaces,'' J. Mech. Phys. Solids, **47**, pp. 2411–2449.
[25] Leigh, S. H., and Berndt, C. C., 1997, ''Elastic Response of Thermal Spray Deposits Under Indentation Tests,'' J. Am. Ceram. Soc., **80**, pp. 2093–2099.

A. V. Pesterev
Senior Researcher
Institute for Systems Analysis,
Russian Academy of Sciences,
Moscow, Russia

C. A. Tan
Associate Professor
Department of Mechanical Engineering,
Wayne State University,
Detroit, MI 48202
Mem. ASME

L. A. Bergman
Professor
Aeronautical and Astronautical
Engineering Department,
University of Illinois,
Urbana, IL 61801
Fellow ASME

A New Method for Calculating Bending Moment and Shear Force in Moving Load Problems

In this paper, a new series expansion for calculating the bending moment and the shear force in a proportionally damped, one-dimensional distributed parameter system due to moving loads is suggested. The number of moving forces, which may be functions of time and spatial coordinate, and their velocities are arbitrary. The derivation of the series expansion is not limited to moving forces that are a priori known, making this method also applicable to problems in which the moving forces depend on the interactions between the continuous system and the subsystems it carries, e.g., the moving oscillator problem. A main advantage of the proposed method is in the accurate and efficient evaluation of the bending moment and shear force, and in particular, the shear jumps at the locations where the moving forces are applied. Numerical results are presented to demonstrate the rapid convergence of the new series representation.
[DOI: 10.1115/1.1356028]

1 Introduction

The problem of loads traveling along a distributed parameter system is commonly encountered in many important engineering systems. Examples include the design of railroad tracks with high-speed trains and highway bridges with moving vehicles ([1,2]), high-speed precision machining ([3]), circular saw blades ([4]) computer disk drives ([5]), and cables transporting humans/materials ([6]). The accurate prediction of the stresses developed in the continuous system due to moving loads is crucial as a miscalculation may lead to undesirable human casualty and loss of important data and information. The collapse of the U.S. Silver Bridge in 1967 that claimed 46 lives remains a chilling warning to bridge design engineers ([7]).

In this paper, a new method is proposed to calculate the bending moment and shear force of a proportionally damped beam due to moving concentrated loads. The term ''moving concentrated load'' is used to denote either a moving force that is a priori known or one that depends on the interactions between the beam and the moving subsystems it carries. Hereafter, when the moving force is a priori known, the problem is termed the ''moving force problem.'' The solution in the form of a series representation is first derived for arbitrary moving forces and then extended to the moving oscillator problem in which the moving forces depend on the responses of the beam and oscillators.

It was shown that the response and slope of the beam can be accurately determined by using only few terms of a conventional eigenfunction series ([8–10]). However, higher order derivatives of the series (required for calculating the bending moment and shear force) converge poorly and cannot capture the jumps in the shear forces. In this work, the eigenfunction expansion is improved by a ''correction function'' which bears information about the shear force jumps at the locations where the moving loads are applied and includes the contributions of the truncated higher modes in the series. This results in a better and more efficient evaluation of the bending moment and shear force.

The genesis of this technique can be traced to accelerating the convergence of the modal series (spectral) representation for the Green's function (or dynamic flexibility or reacceptance) which has been known for some years as the ''mode-acceleration'' method ([11]). Interested readers are referred to the papers by Dowell [12] Palazzolo et al., [13] (general case of a nonconservative finite-dimensional system), Pesterev and Tavrizov [14] (free-free conservative distributed parameter systems), and the references therein. The mode-acceleration technique has been applied to problems related to the steady-state vibration of structures due to harmonic excitations. However, the extension of this technique to the moving loads problem (which is transient in nature) is not trivial and has not been discussed. As discussed in Palazzolo et al. [13], the improved representation for receptances can be expressed in two equivalent forms: (1) in terms of a series with accelerated convergence or (2) as the sum of a conventional spectral representation and a ''residual flexibility'' which accounts for the truncated higher order modes. The ''correction function'' derived in this paper may thus be viewed as an extension of the notion of residual flexibility for moving load problems.

This paper is organized as follows. In the next section, a mathematical formulation of the problem is given. In Section 3, response solution for damped continua in terms of the conventional series is discussed and the modal representation for the static Green's function is given. The improved series representation for a proportionally damped beam is derived in Section 4. In Section 5, the application of the method to the moving oscillators problem is discussed. The efficiency of the new representation is illustrated by numerical results in Section 6.

2 Problem Statement

The vibration of a spatially one-dimensional, damped distributed parameter system due to moving loads is governed by

$$\rho \frac{\partial^2}{\partial t^2} w(x,t) + D \frac{\partial}{\partial t} w(x,t) + K w(x,t) = \sum_{i=1}^{l} F_i(x,t)\, \delta(x - \zeta_i(t)), \quad x \in [0,L], \tag{1}$$

subject to given boundary and initial conditions. Here, L is the length of the continuum; $w(x,t)$ is the transverse displacement of the continuum; ρ, D, and K are spatial differential operators representing inertia, damping, and stiffness of the system, respectively, ρ and K are positive definite and D is positive semidefinite;

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 10, 2000; final revision, Aug. 18, 2000. Associate Editor: A. K. Mal. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$\delta(x)$ is the Dirac delta-function; and $\zeta_i(t)$ are the time-dependent coordinates at which the forces are applied. The functions $F_i(x,t)$ are assumed to be twice differentiable with respect to both arguments for $0<x<L$ and for t satisfying the inequalities $0<\zeta_i(t)<L$, and are not required to be a priori known. In this work, we restrict our consideration to systems with stiffness operator of order four. The reason for this is explained in Section 4.3.

We consider homogeneous boundary conditions and, in particular, assume that the continuum has no rigid-body modes and its ends are fixed, $w(0,t)=w(L,t)=0$. We further assume zero initial conditions, which implies that the continuum is at rest for $t\leqslant 0$. Note that this assumption merely simplifies the notation and does not affect the idea behind the development of the new series representation. We first consider the case of one force moving with a constant velocity v; i.e., $l=1$ and $\zeta_1(t)=vt$. This requirement, in fact, is not needed for the analysis, and the resulting equations are easily extended to the case of many forces moving with arbitrarily varying velocities (see Remarks 2 and 3 in Section 4.4).

It is well known that the solution to Eq. (1) can be expanded in terms of the eigenfunctions of the distributed system. However, a disadvantage of using this expansion is the poor convergence of the series in calculating the bending moment and shear force because of the moving singularities on the right-hand side of Eq. (1). As a result, these calculations are prohibitedly expensive in terms of the number of terms required. In what follows, a ''correction function'' is derived to accelerate the convergence of the series, which is expressed in terms of the static Green's function of the continuum and its modal parameters. When deriving the improved series representation, we need some results concerning the conventional series expansion, which are summarized in the next section.

3 Conventional Series Expansion

In view of the assumptions stated above, we will look for solution to the equation

$$\rho\frac{\partial^2}{\partial t^2}w(x,t)+D\frac{\partial}{\partial t}w(x,t)+Kw(x,t)=F(x,t)\,\delta(x-vt) \quad (2)$$

subject to given homogeneous boundary and zero initial conditions.

3.1 General Case of Damping. It is well known that the solution to Eq. (2) can be written in terms of the dynamic Green's function $g(x,\eta,t)$ of the distributed system as (see, e.g., [15])

$$w(x,t)=\int_{-\infty}^{t}d\tau\int_0^L g(x,\eta,t-\tau)F(\eta,\tau)\,\delta(\eta-v\tau)\,d\eta \equiv\int_{-\infty}^{t}g(x,v\tau,t-\tau)F(v\tau,\tau)\,d\tau. \quad (3)$$

In practice, the Green's function is represented by the truncated modal series

$$g(x,\eta,t)=\frac{1}{2}\sum_{n=\pm1}^{\pm N}\frac{1}{\lambda_n}e^{\lambda_n t}\varphi_n(x)\varphi_n(\eta) =\sum_{n=1}^{N}\mathrm{Re}\left[\frac{1}{\lambda_n}e^{\lambda_n t}\varphi_n(x)\varphi_n(\eta)\right], \quad (4)$$

where complex λ_n and $\varphi_n(x)$ are, respectively, the nth eigenvalue and eigenfunction of the distributed system. In addition, $\varphi_n(x)$ must satisfy the normalization condition ([15])

$$\int_0^L\varphi_n(x)\rho\varphi_n(x)dx-\frac{1}{\lambda_n^2}\int_0^L\varphi_n(x)K\varphi_n(x)dx=2. \quad (5)$$

Thus, we arrive at the approximation of the response of the system by the series expansion

$$w(x,t)=\sum_{n=1}^{N}\mathrm{Re}[\varphi_n(x)q_n(t)], \quad (6)$$

where the time-dependent coefficients $q_n(t)$ are given by

$$q_n(t)=\frac{1}{\lambda_n}\int_0^t e^{\lambda_n(t-\tau)}\varphi_n(v\tau)F(v\tau,\tau)\,d\tau. \quad (7)$$

3.2 Static Green's Function. By definition, the static Green's function $G(x,\xi)$ is the solution to the equation

$$KG(x,\xi)=\delta(x-\xi) \quad (8)$$

and, for a fixed value of ξ, $0<\xi<L$, satisfies the given boundary conditions. For a string or a beam with arbitrary boundary conditions, the static Green's function $G(x,\xi)$ can easily be obtained either in the form of a polynomial (see, e.g., Appendix II of [16]) for a uniform structure or in terms of quadratures for nonuniform structures.

In what follows, we will also need the modal series representation for the static Green's function. It is given in terms of the eigenvalues $\tilde{\lambda}_n=i\tilde{\omega}_n$ and eigenfunctions $\tilde{\varphi}_n(x)$ of the conservative continuum associated with the damped one under consideration, which are solutions to the eigenvalue problem

$$\{\tilde{\lambda}_n^2\rho+K\}\tilde{\varphi}_n(x)=0, \quad (9)$$

and $\tilde{\varphi}_n(x)$ satisfy the conventional orthonormality relations for conservative systems

$$\int_0^L\tilde{\varphi}_n(x)\rho\tilde{\varphi}_j(x)dx=\delta_{nj,}, \quad (10)$$

where δ_{nj} is the Kronecker delta. Thus, $G(x,\xi)$ can be approximated by the modal series

$$G(x,\xi)=\sum_{n=1}^{N}\frac{\tilde{\varphi}_n(x)\tilde{\varphi}_n(\xi)}{\tilde{\omega}_n^2}\equiv\sum_{n=1}^{N}\frac{\tilde{\varphi}_n(x)\tilde{\varphi}_n(\xi)}{\lambda_n\overline{\lambda}_n}. \quad (11)$$

The derivation of the improved solution relies on the modal series representations for the dynamic and static Green's functions. As can be seen from (4) and (11), these representations are given in terms of different sets of eigenfunctions, which makes the analysis of the general case of damping in the system rather complicated. In this work, we confine our efforts to the case of a proportionally damped continuum, for which the relationship between these sets can be easily found.

3.3 Proportionally Damped Continuum. It is well known ([17,18]) that, if the system is proportionally damped, the system eigenvalues are complex, $\lambda_n=\alpha_n+i\omega_n$, but the eigenfunctions can be taken as real. However, as can be easily seen, no real functions satisfy the normalization condition (5) (since λ_ns are complex), and we need either to use complex eigenfunctions to take advantage of the modal series representation (4) for the dynamic Green's function or to find its equivalent representation in terms of the real eigenfunctions $\tilde{\varphi}_n(x)$. We will look for the eigenfunctions of the damped system in the form

$$\varphi_n(x)=c_n\tilde{\varphi}_n(x), \quad (12)$$

with a complex multiplier c_n being chosen from the condition that $\varphi_n(x)$ satisfies (5). Substituting (12) into (5) and using Eq. (9) and the relation $\tilde{\lambda}_n^2=-\lambda_n\overline{\lambda}_n$, we find that $c_n^2=-i\lambda_n/\omega_n$.

Substituting (12) into (4), we get the modal series representation for the dynamic Green's function in terms of the real eigenfunctions of the corresponding conservative continuum as

$$g(x,\eta,t)=\sum_{n=1}^{N}\mathrm{Re}\left[\frac{1}{i\omega_n}e^{\lambda_n t}\right]\tilde{\varphi}_n(x)\tilde{\varphi}_n(\eta). \quad (13)$$

The solution to Eq. (2) is given then by

$$w(x,t)=\sum_{n=1}^{N}\tilde{\varphi}_n(x)q_n^R(t) \tag{14}$$

where $q_n^R(t)$ is the real part of the integral

$$q_n(t)=\frac{1}{i\omega_n}\int_0^t e^{\lambda_n(t-\tau)}\tilde{\varphi}_n(v\tau)F(v\tau,\tau)d\tau. \tag{15}$$

The slope, bending moment, and shear force are obtained by the term-wise differentiation of the series in (14) with respect to x. For example, the shear force for a uniform beam is given by

$$EIw_{xxx}(x,t)=\sum_{n=1}^{N}EI\tilde{\varphi}_n'''(x)q_n^R(t)\,, \tag{16}$$

where EI is the flexural rigidity of the beam. As mentioned before, because of the jump in the shear force, series (16) converges poorly and an accurate approximation of the shear force requires a large number of terms in the series. In the next section, we derive a new representation which explicitly takes into account this jump.

4 Improved Solution Representation for a Proportionally Damped Beam

4.1 General Idea of the Approach to be Used. As aforementioned, the poor convergence of series (16) is associated with the moving singularity on the right-hand side of Eq. (2). This suggests that one possible way to improve the solution is to try to remove the singularity, i.e., to reduce the problem to that of finding the solution of the original equation with the right-hand side free of the moving singularity. This can be achieved if the desired solution is represented as a sum of two functions such that one of these functions is ''responsible'' for the singularity and can easily be determined. Then, the second function satisfies the original equation with the right-hand side free of the singularity and, thus, can be better approximated by the series in terms of the continuum eigenfunctions. To remove the moving singularity, the concept of quasi-static solution introduced in Pesterev and Bergman ([16]) for the case of a constant moving force is extended to the case of varying moving forces.

4.2 Quasi-Static Solution. The *quasi-static solution* $w_{qs}(x,t)$ is defined as

$$w_{qs}(x,t)=F(vt,t)G(x,vt)[h(t)-h(t-L/v)], \tag{17}$$

where $h(t)$ is the Heaviside unit step function. In view of (8), it is evident that this function satisfies the equation

$$Kw_{qs}(x,t)=F(x,t)\,\delta(x-vt) \tag{18}$$

and gives the response of the distributed system due to the moving force $F(x,t)$ if we neglect the inertia of the system.

4.3 Derivation of the Improved Representation. We will look for the solution to problem (2) in the interval $[0,L/v]$ in the form

$$w(x,t)=\tilde{w}(x,t)+w_{qs}(x,t). \tag{19}$$

Introducing the notation

$$H(x,t)=F(vt,t)G(x,vt), \tag{20}$$

we can write the quasi-static solution for $t<L/v$ *as*

$$w_{qs}(x,t)=H(x,t)h(t). \tag{21}$$

The substitution of (19) into (2) with regard to (18), (20), and (21) results in the equation

$$\begin{aligned}\rho\frac{\partial^2}{\partial t^2}\tilde{w}(x,t)+D\frac{\partial}{\partial t}\tilde{w}(x,t)+K\tilde{w}(x,t)\\ =-\rho(H_{tt}(x,t)h(t)+2H_t(x,t)\,\delta(t)+H(x,t)\,\delta'(t))\\ -D(H_t(x,t)h(t)+H(x,t)\,\delta(t)),\end{aligned} \tag{22}$$

where $H_t(x,t)$ and $H_{tt}(x,t)$ are the first and second derivatives, respectively, of $H(x,t)$ with respect to time. If the order of the highest derivative in the stiffness operator is four, then the right-hand side of Eq. (22) has no moving singularity, and hence, the function $\tilde{w}(x,t)$ can be better approximated by the series in terms of the eigenfunctions of the continuum compared to $w(x,t)$. The condition imposed on the stiffness operator is essential. Indeed, the function $H_{tt}(x,t)$ contains the second derivative of the static Green's function with respect to the second variable, $G_{\xi\xi}(x,vt)$. If the differential order of the stiffness operator K is two, then it follows from Eq. (8) and the symmetry of $G(x,\xi)$ that the right-hand side of (22) contains a moving singularity, the function $\delta(x-vt)$. This implies that the method to be presented cannot be directly applied (at least, in the form described below) to systems that have differential stiffness operator of order two, e.g., to strings or rods.

We will expand the solution to (22) in the series of N eigenfunctions of the distributed system and write it in the form

$$\begin{aligned}\tilde{w}(x,t)=-\int_{-\infty}^{t}d\tau\int_0^L g(x,\eta,t-\tau)\rho(H_{tt}(\eta,\tau)h(\tau)\\ +2H_t(\eta,\tau)\,\delta(\tau)+H(\eta,\tau)\,\delta'(\tau))d\eta\\ -\int_{-\infty}^{t}d\tau\int_0^L g(x,\eta,t-\tau)D(H_t(\eta,\tau)h(\tau)\\ +H(\eta,\tau)\,\delta(\tau))d\eta,\end{aligned} \tag{23}$$

where $g(x,\eta,t)$ is given by (13). By using the integration by parts, the right-hand side of Eq. (23) can be transformed to a form free of delta-functions and derivatives of $H(x,t)$,

$$\begin{aligned}\tilde{w}(x,t)=-\int_0^L g_t(x,\eta,0)\rho H(\eta,t)d\eta-\int_0^t d\tau\int_0^L g_{tt}(x,\eta,t-\tau)\\ \times\rho H(\eta,\tau)d\eta-\int_0^t d\tau\int_0^L g_t(x,\eta,t-\tau)DH(\eta,\tau)d\eta.\end{aligned} \tag{24}$$

The proof of this is given in the Appendix.

Now, we apply modal series representations (11) and (13), to evaluate the integrals over η on the right-hand side of (24). Using orthogonality relations (10), we find that

$$\int_0^L g_t(x,\eta,0)\rho H(\eta,t)d\eta=F(vt,t)\sum_{n=1}^{N}\frac{\tilde{\varphi}_n(x)\tilde{\varphi}_n(vt)}{\lambda_n\bar{\lambda}_n}. \tag{25}$$

Similarly,

$$\begin{aligned}\int_0^L g_{tt}(x,\eta,t-\tau)\rho H(\eta,\tau)d\eta\\ =F(v\tau,\tau)\sum_{n=1}^{N}\ \mathrm{Re}\left[\frac{\lambda_n}{i\omega_n\bar{\lambda}_n}e^{\lambda_n(t-\tau)}\right]\tilde{\varphi}_n(x)\tilde{\varphi}_n(v\tau).\end{aligned} \tag{26}$$

Using the equation $\int_0^L\tilde{\varphi}_n(x)D\tilde{\varphi}_j(x)dx=-2\alpha_n\delta_{jn}$, we find that

$$\int_0^L g_t(x,\eta,t-\tau)DH(\eta,\tau)d\eta$$

$$=-F(v\tau,\tau)\sum_{n=1}^{N}\ \mathrm{Re}\left[\frac{2\alpha_n}{i\omega_n\overline{\lambda_n}}e^{\lambda_n(t-\tau)}\right]\tilde{\varphi}_n(x)\tilde{\varphi}_n(v\tau). \quad (27)$$

Adding integrals (26) and (27), we get

$$\int_0^L g_{tt}(x,\eta,t-\tau)\rho H(\eta,\tau)d\eta+\int_0^L g_t(x,\eta,t-\tau)DH(\eta,\tau)d\eta$$

$$=F(v\tau,\tau)\sum_{n=1}^{N}\ \mathrm{Re}\left[\frac{1}{i\omega_n\overline{\lambda_n}}(\lambda_n-2\alpha_n)e^{\lambda_n(t-\tau)}\right]\tilde{\varphi}_n(x)\tilde{\varphi}_n(v\tau)$$

$$=-g(x,v\tau,t-\tau)F(v\tau,\tau).$$

It follows from the last equation and Eqs. (24) and (25) that

$$\tilde{w}(x,t)=\int_0^t g(x,v\tau,t-\tau)F(v\tau,\tau)d\tau-F(vt,t)\sum_{n=1}^{N}\frac{\tilde{\varphi}_n(x)\tilde{\varphi}_n(vt)}{\lambda_n\overline{\lambda_n}}$$

$$=\sum_{n=1}^{N}\tilde{\varphi}_n(x)q_n^R(t)-F(vt,t)\sum_{n=1}^{N}\frac{\tilde{\varphi}_n(x)\tilde{\varphi}_n(vt)}{\tilde{\omega}_n^2},$$

where $q_n^R(t)$ is the real part of integral (15). As can be seen, the first term represents the conventional series expansion. Using (19), we arrive at a compact formula for the desired solution

$$w(x,t)=\sum_{n=1}^{N}\tilde{\varphi}_n(x)q_n^R(t)$$

$$+F(vt,t)\left(G(x,vt)-\sum_{n=1}^{N}\frac{\tilde{\varphi}_n(x)\tilde{\varphi}_n(vt)}{\tilde{\omega}_n^2}\right). \quad (28)$$

4.4 Discussions and Extensions of the Improved Representation. As can be seen from Eq. (28), the improved solution involves no additional computations compared to the conventional series expansion (14). The function in the parenthesis, which may be termed *correction function*, or *dynamic flexibility*, is easily calculated given that the static Green's function is known. This function bears information about the truncated higher modes.

Remark 1. In the above analysis, we considered the time interval $[0,L/v]$, when the force is on the continuum. To extend it to the values of time greater than L/v (when the force leaves the continuum), we need to take into account both unit step functions in the definition of the quasi-static solution (17), which results in the additional term

$$\rho(H_{tt}(x,t)h(t-L/v)+2H_t(x,t)\delta(t-L/v)+H(x,t)\delta'(t-L/v))$$
$$+D(H_t(x,t)h(t-L/v)+H(x,t)\delta(t-L/v))$$

on the right-hand side of (22). Repeating the above calculations for this case and using additionally the assumption that the right end of the continuum is fixed, we obtain, as could be expected, the solution in the form of the conventional series

$$w(x,t)=\sum_{n=1}^{N}\tilde{\varphi}_n(x)q_n^R(t),\quad t>L/v,$$

where $q_n^R(t)$ is again given by the real part of (15) if we set $F(vt,t)=0$ for $t>L/v$. Equations (15) and (28) can be made valid for all values of t by the use of extended eigenfunctions introduced in Pesterev, et al. [19] i.e., for $x<0$ and $x>L$, $\tilde{\varphi}_n(x)\equiv 0$ and $G(x,\xi)\equiv 0$.

Remark 2. Note that none of the derivations employ the assumption of constant velocity of the moving force. It can be easily checked, that all calculations remain valid if the velocity varies. In that case, we simply need to substitute the function $\zeta(t)$ for vt in (28) (and to append the equation governing the variation of $\zeta(t)$ if this function is not specified explicitly).

Remark 3. The case of many forces traversing the beam can be treated in the same way as the case of one force. In view of Remarks 1 and 2, solution to Eq. (1), for any $t>0$, can be written in the form

$$w(x,t)=\sum_{n=1}^{N}\tilde{\varphi}_n(x)q_n^R(t)+\sum_{i=1}^{l}F_i(\zeta_i(t),t)$$
$$\times\left(G(x,\zeta_i(t))-\sum_{n=1}^{N}\frac{\tilde{\varphi}_n(x)\tilde{\varphi}_n(\zeta_i(t))}{\tilde{\omega}_n^2}\right), \quad (29)$$

where Eq. (15) for the time-dependent coefficients $q_n(t)$ now takes the form

$$q_n(t)=\frac{1}{i\omega_n}\int_0^t e^{\lambda_n(t-\tau)}\sum_{j=1}^{l}\tilde{\varphi}_n(\zeta_j(\tau))F_j(\zeta_j(\tau),\tau)d\tau.$$

Note that the use of the extended eigenfunctions ([19]) in the last equation takes care of how many forces are on the beam at a current time t such that the fact that a certain pth force has already left the beam or has not come yet is automatically taken into account since the functions $\varphi_j(\zeta_p(\tau))$, $j=1,\ldots,N$, vanish in these cases.

For a uniform beam, differentiating both sides of (29) gives the improved representation for the bending moment

$$EIw_{xx}(x,t)=EI\sum_{n=1}^{N}\tilde{\varphi}_n''(x)q_n^R(t)+\sum_{i=1}^{l}F_i(\zeta_i(t),t)EI$$
$$\times\left(G_{xx}(x,\zeta_i(t))-\sum_{n=1}^{N}\frac{\tilde{\varphi}_n''(x)\tilde{\varphi}_n(\zeta_i(t))}{\tilde{\omega}_n^2}\right), \quad (30)$$

and the shear force

$$EIw_{xxx}(x,t)=EI\sum_{n=1}^{N}\tilde{\varphi}_n'''(x)q_n^R(t)+\sum_{i=1}^{l}F_i(\zeta_i(t),t)EI$$
$$\times\left(G_{xxx}(x,\zeta_i(t))-\sum_{n=1}^{N}\frac{\tilde{\varphi}_n'''(x)\tilde{\varphi}_n(\zeta_i(t))}{\tilde{\omega}_n^2}\right). \quad (31)$$

The jumps in the shear force at the points $x_i(t)=\zeta_i(t)$ are calculated exactly by virtue of the static Green's function and equal to $F_i(\zeta_i(t),t)$, since $EI(G_{xxx}(\zeta_i^+(t),\zeta_i(t))-G_{xxx}(\zeta_i^-(t),\zeta_i(t)))=1$.

5 Application to the Moving Oscillator Problem

The general formulas obtained in the previous section are valid independent of the fact whether the functions $F_i(x,t)$ are a priori known or not (we did not use the explicit dependence of these functions on time or spatial coordinate). If the functions $F_i(x,t)$ are a priori known, then the improved solution is obtained as easily as in the case of the constant moving force ([16]). The situation becomes more difficult if we deal with the moving oscillator problem. In this case, $F_i(x,t)$ depend on the response of the continuum and on other unknowns such as vertical displacements of the oscillators, the equations for which are to be appended to (1). For example, for the problem where several conservative oscillators traverse the continuum, we have $F_i(x,t)=-m_ig-k_i(w(x,t)-z_i(t))$, where m_i and k_i are the mass and the spring stiffness of the ith oscillator and $z_i(t)$ are the unknown vertical displacements of the oscillators, which require additional equations. For the problem of damped oscillators moving with a constant velocity v along an even beam surface with the profile $\varepsilon(x)$, $F_i(x,t)$ are given by

$$F_i(x,t) = -m_i g - k_i(w(x,t) - z_i(t)) - c_i \frac{d}{dt}(w(x,t) - z_i(t)) - k_i \varepsilon(x) - c_i v \varepsilon'(x),$$

where c_i are the damper coefficients. Thus, we see that, in order to calculate the forces $F_i(\zeta_i(t),t)$ acting on the beam at the points of the oscillator attachments, we need to know the displacements of the beam $w(\zeta_i(t),t)$ at these points, which, as can be seen from (29), depend in their turn on the forces $F_i(\zeta_i(t),t)$. Instead of trying to find an accurate solution to this problem, we suggest the following approach.

The response and slope can be accurately determined by using the conventional series expansion (14). The high accuracy of calculation is explained by the fast convergence of the series in the case of a beam ($\tilde{\omega}_n$s are proportional to n^2) and is confirmed by our previous results ([8–10]). Since the interaction forces depend on the beam response and, in the case of a damped oscillator, on the slope of the beam, we suggest to first determine the forces by using the conventional series ([19]). Then, substitute the interaction forces obtained for $F_i(\zeta_i(t),t)$ into the improved representations (30) or (31) to accurately calculate the bending moment or shear force. The program implementation of this approach is extremely easy and suggests the use of the programs implementing the earlier methods with the subsequent correction of the solutions obtained. The results of our numerical experiments shown in the next section demonstrate that the new series converges rapidly, which substantiates the efficiency of the new representation and justifies the use of the approach suggested above in the moving oscillator problem. Since the interaction forces are calculated approximately, there may appear a question of whether the improved series converges to the solution of Eq. (1). This question is easily answered. Indeed, let the number N of the series go to infinity. Then, the function in the parentheses in (28) tends to zero (the infinite series in the parenthesis equals the static Green's function), and the improved representation (28) reduces to the conventional series, which is known to converge to the desired solution.

6 Numerical Examples

The aim of our numerical experiments was to examine the convergence of the improved series representation (31) for the shear force distribution and to provide comparison with the solution obtained through the use of the conventional series expansion. The latter was calculated by the method described in Pesterev and Bergman [10]. We refer to this solution as ''conventional solution.'' The solution obtained through the use of (31) is referred to as the ''new solution.'' The static Green's function of a simply supported beam required in (31) was evaluated by means of the analytical formula given in Pesterev and Bergman [16].

We considered five damped oscillators with equal masses m traversing a simply supported damped beam with the velocity $v = 6$ m/s and the arrival time intervals 0.2 s. The beam parameters are the same as those employed in Sadiku and Leipholz [20] and in Pesterev and Bergman [8–10,16]: $L = 6$ m, $EI/\rho = 275.4408\ \mathrm{m}^4/\mathrm{s}^2$, $m/\rho L = 0.2$. We introduced moderate proportional damping into the beam model by setting $D/\rho = 2.0\ \mathrm{s}^{-1}$ (the critical value of damping for this beam is equal to $9.1\ \mathrm{s}^{-1}$) and dampers in all oscillator models with the damping coefficients $c_i = 2\ \mathrm{N\cdot s/m}$, such that the fourth oscillator is overdamped and the others are underdamped. The spring stiffness coefficients are $k_1 = 20$, $k_2 = 30$, $k_3 = 40$, $k_4 = 4$, and $k_5 = 20$ (N/m). The results related to this system are shown in Figs. 1–3. The forces acting on the beam from the oscillators (each force is the sum of the oscillator weight and the elastic and damping forces) in the time interval 0 and 1.8 s (when at least one oscillator is on the beam) were calculated with the use of the conventional series expansion. Figure 1 shows the exact values of the forces (solid lines) and their approximations by two terms of the series (dashed lines). The convergence of the conventional series for the response is so good that, beginning with $N = 4$, all approximations result in the same curves and may be considered as accurate. These forces and the time-dependent coefficients $q_n(t)$ of the conventional series expansion were substituted into Eq. (31) to calculate the shear force distribution at $t = 0.9$ s by the proposed method. Figure 2 demonstrates the convergence of the new series expansion: the solid and

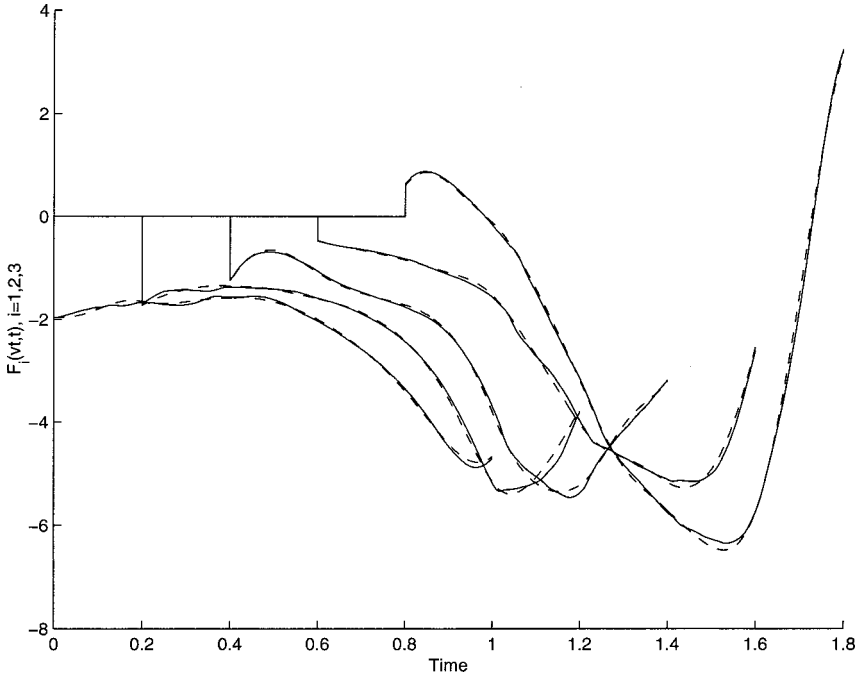


Fig. 1 Forces acting on the beam from the oscillators

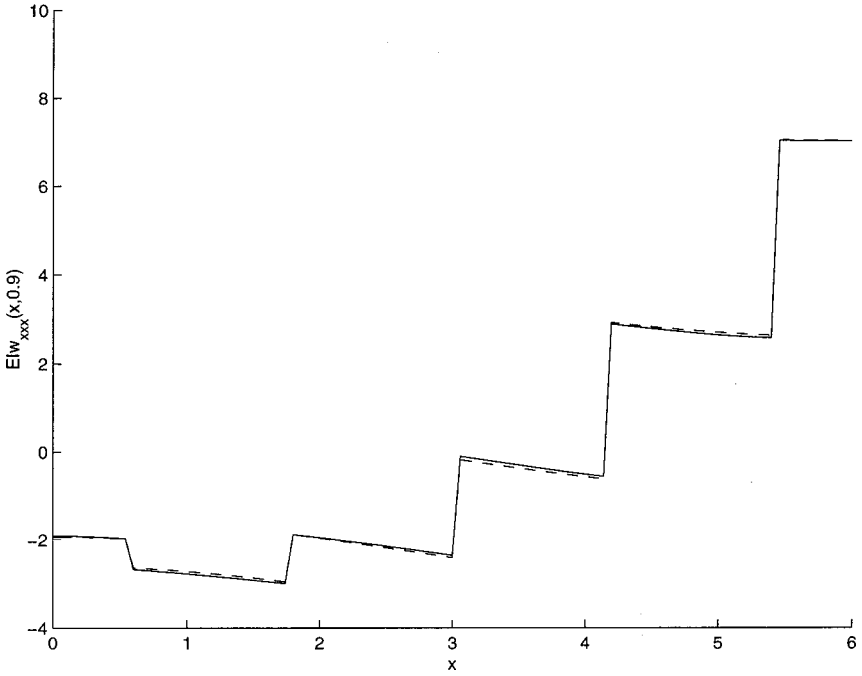


Fig. 2 Shear force distribution at $t=0.9$ s: exact solution (solid line) and solution by two terms of the new series (dashed line)

dashed lines depict the accurate solution and its approximation by two terms of (31), respectively. Beginning with $N=4$, the curves corresponding to different approximations with N terms of series (31) coincide. Figure 3 demonstrates the convergence of the conventional series and shows the accurate solution (solid line) and the approximations obtained by using 10 (dash-dotted line) and 20 (dashed line) terms. The difference in the convergence of two series is easily seen and self-explanatory.

The results presented show the superiority of the new representation (31) over the conventional one: two–four terms of the new expansion were sufficient to get nearly exact solution. On the other hand, the conventional series is not able to provide a good approximation for the shear force: even with 20 terms, the solution obtained is still far from the accurate one. Although not shown, results using the new series also converge faster than those by the conventional series in the calculation of the bending moment. For instance, the bending moment distribution in the neighborhood of the location of the moving force is poorly represented by the conventional series but is accurately calculated by the new series.

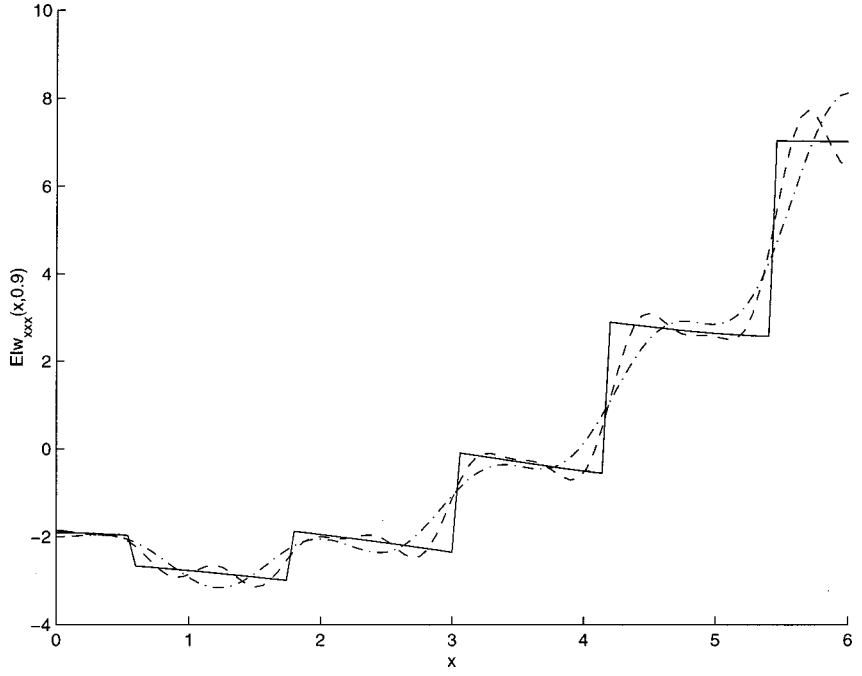


Fig. 3 Shear force distribution at $t=0.9$ s: exact solution (solid line) and solution by the conventional series with 10 terms (dash-dot line) and 20 terms (dashed line)

7 Conclusions

An improved series expansion of the solution to the problem of vibration of a proportionally damped beam subject to an arbitrary number of moving loads has been derived. The forces acting on the beam may depend on time and spatial coordinate and are allowed to move with different and arbitrarily varying velocities. The improved representation is valid even if the moving forces are not a priori known, which made it possible to apply it to the problem of multiple moving oscillators. The convergence rate of the new expansion is considerably better than that of the conventional series expansion.

The advantages of the new technique are most pronounced when the term-wise differentiation of the response solution is required to calculate the shear force distribution, which is a discontinuous function. The jumps in the shear force at the points where the forces are applied are explicitly and accurately taken into account by the quasi-static solution.

Numerical results have been presented that clearly demonstrate the improved convergence of the new representation. Based on these and other results, not included in the paper, we can state that, even with 25 terms, the approximation by the conventional series is worse than the three-term approximation by the new method. Note that the number of first-order ordinary differential equations required to solve the multiple moving oscillator problem is equal to $2(N+l)$, where l is the number of the oscillators. Thus, the difference in the computational complexity of the methods based on the improved and conventional series is considerable.

Acknowledgments

The authors wish to acknowledge the support of the National Science Foundation through grant number CMS-9800136 (Dr. P. P. Nelson, Division Director, Civil and Mechanical Systems). The first author especially acknowledges the International Programs Division of the National Science Foundation for providing the grant supplement necessary to facilitate his extended visit to the United States and this collaboration.

Appendix

Proof of Representation (24). Using the well-known properties of the functions $h(t)$, $\delta(t)$, and $\delta'(t)$, we get

$$\begin{aligned}\tilde{w}(x,t) = &-\int_{-\infty}^{t} d\tau \int_0^L g(x,\eta,t-\tau)\rho(H_{tt}(\eta,\tau)h(\tau)\\ &+2H_t(\eta,\tau)\delta(\tau)+H(\eta,\tau)\delta'(\tau))d\eta\\ &-\int_{-\infty}^{t} d\tau \int_0^L g(x,\eta,t-\tau)D(H_t(\eta,\tau)h(\tau)\\ &+H(\eta,\tau)\delta(\tau))d\eta\\ =&-\int_0^t d\tau \int_0^L g(x,\eta,t-\tau)\rho H_{tt}(\eta,\tau)d\eta\\ &-2\int_0^L g(x,\eta,t)\rho H_t(\eta,0)d\eta\\ &+\int_0^L \frac{\partial}{\partial\tau}(g(x,\eta,t-\tau)\rho H(\eta,\tau))\big|_{\tau=0}d\eta\\ &-\int_0^t d\tau \int_0^L g(x,\eta,t-\tau)DH_t(\eta,\tau)d\eta\\ &-\int_0^L g(x,\eta,t)DH(\eta,0)d\eta. \end{aligned} \tag{32}$$

Rewrite the third integral in the right-hand side of (32) as

$$\begin{aligned}&\int_0^L \frac{\partial}{\partial\tau}(g(x,\eta,t-\tau)\rho H(\eta,\tau))\big|_{\tau=0}d\eta\\ &\quad=\int_0^L g_t(x,\eta,t)\rho H(\eta,0)d\eta+\int_0^L g(x,\eta,t)\rho H_t(\eta,0)d\eta.\end{aligned} \tag{33}$$

Since the left end of the continuum is assumed to be fixed, $G(\eta,0)=G(0,\eta)=0$, and, hence, $H(\eta,0)=0$. Thus, the first integral on the right-hand side of (33) vanishes, and (32) reduces to

$$\begin{aligned}\tilde{w}(x,t) = &-\int_0^t d\tau \int_0^L g(x,\eta,t-\tau)\rho H_{tt}(\eta,\tau)d\eta\\ &-\int_0^t d\tau \int_0^L g(x,\eta,t-\tau)DH_t(\eta,\tau)d\eta\\ &-\int_0^L g(x,\eta,t)\rho H_t(\eta,0)d\eta-\int_0^L g(x,\eta,t)DH(\eta,0)d\eta.\end{aligned} \tag{34}$$

Changing the order of the integration in the first two integrals and taking the internal integrals by parts, we obtain

$$\begin{aligned}&\int_0^t g(x,\eta,t-\tau)\rho H_{tt}(\eta,\tau)d\tau\\ &\quad=g(x,\eta,0)\rho H_t(\eta,t)-g(x,\eta,t)\rho H_t(\eta,0)\\ &\qquad+g_t(x,\eta,0)\rho H(\eta,t)-g_t(x,\eta,t)\rho H(\eta,0)\\ &\qquad+\int_0^t g_{tt}(x,\eta,t-\tau)\rho H(\eta,\tau)d\tau,\end{aligned} \tag{35}$$

and

$$\begin{aligned}&\int_0^t g(x,\eta,t-\tau)DH_t(\eta,\tau)d\tau\\ &\quad=g(x,\eta,0)DH(\eta,t)-g(x,\eta,t)DH(\eta,0)\\ &\qquad+\int_0^t g_t(x,\eta,t-\tau)DH(\eta,\tau)d\tau.\end{aligned} \tag{36}$$

The fourth addend on the right-hand side of Eq. (35) vanishes since $H(\eta,0)=0$. The first addends on the right-hand sides of (35) and (36) are equal to zero since $g(x,\eta,0)$ is zero. Substituting the resulting equations into (34), we get (24).

References

[1] Diana, G., and Cheli, F., 1989, ''Dynamic Interaction of Railway Systems With Large Bridges,'' Veh. Syst. Dyn., **18**, pp. 71–106.
[2] Taheri, M. R., Ting, E. C., and Kukreti, A. R., 1990, ''Vehicle-Guideway Interactions: A Literature Review,'' Shock Vib. Dig., **22**, pp. 3–9.
[3] Chen, C. H., and Wang, K. W., 1994, ''An Integrated Approach Toward the Modeling and Dynamic Analysis of High-Speed Spindles, Part II: Dynamics Under Moving End Load,'' ASME J. Vibr. Acoust., **116**, pp. 514–522.
[4] Mote, C. D., Jr., 1970, ''Stability of Circular Plates Subjected to Moving Loads,'' J. Franklin Inst., **290**, No. 4, pp. 329–344.
[5] Weisenel, G. N., and Schlack, A. L., 1993, ''Response of Annular Plates to Circumferentially and Radially Moving Loads,'' ASME J. Appl. Mech., **60**, pp. 649–661.
[6] Zhu, W. D., and Mote, C. D., Jr., 1994, ''Free and Forced Response of an Axially Moving String Transporting a Damped Linear Oscillator,'' J. Sound Vib., **177**, No. 5, pp. 591–610.
[7] C-J Online News, 1997, http://www.cjonline.com/stories/110397/new_usbridges.html.
[8] Pesterev, A. V., and Bergman, L. A., 1997, ''Vibration of Elastic Continuum Carrying Moving Linear Oscillator,'' J. Eng. Mech., **123**, pp. 878–884.
[9] Pesterev, A. V., and Bergman, L. A., 1997, ''Vibration of Elastic Continuum Carrying Accelerating Oscillator,'' J. Eng. Mech., **123**, pp. 886–889.
[10] Pesterev, A. V., and Bergman, L. A., 1998, ''Response of a Nonconservative

Continuous System to a Moving Concentrated Load,'' ASME J. Appl. Mech., **65**, pp. 436–444.
[11] Bisplinghoff, R. L., Ashley, H., and Halfman, R. L., 1955, *Aeroelasticity*, Addison-Wesley, Boston (reprinted 1996, Dover, New York).
[12] Dowell, E. H., 1996, ''Comment on Energy Flow Predictions in a Structure of Rigidly Joined Beams Using Receptance Theory,'' J. Sound Vib., **194**, pp. 445–447.
[13] Palazzolo, A. B., Wang, B. P., and Pilkey, W. D., 1982, ''A Receptance Formula for General Second-Degree Square Lambda Matrices,'' Int. J. Numer. Methods Eng., **18**, pp. 829–843.
[14] Pesterev, A. V., and Tavrizov, G. A., 1994, ''Vibrations of Beams With Oscillators, I: Structural Analysis Method for Solving the Spectral Problem,'' J. Sound Vib., **170**, pp. 521–536.
[15] Yang, B., 1996, ''Closed-Form Transient Response of Distributed Damped Systems, Part I: Modal Analysis and Green's Function Formula,'' ASME J. Appl. Mech., **63**, pp. 997–1003.
[16] Pesterev, A. V., and Bergman, L. A., 2000, ''An Improved Series Expansion of the Solution to the Moving Oscillator Problem,'' ASME J. Vibr. Acoust., **122**, pp. 54–61.
[17] Caughey, T. K., and O'Kelly, M. E. J., 1965, ''Classical Normal Modes in Damped Linear Dynamic Systems,'' ASME J. Appl. Mech., **32**, pp. 583–588.
[18] Yang, B., 1996, ''Integral Formulas for Non-Self-Adjoint Distributed Dynamic Systems,'' AIAA J., **34**, pp. 2132–2139.
[19] Pesterev, A. V., Yang, B., Bergman, L. A., and Tan, C. A., 2001, ''Response of an Elastic Continuum Carrying Multiple Moving Oscillators,'' J. Eng. Mech., in press.
[20] Sadiku, S., and Leipholz, H. H. E., 1987, ''On the Dynamics of Elastic Systems With Moving Concentrated Masses,'' Ing.-Arch., **57**, pp. 223–242.

Thin-Walled Multicell Beam Analysis for Coupled Torsion, Distortion, and Warping Deformations

J. H. Kim
Research Associate,
e-mail: nockim@idealab.snu.ac.kr

Y. Y. Kim[1]
Associate Professor,
e-mail: yykim@snu.ac.kr

School of Mechanical and Aerospace Engineering and Institute of Advanced Machinery and Design,
Seoul National University,
Kwanak-Gu, Shinlim-Dong, San 56-1,
Seoul 151-742, Korea

Due to the complicated deformations occurring in thin-walled multicell beams, no satisfactory one-dimensional beam theory useful for general quadrilateral multicells appears available. In this paper, we present a new systematic approach to analyze the coupled deformations of torsion, distortion, and the related warping. To develop a one-dimensional thin-walled multicell beam theory, the method to determine the section deformation functions associated with distortion and distortional warping is newly developed. In order to guarantee the singlevaluedness of the distortional warping function in multicells, distortional shear flows have been utilized. The superior result by the present one-dimensional theory is demonstrated with various examples.
[DOI: 10.1115/1.1357166]

1 Introduction

Thin-walled closed beams are very useful structural members in civil, automobile, and aerospace industries due to their high rigidity and lightweightness. Though the analysis of open section beams is quite well known ([1,2]), the analysis of closed section beams is still of a major research topic because of their complicated deformations such as distortion. The analysis related to distortion and distortional warping is, perhaps, the key issue in the analysis of thin-walled closed beams.

The fundamental treatment on this subject goes back to Vlasov [2], Křístek [3], and Wright [4]. Boswell and Zhang [5] discussed sectional deformation functions for distortion and distortional warping for monosymmetric cross sections. Balch and Steele [6] analyzed the local effects associated with warping and distortion near the T-joint of thin-walled closed beams. Hsu et al. [7] used an equivalent beam-on-elastic-foundation method to account for the distortion deformation effect of the cross section with rigid or flexible diaphragms. Recently, the present authors ([8]) develop a new approach to determine section deformation functions for distortion and distortional warping of quadrilateral cross sections. A beam-frame model approach for general cross sections is also formulated ([9]) to find distortional functions for generally shaped cross sections. The importance of the consideration of the shear flow is discussed for the singlevaluedness of the distortional warping function ([9,10]). Jönsson [10] used local axial equilibrium to derive differential equations for the determination of the distortional warping and shear distributions. Razaqpur and Li [11,12] performed a rigorous study on multicell box beams, but their methods appear difficult to extend for the analysis of arbitrarily shaped multicell beams.

The aim of this work is to carry out the one-dimensional analysis of coupled deformations of torsion, distortion, and warping in general multicell box beams. To this end, we propose a systematic method to determine the section function for distortional deformation of multicell beams by extending the method used for single-cell beams ([8]). Then a procedure to determine the section deformation function for distortional warping is presented. In deriving the distortional warping function, the importance of a distortional shear flow is addressed. The numerical analysis is carried out by developing C^0-finite elements based on the present approaches.

2 Field Approximation in Single-Cell Box Beams

We review underlying displacement approximations for the analysis of thin-walled simply connected closed beams, which have been proposed by Kim and Kim [8,9]. The displacement field for multiply connected or multicell sections, which will be the major subject of the present work, is presented in the next section.

The shell displacements of a point on the cross section contour are expressed in terms of the normal $u_n(s,z)$, tangential $u_s(s,z)$ and axial $u_z(s,z)$ components as in Fig. 1. In Fig. 1, the distance ON_i from the shear center O to the point N_i of the ith edge is denoted by r_i. The normal passing through the shear center defines the point N_i. The length of each wall is denoted by b_i, and the distance from the ith corner to the point N_i by l_i.

The shell displacement can be expressed as the sum of beam deformations and corresponding sectional deformation shape functions. The beam deformations in consideration include rotation $\theta(z)$, torsional warping $U^\theta(z)$, distortion $\chi(z)$, and distortional warping $U^\chi(z)$. Figure 2 shows the beam deformations of a rectangular cross section.

Once the corresponding section deformation shapes $\psi(s)$ are determined, one can express the shell displacement as ([8,9])

$$u_s(s,z)=\psi_s^\theta(s)\,\theta(z)+\psi_s^\chi(s)\chi(z) \tag{1a}$$

$$u_n(s,z)=\psi_n^\theta(s)\,\theta(z)+\psi_n^\chi(s)\chi(z) \tag{1b}$$

$$u_z(s,z)=\psi_z^{U^\theta}(s)U^\theta(z)+\psi_z^{U^\chi}(s)U^\chi(z). \tag{1c}$$

In Eqs. (1), only nonvanishing section deformations $\psi(s)$ are considered. The section shapes associated with the rotation $\theta(z)$ are

$$\psi_s^\theta(s)=r_i, \quad \psi_n^\theta(s)=-l_i+s. \tag{2}$$

For simply connected or single-cell sections, the sectional deformation shape $\psi_{z\,(s)}^{U^\theta}$ associated with the torsional warping is

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, February 8, 2000; final revision, October 4, 2000. Associate Editor: R. C. Benson. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

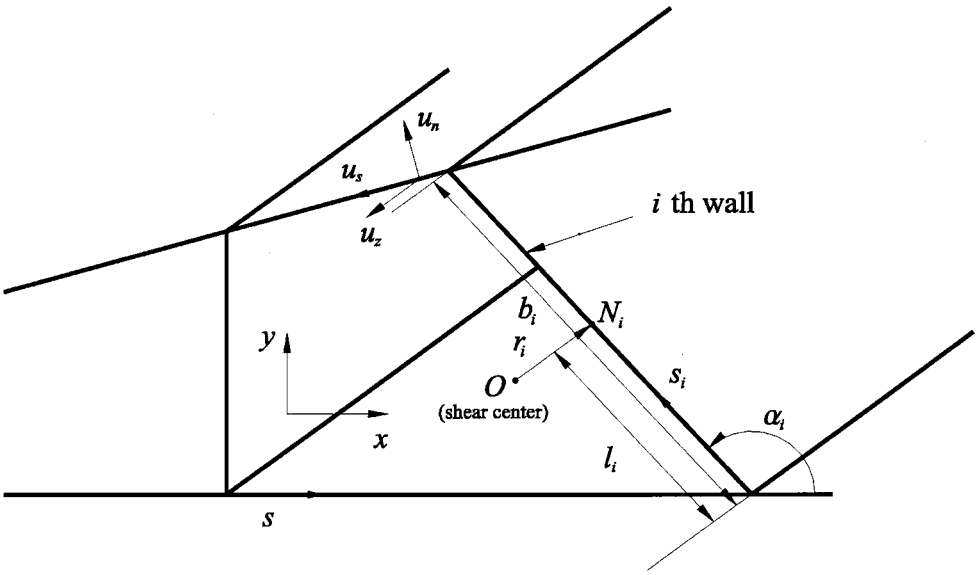
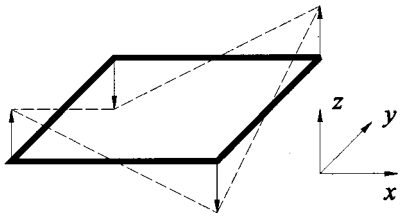
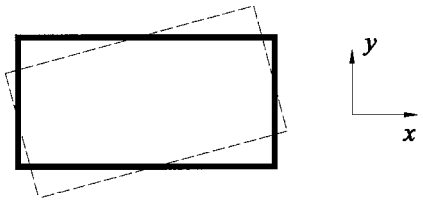


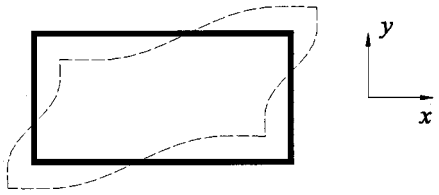
Fig. 1 Displacements at an arbitrary point of a thin-walled beam



(a) warping



(b) rotation



(c) distortion

Fig. 2 Section deformations of a rectangular box beam for (*a*) warping, (*b*) rotation and (*c*) distortion. Dotted lines denote the deformed shapes. In this case, no distinction between torsional and distortional warping deformations is needed as they are identical.

$$\psi_z^{U^\theta}(s) = -\int_0^s (r_i - r_n)ds + \psi_{z0}^{U^\theta} \quad \text{for the } i\text{th wall.} \tag{3}$$

The constant $\psi_{z0}^{U^\theta}$ in Eq. (3) is selected to satisfy the singlevaluedness condition

$$\oint \psi_z^{U^\theta}(s)ds = 0$$

and r_n is defined as

$$r_n = \frac{2\overline{A}}{\oint_c ds}$$

where the area enclosed by and the total length of the contour C are denoted by $\overline{A}$ and $\oint_C ds$, respectively. The sectional deformation shapes $\psi_s^\chi(s)$ and $\psi_z^{U^\chi}(s)$ by Kim and Kim [8,9] are somewhat complicated even for simply connected sections, and thus they are not repeated here. However, the next section will discuss these shapes for multiply connected sections in detail.

On the assumption that the thickness of the beam is much smaller than other sectional dimensions, the three-dimensional displacements $\tilde{u}_n$, $\tilde{u}_s$, $\tilde{u}_z$ may be written as

$$\tilde{u}_n(n,s,z) \approx u_n(s,z) = \psi_n^\theta(s)\,\theta(z) + \psi_n^\chi(s)\chi(z)$$

$$\tilde{u}_s(n,s,z) \approx u_s(s,z) + n\frac{\partial u_n(s,z)}{\partial s}$$

$$\approx \psi_s^\theta(s)\chi(z) + \psi_s^\chi(s)\chi(z) + n\frac{d\psi_n^\chi(s)}{ds}\chi(z)$$

$$\tilde{u}_z(n,s,z) \approx u_z(s,z) = \psi_z^{U^\theta}(s)U^\theta(z) + \psi_z^{U^\chi}(s)U^\chi(z). \tag{4}$$

The normal coordinate n in Eq. (4) is measured from the middle line of each wall. Nonegligible three-dimensional strain components obtained from the displacement field in Eq. (4) can be found as

$$\epsilon_{zz}=\frac{\partial \tilde{u}_z}{\partial z}\approx \psi_z^{U^\theta}(s)\frac{dU^\theta(z)}{dz}+\psi_z^{U^\chi}(s)\frac{dU^\chi(z)}{dz} \tag{5a}$$

$$\epsilon_{zs}=\frac{1}{2}\left(\frac{\partial \tilde{u}_z}{\partial s}+\frac{\partial \tilde{u}_s}{\partial z}\right)\approx \frac{1}{2}\left[\frac{d\psi_z^{U^\theta}(s)}{ds}U^\theta(z)+\frac{d\psi_z^{U^\chi}(s)}{ds}U^\chi(z)+\psi_s^\theta(s)\frac{d\theta(z)}{dz}+\psi_s^\chi(s)\frac{d\chi(z)}{dz}\right] \tag{5b}$$

$$\epsilon_{ss}=\frac{\partial \tilde{u}_s}{\partial s}\approx n\frac{d^2\psi_n^\chi}{ds^2}\chi(z). \tag{5c}$$

The nonvanishing stress components are simply

$$\sigma_{zz}=\frac{E}{1-\nu^2}(\epsilon_{zz}+\nu\epsilon_{ss}), \tag{6a}$$

$$\sigma_{ss}=\frac{E}{1-\nu^2}(\epsilon_{ss}+\nu\epsilon_{zz}), \tag{6b}$$

$$\sigma_{sz}=2G\epsilon_{sz} \tag{6c}$$

Integrating over the beam cross section the following expression for the system potential energy Π:

$$\Pi=\frac{1}{2}\int \sigma_{ij}\epsilon_{ij}dV-\int (p\tilde{u}_z+q\tilde{q}_s)dV, \tag{7}$$

one can find the potential energy in one-dimensional form (see [8,9,13] for the explicit expression for simply connected sections).

$$\Pi[\theta,\chi,U^\theta,U^\chi]=\frac{1}{2}\int F(z,\theta,\chi,U^\theta,U^\chi,\cdots)dz$$

3 Section Displacements in Multicell Box Beams

The previous section summarized how three-dimensional displacements can be obtained from the section deformation shapes $\psi(s)$ in single-cell box beams. In this section, we present the section deformation shapes for multicell box beams. In particular, we focus on the analytic derivation of distortion and distortional warping deformation shapes for multicells, which is the main contribution of this work.

3.1 Distortion Functions $\psi_s^\chi(s)$, $\psi_n^\chi(s)$. Extending the theory to determine the distortion functions for simply connected closed beams ([8]), we present the technique to determine the distortion functions for multicell box beams. The present study will be given for single-layer multicells composed of general quadrilateral cells.

In determining the distortion functions, a two-stage procedure is employed. In the first stage, the corner rigidity resisting the distortional deformation is not taken into account; thus the cross section of a box beam is assumed to deform as hinged links where corners are regarded simply as hinges. In this setting, the distortion can be described by the amount of rigid-body tangential ($\bar{\psi}_{si}^\chi$) and normal ($\bar{\psi}_{ni}^\chi$) translations and a rigid-body rotation ($\bar{\psi}_{\theta i}^\chi$) of the ith wall. The normal $\psi_{ni}^\chi(s_i)$ and tangential $\psi_{si}^\chi(s_i)$ displacements of the ith wall can be written in the following form:

$$\begin{aligned}\psi_{si}^\chi(s_i)&=\bar{\psi}_{si}^\chi\\ \psi_{ni}^\chi(s_i)&=\bar{\psi}_{ni}^\chi-(s_i-b_i/2)\bar{\psi}_{\theta i}^\chi\\ 0&\le s_i\le b_i,\quad i=1,\cdots m\end{aligned} \tag{8}$$

where b_i denotes the length of the ith wall and m is the number of walls of the multicell section. It is worth noting that the number m of walls for a cross section having M cells is $3M+1$. Therefore there are $9M+3$ unknowns since each wall has three unknowns ($\bar{\psi}_{si}^\chi,\bar{\psi}_{ni}^\chi,\bar{\psi}_{\theta i}^\chi$). Though the present work is limited to commonly used single-layer multicells, the approach developed in this work can be extended to multilayer multicells.

To determine the distortion functions, the following conditions must be satisfied:

(i) Corner displacement conditions (in-plane components): At corners where two walls join (see Fig. 3(*a*))

$$\psi_{si}^\chi|_{s_i=b_i}=\psi_{sj}^\chi|_{s_j=0}\cos(\alpha_j-\alpha_i)-\psi_{nj}^\chi|_{s_j=0}\sin(\alpha_j-\alpha_i) \tag{9a}$$

and

$$\psi_{ni}^\chi|_{s_i=b_i}=\psi_{sj}^\chi|_{s_j=0}\sin(\alpha_j-\alpha_i)+\psi_{nj}^\chi|_{s_j=0}\cos(\alpha_j-\alpha_i). \tag{9b}$$

At corners where three walls join (see Fig. 3(*b*))

$$\psi_{si}^\chi|_{s_i=b_i}=\psi_{sj}^\chi|_{s_j=0}\cos(\alpha_j-\alpha_i)-\psi_{nj}^\chi|_{s_j=0}\sin(\alpha_j-\alpha_i), \tag{10a}$$

$$\psi_{ni}^\chi|_{s_i=b_i}=\psi_{sj}^\chi|_{s_j=0}\sin(\alpha_j-\alpha_i)+\psi_{nj}^\chi|_{s_j=0}\cos(\alpha_j-\alpha_i) \tag{10b}$$

and

$$\psi_{si}^\chi|_{s_i=b_i}=\psi_{sk}^\chi|_{s_k=0}\cos(\alpha_k-\alpha_i)-\psi_{nk}^\chi|_{s_k=0}\sin(\alpha_k-\alpha_i), \tag{10c}$$

$$\psi_{ni}^\chi|_{s_i=b_i}=\psi_{sk}^\chi|_{s_k=0}\sin(\alpha_k-\alpha_i)+\psi_{nk}^\chi|_{s_k=0}\cos(\alpha_k-\alpha_i). \tag{10d}$$

In Eqs. (9), (10), α_i is the angle between the horizontal x-axis and the tangential direction of the ith wall. Note that Eqs. (9), (10) provide $8M$ conditions for an M-cell section.

(ii) Zero virtual work conditions: The stress field by the distortion deformation should not produce any virtual work under virtual rigid-body rotation (see Eq. (11*a*)) and displacements (see Eqs. (11b,c)).

$$\int r\cdot\psi_s^\chi(s)dA=\sum_{i=1}^{m} r\cdot\bar{\psi}_{si}^\chi b_i t=0 \tag{11a}$$

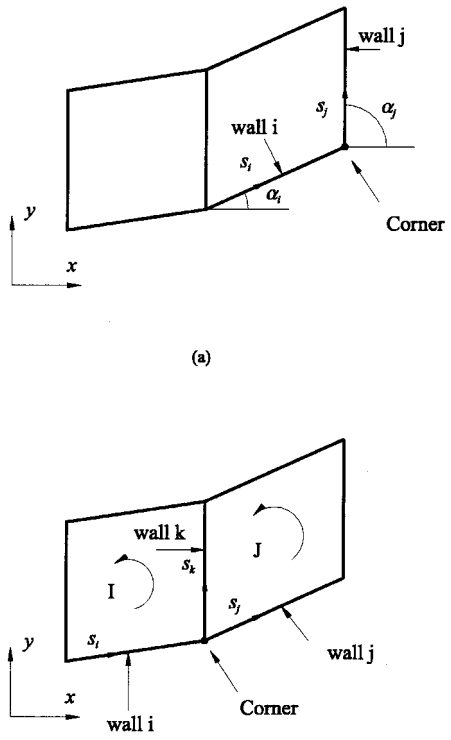


Fig. 3 (*a*) A corner at which two walls meet and (*b*) a corner at which three walls meet

$$\int \cos\alpha \cdot \psi_s^{\chi}(s)\,dA = \sum_{i=1}^{m} \cos\alpha_i \cdot \overline{\psi}_{si}^{\chi} b_i t = 0 \qquad (11b)$$

$$\int \sin\alpha \cdot \psi_s^{\chi}(s)\,dA = \sum_{i=1}^{m} \sin\alpha_i \cdot \overline{\psi}_{si}^{\chi} b_i t = 0 \qquad (11c)$$

Note that r in Eq. (11*a*) represents the rigid-body rotation field and $\cos\alpha$ and $\sin\alpha$ are used to express the resultant forces by σ_{sz}^{χ} in the direction of the x and y-axes. The conditions stated as Eqs. (11) may be viewed as the orthogonality conditions of the distortional deformation with respect to three in-plane rigid-body motions. Since Eqs. (9) through (11) give $8M+3$ conditions for $9M+3$ unknowns, M different distortion functions are possible.

CASES STUDY: Two-Cell Box Beam. To understand the nature of the distortion functions, we consider a specific example of a two-cell box beam shown in Fig. 4. Since $M=2$, there will be two independent distortion functions ψ^1 and ψ^2. Using Eq. (8) and conditions (9)–(11), one can find

$$\begin{Bmatrix} \overline{\psi}_{s1}^{\chi} \\ \overline{\psi}_{s2}^{\chi} \\ \overline{\chi}_{s3}^{\chi} \\ \overline{\psi}_{s4}^{\chi} \\ \overline{\psi}_{s5}^{\chi} \\ \overline{\psi}_{s6}^{\chi} \\ \overline{\psi}_{s7}^{\chi} \end{Bmatrix}^{1} = \begin{Bmatrix} 0 \\ 0 \\ -\frac{10}{3} \\ 0 \\ 0 \\ \frac{10}{3} \\ \frac{20}{3} \end{Bmatrix} \times K_1,$$

$$\begin{Bmatrix} \overline{\psi}_{n1}^{\chi} \\ \overline{\psi}_{n2}^{\chi} \\ \overline{\chi}_{n3}^{\chi} \\ \overline{\psi}_{n4}^{\chi} \\ \overline{\psi}_{n5}^{\chi} \\ \overline{\psi}_{n6}^{\chi} \\ \overline{\psi}_{n7}^{\chi} \end{Bmatrix}^{1} = \begin{Bmatrix} -\frac{5}{3} \\ -\frac{5}{3} \\ 0 \\ \frac{5}{3} \\ \frac{5}{3} \\ 0 \\ 0 \end{Bmatrix} \times K_1,$$

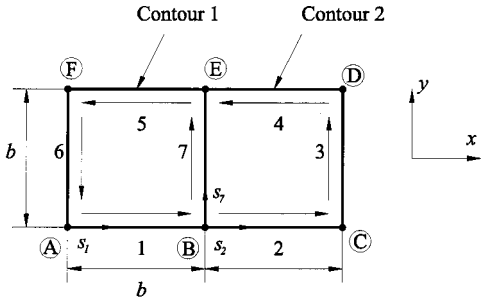


Fig. 4 A typical two-cell box beam

$$\begin{Bmatrix} \overline{\psi}_{\theta 1}^{\chi} \\ \overline{\psi}_{\theta 2}^{\chi} \\ \overline{\chi}_{\theta 3}^{\chi} \\ \overline{\psi}_{\theta 4}^{\chi} \\ \overline{\psi}_{\theta 5}^{\chi} \\ \overline{\psi}_{\theta 6}^{\chi} \\ \overline{\psi}_{\theta 7}^{\chi} \end{Bmatrix}^{1} = \begin{Bmatrix} 1 \\ -1 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{Bmatrix} \times K_1, \qquad (12)$$

and

$$\begin{Bmatrix} \overline{\psi}_{s1}^{\chi} \\ \overline{\psi}_{s2}^{\chi} \\ \overline{\chi}_{s3}^{\chi} \\ \overline{\psi}_{s4}^{\chi} \\ \overline{\psi}_{s5}^{\chi} \\ \overline{\psi}_{s6}^{\chi} \\ \overline{\psi}_{s7}^{\chi} \end{Bmatrix}^{2} = \begin{Bmatrix} -10 \\ -10 \\ 10 \\ -10 \\ 10 \\ 0 \end{Bmatrix} \times K_2,$$

$$\begin{Bmatrix} \overline{\psi}_{n1}^{\chi} \\ \overline{\psi}_{n2}^{\chi} \\ \overline{\chi}_{n3}^{\chi} \\ \overline{\psi}_{n4}^{\chi} \\ \overline{\psi}_{n5}^{\chi} \\ \overline{\psi}_{n6}^{\chi} \\ \overline{\psi}_{n7}^{\chi} \end{Bmatrix}^{2} = \begin{Bmatrix} 5 \\ -5 \\ 0 \\ 5 \\ -5 \\ 0 \\ 0 \end{Bmatrix} \times K_2,$$

$$\begin{Bmatrix} \overline{\psi}_{\theta 1}^{\chi} \\ \overline{\psi}_{\theta 2}^{\chi} \\ \overline{\chi}_{\theta 3}^{\chi} \\ \overline{\psi}_{\theta 4}^{\chi} \\ \overline{\psi}_{\theta 5}^{\chi} \\ \overline{\psi}_{\theta 6}^{\chi} \\ \overline{\psi}_{\theta 7}^{\chi} \end{Bmatrix}^{2} = \begin{Bmatrix} 1 \\ 1 \\ -2 \\ 1 \\ 1 \\ -2 \\ -2 \end{Bmatrix} \times K_2, \qquad (13)$$

where K_1 and K_2 are arbitrary constants and will be set equal to 1 without loss of generality. The deformation shapes corresponding to ψ^1 and ψ^2 are plotted in Figs. 5(*a*) and (*b*), respectively. It is apparent that the deformation shapes shown in Figs. 5(*a*) and (*b*) alone violate the slope continuity at corners. Since the use of these distortional deformation shapes results in serious distortional stiffness reduction, some modifications must be made; both slope and moment continuities will be imposed at the corners.

To this end, we replace the normal displacement $\psi_n^{\chi}(s)$ in Eq. (8) by cubic polynomial functions as

$$\psi_{ni}^{\chi}(s) = \left(\frac{\beta_{i(0)} + \beta_{i(b_i)} + 2\overline{\psi}_{\theta i}^{\chi}}{b_i^2} \right) s_i^3 - \left(\frac{2\beta_{i(0)} + \beta_{i(b_i)} + 3\overline{\psi}_{\theta i}^{\chi}}{b_i} \right) s_i^2 + \beta_{i(0)} s_i + \overline{\psi}_{\theta i}^{\chi} \frac{b_i}{2} + \overline{\psi}_{ni}^{\chi} \qquad (14)$$

where new variables $\beta_{i(0)}$ and $\beta_{i(b_i)}$ represent the amounts of the slope of the ith wall at both ends, which are defined as

$$\beta_{i(0)} = \left. \frac{d\psi_{nj}^{\chi}}{ds} \right|_{s_i=0}$$

$$\beta_{i(b_i)} = \left. \frac{d\psi_{ni}^{\chi}}{ds} \right|_{s_i=b_i}. \qquad (15)$$

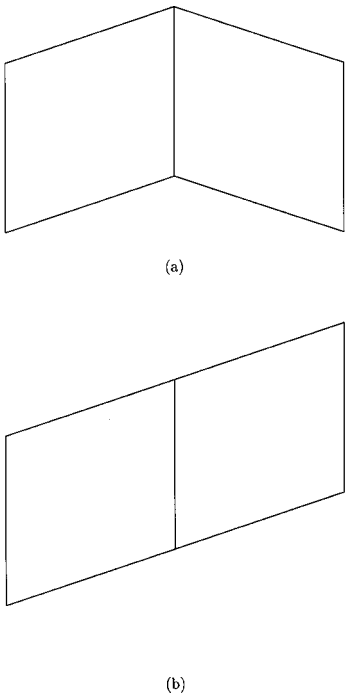


Fig. 5 Two independent preliminary distortion modes. These modes do not satisfy the rotation continuity and moment equilibrium at corners.

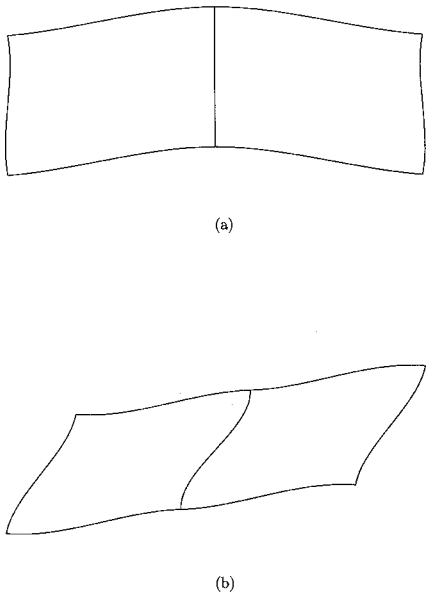


Fig. 6 Two independent distortion modes satisfying the rotation continuity and moment equilibrium at corners

Note that $\psi_{ni}^{\chi}(s_i)$ in Eq. (14) are selected to yield exactly the same corner displacements as those by Eq. (8):

$$\psi_{ni}^{\chi}(0)=\overline{\psi}_{ni}^{\chi}+\frac{b_i}{2}\overline{\psi}_{\theta i}^{\chi};\quad \psi_{ni}^{\chi}(b_i)=\overline{\psi}_{nj}^{\chi}-\frac{b_i}{2}\overline{\psi}_{\theta i}^{\chi}.$$

Furthermore, $\overline{\psi}_{ni}^{\chi}$ and $\overline{\psi}_{\theta i}^{\chi}$ in Eq. (14) are the constants that have been determined by Eqs. (8)–(11). Therefore, $\psi_{ni}^{\chi}(s_i)$ of Eq. (14) automatically satisfy the conditions (9) through (11) for any values of $\beta_{i(0)}$ and $\beta_{i(b_i)}$.

To determine the remaining unknowns $\beta_{i(0)}$ and $\beta_{i(b_i)}$ in Eq. (14), the slope continuity and moment equilibrium conditions are imposed at each corner.

For corners where two walls meet (see Fig. 3(*a*))

$$\beta_{i(b_i)}=\beta_{j(0)} \tag{16a}$$

$$M_i^b|_{s_i=b_i}=M_j^b|_{s_j=0} \tag{16a}$$

and for corners where three walls meet (see Fig. 3(*b*))

$$\beta_{i(b_i)}=\beta_{j(0)} \tag{17a}$$

$$\beta_{i(b_i)}=\beta_{k(0)} \tag{17b}$$

$$M_i^b|_{s_i=b_i}=M_j^b|_{s_j=0}+M_k^b|_{s_k=0} \tag{17c}$$

where M_i^b is the (in-plane) transverse bending moment defined as

$$M_i^b(s_i)=\frac{Et_i^3}{12}\frac{d^2\psi_{ni}^{\chi}(s_i)}{ds_i^2}.$$

Since there are $4M$ slope continuities and $2M+2$ moment equilibrium conditions for $2m=6M+2$ unknowns ($\beta_{i(0)},\beta_{i(b_i)}$), all of the unknowns are uniquely determined.

CASE STUDY: Two-Cell Box Beam. For the case of the two-cell box beam shown Fig. 4, the modification of ψ_{ni}^{χ} using Eq. (14) yields the results shown in Fig. 6. From Fig. 5 and Fig. 6, it is clear that the consideration of the corner condition is very important to obtain the correct in-plane distortional stiffness of box beams.

3.2 Distortional Warping Function $\psi_z^{U^{\chi}}(s)$. There are some reports ([10,11]) discussing how to find distortional warping functions in some cross sections.

Nonuniform torsion (see [1]) produces an additional shear flow contributing to the twisting moment. This additional shear flow may be called a distortional shear flow, and the negligence of this shear flow would yield incorrect distortional warping deformations. We introduce a constant distortional shear flow in each cell in order to guarantee the singlevaluedness of the distortional warping function.

If we denote the distortional shear flow by q^{χ} in a single-cell box beam, the distortional warping function $\psi_z^{U^{\chi}}$ may be found from

$$\frac{q^{\chi}(z)}{Gt}=\frac{\partial u_z^{U^{\chi}}}{\partial s}+\frac{\partial u_s^{\chi}}{\partial z}. \tag{18}$$

In Eq. (18), u_s^{χ} is the tangential displacement by the distortional deformation, which has already been given as $u_s^{\chi}(s,z)\equiv\psi_s^{\chi}(s)\chi(z)$ (see Eq. (1*a*)). Likewise, the axial displacement $u_z^{U^{\chi}}(s,z)$ may be written as

$$u_x^{U^{\chi}}(s,z)=\psi_z^{U^{\chi}}(s)U^{\chi}(z). \tag{19}$$

To determine $\psi_z^{U^{\chi}}(s)$, the shear stress flow $q^{\chi}(z)$ must be found first from the single-valuedness of the axial displacement $u_z^{U^{\chi}}(s,z)$:

$$0=\oint_C du_z^{U^{\chi}}(s,z)=\oint\frac{\partial u_z^{U^{\chi}}}{\partial s}ds$$

$$=\frac{q^{\chi}(z)}{G}\oint_C\frac{ds}{t}-\frac{d\chi(z)}{dz}\oint_C\psi_s^{\chi}(s)ds. \tag{20}$$

Rearranging Eq. (20) for $q^{\chi}(z)$ yields

$$q^{\chi}=G\frac{\oint_C \psi_s^{\chi}ds}{\oint_C \frac{ds}{t}}\frac{d\chi}{dz}=G\Gamma^{\chi}\frac{d\chi}{dz} \tag{21}$$

where

$$\Gamma^{\chi}=\frac{\Omega^{\chi}}{\oint_C \frac{ds}{t}} \tag{22}$$

and

$$\Omega^{\chi}=\oint_C \psi_s^{\chi}ds. \tag{23}$$

Using $q^{\chi}(z)$ expressed as Eq. (21), one can rewrite Eq. (18) as

$$\frac{\partial u_z^{\chi}(s,z)}{\partial s}=\frac{d\psi_z^{U^{\chi}}(s)}{ds}U^{\chi}(z)\equiv\left(\frac{\Gamma^{\chi}}{t}-\psi_s^{\chi}(s)\right)\frac{d\chi(z)}{dz}. \tag{24}$$

With $U^{\chi}(z)=d\chi(z)/dz$, one obtains

$$\frac{d\psi_z^{U^{\chi}}(s)}{ds}=\frac{\Gamma^{\chi}}{t}-\psi_s^{\chi}(s) \tag{25}$$

and integrating Eq. (25) yields

$$\psi_z^{U^{\chi}}(s)=\int_{s_0}^{s}\left[\frac{\Gamma^{\chi}}{t}-\psi_s^{\chi}(s)\right]ds+\psi_{z0}^{U^{\chi}}. \tag{26}$$

The integration constant $\psi_{z0}^{U^{\chi}}$ can be determined from the condition that the virtual work by the distortion warping stress due to the virtual extensional axial displacement must vanish:

$$\oint_C \psi_z^{U^{\chi}}(s)ds=0. \tag{27}$$

A similar procedure can be applied to M-cell box beams. In this case, M shear flows defined in each of M cells must be considered. In a typical multicell cross section shown in Fig. 7, the shear flows

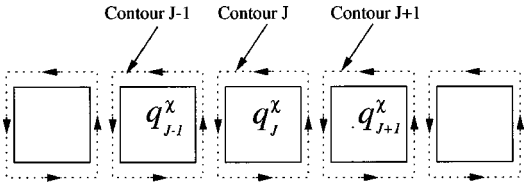


Fig. 7 Shear flows of three adjacent cells $J-1$, J, $J+1$ of a multicell cross section associated with distortion and distortional warping

in three adjacent cells (the $J-1$th, Jth and $J+1$th cells) are denoted by q_{J-1}^{χ}, q_J^{χ}, and q_{J+1}^{χ}.

Applying the singlevaluedness condition expressed as Eq. (20) around the Jth cell yields

$$\begin{aligned}0=G\oint_{C_J}du_z&=-q_{J-1}^{\chi}\int_{C_{J-1J}}\frac{ds}{t}+q_J^{\chi}\oint_{C_J}\frac{ds}{t}-q_{J+1}^{\chi}\\&\quad\times\int_{C_{JJ+1}}\frac{ds}{t}-G\frac{d\chi(z)}{dz}\oint_{C_J}\psi_s^{\chi}ds\\&=-q_{J-1}^{\chi}\int_{C_{J-1J}}\frac{ds}{t}+q_J^{\chi}\oint_{C_J}\frac{ds}{t}\\&\quad-q_{J+1}^{\chi}\int_{C_{JJ+1}}\frac{ds}{t}-G\Omega_J\frac{d\chi(z)}{dz}.\end{aligned} \tag{28}$$

In Eq. (28), the subscript J denotes the quantities associated with the Jth contour, and $\int_{C_{J-1J}}$ represents the integral along the wall common to the $J-1$th and Jth cells. Writing Eq. (28) for $J=1,2,\cdots,M$ and putting the resulting equations in matrix form, one may obtain

$$\mathbf{C}\cdot\mathbf{q}^{\chi}=G\frac{d\chi}{dz}\mathbf{\Omega}^{\chi} \tag{29}$$

where

$$\mathbf{C}=\begin{bmatrix}\oint_{C_1}\frac{ds}{t} & -\int_{C_{12}}\frac{ds}{t} & 0 & 0 & \cdots & 0\\ -\int_{C_{21}}\frac{ds}{t} & \oint_{C_2}\frac{ds}{t} & -\int_{C_{23}}\frac{ds}{t} & 0 & \cdots & 0\\ & \ddots & \ddots & \ddots & & \\ 0 & -\int_{C_{J-1J}}\frac{ds}{t} & \oint_{C_J}\frac{ds}{t} & -\int_{C_{JJ+1}}\frac{ds}{t} & \cdots & 0\\ & & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 0 & \int_{C_{M-1M}}\frac{ds}{t} & \oint_{C_M}\frac{ds}{t}\end{bmatrix}, \tag{30}$$

$$\mathbf{q}^{\chi}=\{q_1^{\chi},q_2^{\chi},\cdots,q_M^{\chi}\}^T, \tag{31}$$

$$\mathbf{\Omega}^{\chi}=\{\Omega_1^{\chi},\Omega_2^{\chi},\cdots,\Omega_M^{\chi}\}^T. \tag{32}$$

Inverting Eq. (29) for $\mathbf{q}^{\chi}$ yields

$$\mathbf{q}^{\chi}=G\mathbf{C}^{-1}\mathbf{\Omega}^{\chi}\frac{d\chi}{dz}=G\mathbf{\Gamma}^{\chi}\frac{d\chi}{dz} \tag{33}$$

where

$$\mathbf{\Gamma}^{\chi}=(\mathbf{C})^{-1}\mathbf{\Omega}^{\chi}. \tag{34}$$

Since $\mathbf{\Gamma}^{\chi}$ can be determined for all cells, the distortional warping function $\psi_z^{U^{\chi}}(s)$ can be found by extending Eq. (26) to the case of a multicell section. Referring to Fig. 3(*b*), one can find

$$\psi_z^{U^\chi}(s_i)=\int_{s_0}^{s_i}\left[\frac{\Gamma_I^\chi}{t}-\psi_s^\chi(s_i)\right]ds+\psi_{zi0}^{U^\chi} \tag{35a}$$

and

$$\psi_z^{U^\chi}(s_k)=\int_{s_0}^{s_k}\left[\frac{\Gamma_I^\chi-\Gamma_J^\chi}{t}-\psi_s^\chi(s_k)\right]ds+\psi_{zk0}^{U^\chi}. \tag{35b}$$

In order to determine the integration constants $\psi_{zi0}^{U^\chi}$, $\psi_{zk0}^{U^\chi}$, Eq. (27) and the continuity conditions of $\psi_z^{U^\chi}$ at corners must be used.

CASE STUDY: Two-cell Box Beam. We reconsider the two-cell box beam shown in Fig. 4 in order to give the distortional warping function explicitly. In this case, there exist two distortional warping functions $\psi_z^{U^{\chi_1}}$, $\psi_z^{U^{\chi_2}}$ corresponding to the two distortion functions $\psi_z^{\chi_1}$, $\psi_z^{\chi_2}$. Using $\overline{\psi}_{si}^{\chi_1}$ (with $K_1=1$) in Eq. (12), $\Omega_J^{\chi_1}$ are determined as

$$\Omega_1^{\chi_1}=\oint_{C_1}\psi_s^{\chi_1}ds=\overline{\psi}_{s1}^{\chi_1}b_1+\overline{\psi}_{s7}^{\chi_1}b_7+\overline{\psi}_{s5}^{\chi_1}b_5+\overline{\psi}_{s6}^{\chi_1}b_6$$

$$=0+\frac{20}{3}\cdot 10+0+\frac{10}{3}\cdot 10=100 \tag{36a}$$

$$\Omega_2^{\chi_1}=\oint_{C_2}\psi_s^{\chi_2}ds=\overline{\psi}_{s2}^{\chi_1}b_2+\overline{\psi}_{s3}^{\chi_1}b_3+\overline{\psi}_{s4}^{\chi_1}b_4-\overline{\psi}_{s7}^{\chi_1}b_7$$

$$=0-\frac{10}{3}\cdot 10+0-\frac{20}{3}\cdot 10=-100. \tag{36b}$$

The matrix $\mathbf{C}$ in Eq. (29) is simply

$$\mathbf{C}=\begin{bmatrix}\frac{40}{t} & -\frac{10}{t}\\ -\frac{10}{t} & \frac{40}{t}\end{bmatrix}. \tag{37}$$

Substituting Eqs. (36) and (37) in Eq. (34) yields

$$\Gamma_1^{\chi_1}=2t,\quad \Gamma_2^{\chi_1}=-2t. \tag{38}$$

Once Γ_J^χ are determined, the distortional warping function ψ_z^χ can be obtained from Eqs. (35). The explicit expressions for each wall are given below.

$$\psi_{z1}^{U^{\chi_1}}(s_1)=\int_0^{s_1}\left(\frac{\Gamma_1^{\chi_1}}{t}-\psi_s^{\chi_1}(s_1)\right)ds_1+\psi_{z0}^{U^{\chi_1}}$$

$$=2s_1+\psi_{z0}^{U^{\chi_1}}\quad 0\leqslant s_1\leqslant b_1=b=10$$

$$\psi_{z2}^{U^{\chi_1}}(s_2)=\int_0^{s_2}\left(\frac{\Gamma_2^{\chi_1}}{t}-\psi_{s2}^{\chi_1}\right)ds_2+\psi_{z1}^{U^{\chi_1}}(b_1)$$

$$=-2s_2+2b+\psi_{z0}^{U^{\chi_1}}\qquad 0\leqslant s_2\leqslant b_2=b$$

$$\psi_{z3}^{U^{\chi_1}}(s_3)=\int_0^{s_3}\left(\frac{\Gamma_2^{\chi_1}}{t}-\psi_{s3}^{\chi_1}\right)ds_3+\psi_{z2}^{U^{\chi_1}}(b_2)$$

$$=\frac{4}{3}s_3+\psi_{z0}^{U^{\chi_1}}\quad 0\leqslant s_3\leqslant b_3=b$$

$$\psi_{z4}^{U^{\chi_1}}(s_4)=\int_0^{s_4}\left(\frac{\Gamma_2^{\chi_1}}{t}-\psi_{s4}^{\chi_1}\right)ds_4+\psi_{z3}^{U^{\chi_1}}(b_3)$$

$$=-2s_4+\frac{4}{3}b+\psi_{z0}^{U^{\chi_1}}\qquad 0\leqslant s_4\leqslant b_4=b$$

$$\psi_{z5}^{U^{\chi_1}}(s_5)=\int_0^{s_5}\left(\frac{\Gamma_1^{\chi_1}}{t}-\psi_{s5}^{\chi_1}\right)ds_5+\psi_{z4}^{U^{\chi_1}}(b_4)$$

$$=2s_5-\frac{2}{3}b+\psi_{z0}^{U^{\chi_1}}\qquad 0\leqslant s_5\leqslant b_5=b$$

$$\psi_{z6}^{U^{\chi_1}}(s_6)=\int_0^{s_6}\left(\frac{\Gamma_1^{\chi_1}}{t}-\psi_{s6}^{\chi_1}\right)ds_6+\psi_{z5}^{U^{\chi_1}}(b_5)$$

$$=-\frac{4}{3}s_6+\frac{4}{3}b+\psi_{z0}^{U^{\chi_1}}\qquad 0\leqslant s_6\leqslant b_6=b$$

$$\psi_{z7}^{U^{\chi_1}}(s_7)=\int_0^{s_7}\left(\frac{\Gamma_1^{\chi_1}-\Gamma_2^{\chi_1}}{t}-\psi_{s7}^{\chi_1}\right)ds_7+\psi_{z1}^{U^{\chi_1}}(b_1)$$

$$=-\frac{8}{3}s_7+2b+\psi_{z0}^{U^{\chi_1}}\qquad 0\leqslant s_7\leqslant b_7=b$$

The integration constant $\psi_{z0}^{U^{\chi_1}}$ from the condition stated by Eq. (27) is

$$\psi_{z0}^{U^{\chi_1}}=-\frac{20}{3}.$$

Repeating the same procedure for the second distortional warping function, one can find

$$\psi_{z1}^{U^{\chi_2}}(s_1)=\frac{20}{3}s_1-\frac{200}{3}\qquad 0\leqslant s_1\leqslant b_1=b=10$$

$$\psi_{z2}^{U^{\chi_2}}(s_2)=\frac{20}{3}s_2\qquad 0\leqslant s_2\leqslant b_2=b$$

$$\psi_{z3}^{U^{\chi_2}}(s_3)=-\frac{40}{3}s_3+\frac{200}{3}\qquad 0\leqslant s_3\leqslant b_3=b$$

$$\psi_{z4}^{U^{\chi_2}}(s_4)=\frac{20}{3}s_4-\frac{200}{3}\qquad 0\leqslant s_4\leqslant b_4=b$$

$$\psi_{z5}^{U^{\chi_2}}(s_5)=\frac{20}{3}s_5\qquad 0\leqslant s_5\leqslant b_5=b$$

$$\psi_{z6}^{U^{\chi_2}}(s_6)=-\frac{40}{3}s_6+\frac{200}{3}\qquad 0\leqslant s_6\leqslant b_6=b$$

$$\psi_{z7}^{U^{\chi_2}}(s_7)=0\qquad 0\leqslant s_7\leqslant b_7=b$$

Figures 8(*a*) and (*b*) show the first and second distortional warping functions, respectively.

3.3 Torsional Warping Function. Although the torsional warping is well known (see [14]), we repeat the result here for the sake of completeness. As in the case of the distortional warping functions, one can form the following matrix equation for the torsional shear flow q^θ,

$$\mathbf{C}\cdot\mathbf{q}^\theta=G\frac{d\theta}{dz}\mathbf{\Omega} \tag{39}$$

where $\mathbf{C}$ is the same matrix as defined as Eq. (30), and

$$\mathbf{\Omega}^\theta=\left\{\oint_{C_1}rds,\cdots,\oint_{C_M}rds\right\}^T.$$

Replacing $(\mathbf{q}^\chi,\mathbf{\Omega}^\chi)$ by $(\mathbf{q}^\theta,\mathbf{\Omega}^\theta)$, the analysis carried out for the distortional warping functions applies directly in determining the torsional warping functions $\psi_z^{U^\theta}(s)$.

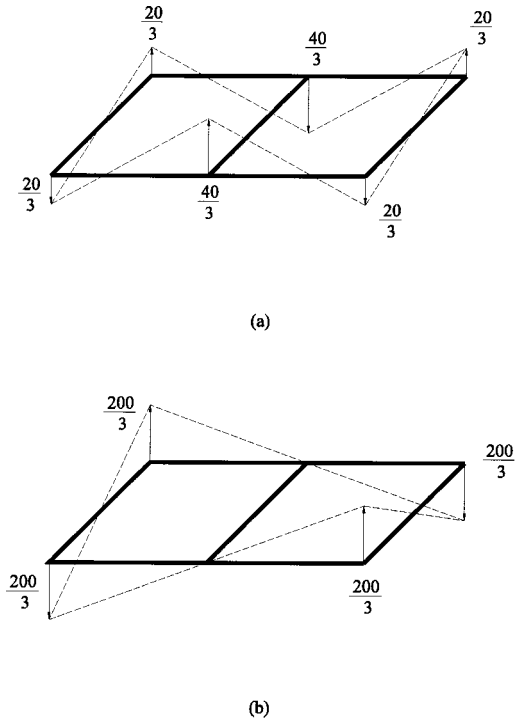


Fig. 8 Two distortional warping functions corresponding to the distortion functions shown in Fig. 6. Numbers denote the relative magnitudes of axial warping displacements.

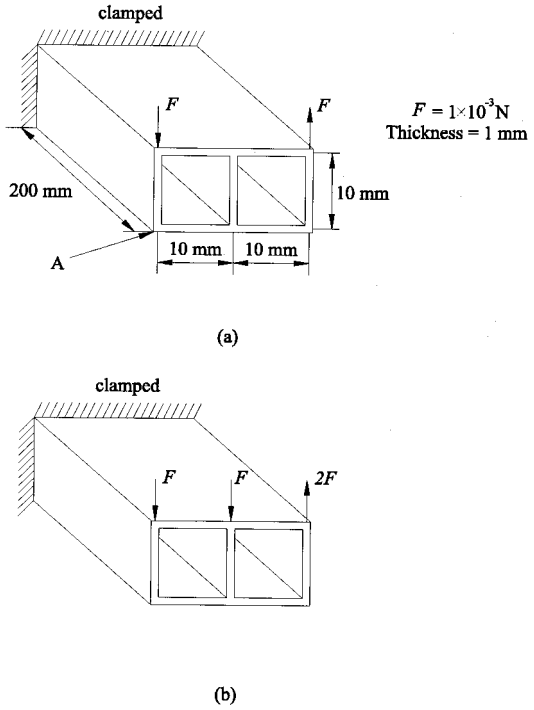


Fig. 9 A two-cell box beam subjected to (*a*) a couple and (*b*) a set of three concentrated loads

4 Numerical Examples

For numerical calculations, a C^0-continuous piecewise-linear one-dimensional finite beam element is implemented using the theory presented in the previous sections. In the case of two-cell beams, the nodal displacement vector **d** has five degrees-of-freedom:

$$\mathbf{d}^T=\{\theta \;\; U^\theta \;\; \chi_1 \;\; U^{\chi_1} \;\; \chi_2 \;\; U^{\chi_2}\}.$$

The rotation and the torsional warping degrees are denoted by θ and U^θ, and the two distortion and distortional warping degrees are denoted by (χ_1,χ_2) and (U^{χ_1},U^{χ_2}), respectively. Since the procedure for the finite element implementation can be found in Kim and Kim [8,9] the detailed procedure is skipped here.

Example 1 Two-Cell Box Beam Subjected to Various Load Conditions. We consider a straight two-cell box beam subjected to a couple (Fig. 9(*a*)) and a set of three point loads (Fig. 9(*b*)). Throughout the numerical examples, the magnitude of the point load F will be taken simply as 1×10^{-3} N, and Young's modulus and Poisson's ratio are taken as 2.0×10^{11} N/m^2 and 0.3, respectively. The vertical displacements u_y along A due to the loads shown in Figs. 9(*a*) and (*b*) are plotted in Figs. 10(*a*) and (*b*), respectively. The results with varying numbers (N_e) of the present beam elements are compared with those by plate elements. Figure 11 compares the present one-dimensional results and the plate finite element results for the distribution of the transverse bending stress σ_{ss} at $z=175$ mm and $n=t/2$ for the load case of Fig. 9(*a*). For the plate element analysis, IDEAS ([15]) was employed and the convergence of the plate element result was confirmed. As the number of the present beam element increases, the present one-dimensional solution converges well to the plate element result.

Example 2 More General Cross Sections. In this example, we consider more general cross sections of two-cell box beams. Figure 12 shows the dimensions of a monosymmetric trapezoidal section and the vertical displacement u_y at point A. Other geometric and boundary conditions except the section shape are the same

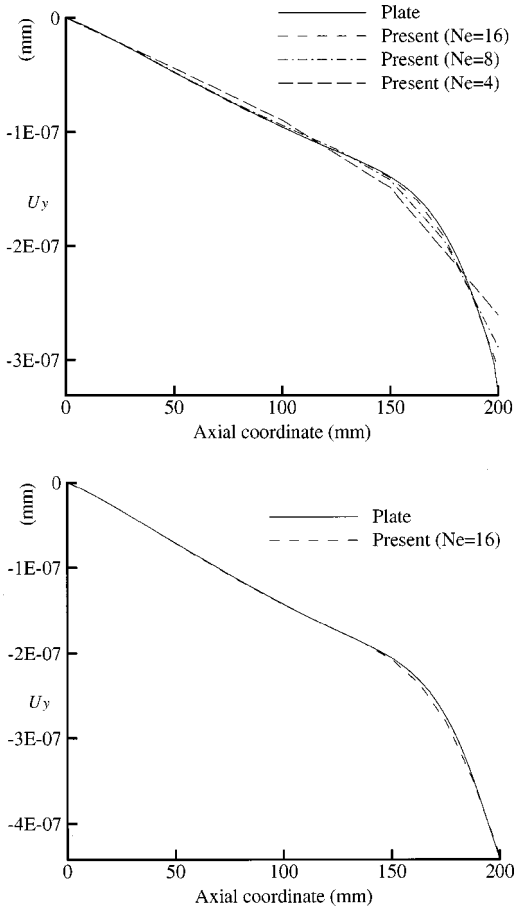


Fig. 10 The vertical displacements u_y along corner *A*. The results (*a*) and (*b*) correspond to the loading cases shown in Figs. 9(*a*) and (*b*), respectively. The number of the present beam finite elements is denoted by N_e.

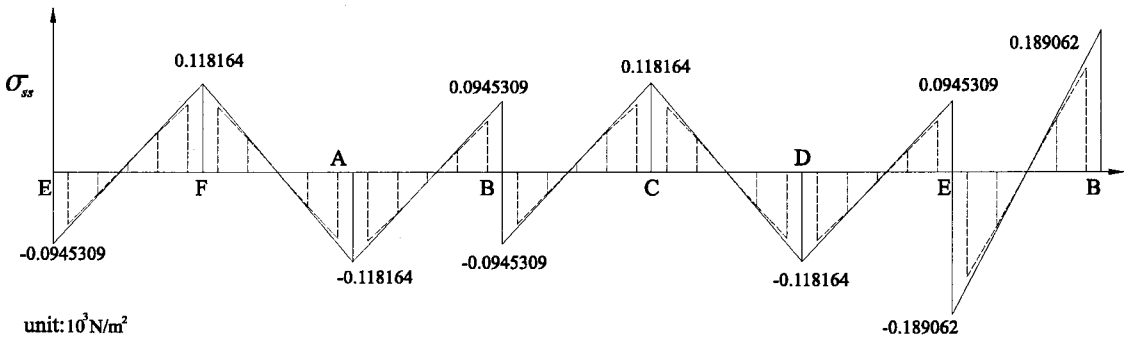


Fig. 11 The distribution of the transverse bending stress σ_{ss} at $z=175$ mm and $n=t/2$. The present results (solid lines) are compared with the finite element results (dotted lines) by IDEAS.

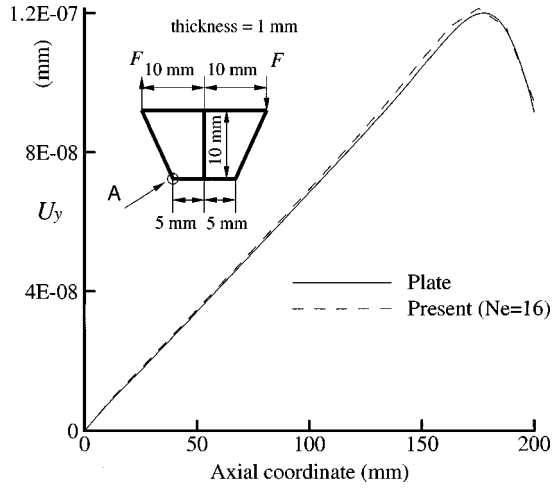


Fig. 12 The vertical displacement u_y along corner A of a trapezoidal two-cell box beam under a couple (beam length = 200 mm)

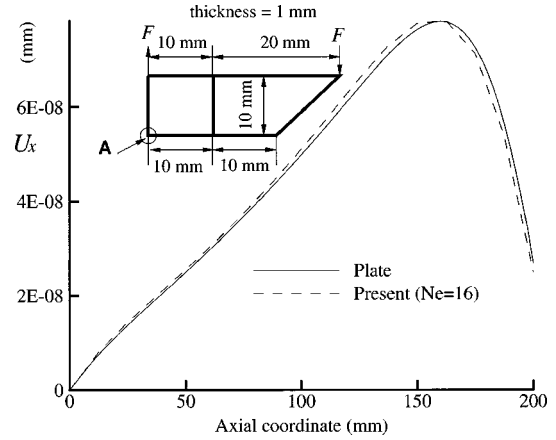


Fig. 13 The horizontal displacement u_x along corner A in an unsymmetric two-cell beam under a couple (beam length = 200 mm)

as those in Example 1. An excellent agreement between the present one-dimensional result and the plate element result is observed. We also treat an unsymmetrically shaped two-cell beam subjected to a couple as shown in Fig. 13. The horizontal displacement u_x at point A is also plotted in Fig. 13. This cross section is seldom used in practice, but it serves to confirm the validity of the present approach.

Table 1 The eigenfrequencies of a freely supported two-cell box beam whose cross section is shown in Fig. 9. (N_e denotes the number of elements.)

Mode	Plate ($N_e=3500$)	Classical Theory	Present ($N_e=20$)	Present ($N_e=40$)
1st torsion	5674.8 Hz	5987.6 Hz	5693.9 Hz	5689.6 Hz
1st distortion	8775.1 Hz	N/A	8807.5 Hz	8786 Hz
2nd distortion	9986.4 Hz	N/A	10111 Hz	9999.5 Hz
3rd distortion	10790.6 Hz	N/A	10817 Hz	10817 Hz
4th distortion	11385.6 Hz	N/A	11398 Hz	11394 Hz
5th distortion	11769.4 Hz	N/A	12330 Hz	11958 Hz
6th distortion	12893.7 Hz	N/A	12418 Hz	12418 Hz
7th distortion	13697.9 Hz	N/A	13164 Hz	13164 Hz

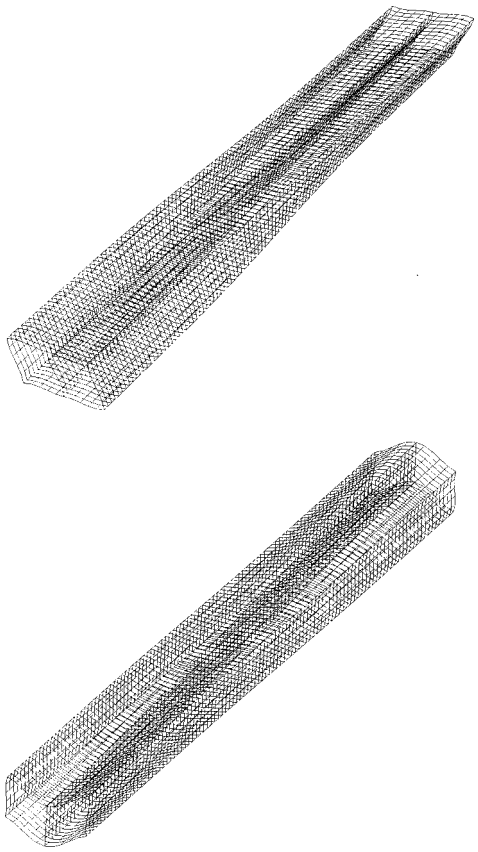


Fig. 14 (*a*) The fourth and (*b*) the seventh distortional eigenmodes of a two-cell box beam with a freely supported condition

It is remarked that the present one-dimensional theory accurately predicts the local end effect near the loaded end. The displacements shown in Figs. 10, 12, and 13 vary rapidly at the end. If one does not consider the distortional deformation, the linear displacement distributions would appear in these figures. See Balch and Steele [6] for more rigorous analysis of end effects in thin-walled closed beams.

Example 3 Free Vibration Analysis of a Two-Cell Box Beam We perform the vibration analysis of a two-cell box beam with its both ends free. The dimensions and material properties of the beam are the same as those used in Example 1. The lowest torsional and distortional eigenfrequencies of the beam are listed in Table 1. The present results show an excellent agreement with those of the plate finite elements while the classical torsion theory by St. Venant cannot predict the distortional eigenfrequencies. Figure 14 shows the fourth and seventh distortional eigenmodes of the two-cell box beam.

5 Conclusions

Section deformation functions associated with distortion and distortional warping for general multicell beams were presented. The distortional shear flow was utilized to derive the distortional warping function. Excellent agreement was found between the present one-dimensional beam results and the plate finite element results for both static and dynamic analyses. In particular, local end-effects were accurately predicted even with the present one-dimensional analysis.

Acknowledgment

This work was supported by the Brain Korea 21 Project.

References

[1] Timoshenko, S. P., and Gere, J. M., 1963, *Theory of Elastic Stability*, 2nd ed, McGraw-Hill, New York.
[2] Vlasov, V. Z., 1961, *Thin Walled Elastic Beams*, Israel Program for Scientific Translations, Jerusalem.
[3] Křístek, V., 1970, ''Tapered Box Girders of Deformable Cross Section,'' J. Struct. Div. ASCE, **96**, No. ST8, Proc. Paper 7489, pp. 1761–1793.
[4] Wright, R. N., Abdel-Samad, S. R., and Robinson, A. R., 1968, ''BEF Analogy for Analysis of Box Girders,'' J. Struct. Div. ASCE, **94**, No. St7, pp. 1719–1743.
[5] Boswell, L. F., and Zhang, S. H., 1984, ''The Effect of Distortion in Thin-Walled Box-Spine Beams,'' Int. J. Solids Struct., **20**, No. 9–10, pp. 845–862.
[6] Balch, C. D., and Steele, C. R., 1987, ''Asymptotic Solutions for Warping and Distortion of Thin-Walled Box Beams,'' ASME J. Appl. Mech., **54**, pp. 165–173.
[7] Hsu, Y. T., Fu, S. S., and Schelling, D. R., 1995, ''EBEF Method for Distortional Analysis of Steel Box Girder Bridges,'' J. Struct. Div. ASCE, **121**, No. 3, pp. 557–566.
[8] Kim, J. H., and Kim, Y. Y., 1999, ''Analysis of Thin-Walled Closed Beams With General Quadrilateral Cross Sections,'' ASME J. Appl. Mech., **66**, pp. 904–912.
[9] Kim, J. H., and Kim, Y. Y., 2000, ''One-Dimensional Analysis of Thin-Walled Closed Beams Having General Cross Sections,'' Int. J. Numer. Methods Eng., **49**, pp. 653–668.
[10] Jönsson, J., 1999, ''Distortional Warping Functions and Shear Distributions in Thin-Walled Beams,'' Thin-Walled Struct., **33**, pp. 245–268.
[11] Razaqpur, A. G., and Li, H., 1991, ''Thin-Walled Multicell Box-Girder Finite Element,'' J. Struc. Eng. ASCE, **117**, No. 10, pp. 2953–2971.
[12] Razaqpur, A. G., and Li, H., 1994, ''Refined Analysis of Curved Thin-Walled Multicell Box-Girders,'' Comput. Struct., **53**, No. 1, pp. 131–142.
[13] Kim, Y. Y., and Kim, J. H., 1999, ''Thin-Walled Closed Box Beam Element for Static and Dynamic Analysis,'' Int. J. Numer. Methods Eng., **45**, pp. 473–490.
[14] Gjelsvik, A., 1981, *The Theory of Thin Walled Bars*, John Wiley and Sons, New York.
[15] SDRC, 1993, *I-DEAS FEM Users' Guide*, Structural Dynamics Research Corporation, Milford.

Hysteresis Behavior and Modeling of Piezoceramic Actuators

X. Zhou

A. Chattopadhyay
Mem. ASME

Department of Mechanical and Aerospace Engineering,
Arizona State University,
Tempe, AZ 85287-6106

A new theory is developed to model the hysteresis relation between polarization and electric field of piezoceramics. An explicit formulation governing the hysteresis is obtained by using saturation polarization, remnant polarization, and coercive electric field. A new form of elastic Gibbs energy is proposed to address the coupling relations between electrical field and mechanical field. The nonlinear constitutive relations are derived from the elastic Gibbs energy and are applicable in the case of high stroke actuation. The hysteresis relations obtained using the current model are correlated with experimental results. The static deflection of a cantilever beam with surface-bonded piezoelectric actuators is analyzed by implementing the current constitutive relations. Numerical results reveal that hysteresis is an important issue in the application of piezoceramics.
[DOI: 10.1115/1.1357168]

Introduction

Piezoelectricity refers to the relationship between pressure and electricity that exists within a unique family of materials, called piezoelectric materials. Piezoelectric materials, especially PZT ceramics, have received considerable attention in the past decade for their wide applications to augment stability and control vibration ([1]). Traditional studies of their applications in smart structures are based on linear piezoelectric model ([2,3]) which implies both low electric fields and low mechanical strains. However, greater actuation authority can be achieved by applying an electric field exceeding the limit of a linear piezoelectric constitutive relation to increase induced strains. Piezoceramics exhibit constitutive nonlinearity of hysteresis due to the variation of polarization if the applied electric field is above the coercive limit. On the other hand, hysteresis behavior is also related to mechanical loading. Different hysteresis loops are observed with varying mechanical stresses.

For the analysis of a smart structural system, it is necessary to develop efficient theories that are capable of addressing material nonlinearities and accurately predicting hysteresis phenomenon in actuators under high electrical excitations accompanied by mechanical stresses. It has been well known that the application of high electric field and mechanical stress has the effect of changing the polar axis of material unit cells. The process of domain polarization reorientation creates repolarization, which introduces the hysteresis loop shown in piezoelectric material experiments ([4,5]). This nonlinear phenomenon of piezoelectric hysteresis has been studied by innumerable researchers at two different scales. Microscopic theories, based on individual domain or a simple collection of domains, present only some physical insights into piezoelectricity and provide no successful model to accurately predict the response of practical actuators. Phenomenological models, on the other hand, are proposed to match experimental results with assumptions motivated by certain piezoelectric characteristics. Devonshire [6,7] formulated a model of ferroelectricity by expanding the free energy as a function of polarization and strain. Further study on thermodynamic energy and invariance requirements was performed by Chowdhury et al. [8] to investigate nonlinear elastic dielectrics including polarization gradient effects. One-dimensional domain switching was studied and a rate law of the effective aligned dipoles was proposed by Chen and Montgomery [9]. Chan and Hagood [4] and Ghandi and Hagood [10] studied polarization variation by using the method of energy and energy barrier. The marcolevel approach combined with reversible and irreversible domain wall motions in response to an applied electric field was used by Smith et al. [11] to investigate the hysteresis between polarization and an electric field. These models are more useful in engineering applications but are not completely satisfactory due to weak physical assumptions or the introduction of material-related parameters without clear physical meanings.

A physically based phenomenological approach to address the hysteresis behavior of piezoceramic actuators is developed in this paper. The piezoelectric actuator is assumed as a continuum with an oriented and switchable dipole microstructure. The internal constraint of only two types of polarization switching, 180 deg and 90 deg switching, in the case of a tetragonal piezoceramic phase under the application of mechanical loading and a high electric field, is considered. The differential equation governing the relation between polarization and an electric field is established by considering the critical energy for 180 deg and 90 deg switching and energy loss due to inclusions in the materials. A new form of elastic Gibbs free energy for piezoceramics is proposed to address the coupling effects between polarization, an electric field, and mechanical strains. A nonlinear constitutive model, indicating gradual development and critical yield of polarization due to polar axis switching of crystal domains, is obtained to simulate electromechanical motion.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 1, 2000; final revision, Aug. 28, 2000. Associate Editor: A. K. Mal. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Model Development

For a piezoelectric elastic body, the elastic Gibbs free energy in isothermal conditions can be written in terms of two independent state variables, polarization P_i and stress σ_{ij}, that is, $G(\sigma_{ij}, P_i)$. Consider the functional π as follows:

$$\pi(u_i,\sigma_{ij},\phi,E_i,D_i,P_i)$$
$$= \int_{t_0}^{t} dt \cdot \left\{ \int_V \left[\frac{1}{2}\rho \dot{u}_i \dot{u}_i + f_i u_i - \rho_e \phi - G(\sigma_{ij},P_i) - D_i \phi_i \right.\right.$$
$$\left. - \sigma_{ij}\frac{1}{2}(u_{i,j}+u_{j,i}) - \left(D_i - \frac{1}{2}\gamma_{ij}E_j - P_i \right) E_i \right] dv$$
$$+ \int_{\Gamma_u} \sigma_{ij} n_j (u_i - \overline{u_i}) ds + \int_{\Gamma_\sigma} \overline{t}_i u_i ds$$
$$\left. + \int_{\Gamma_\phi} D_i n_i (\phi - \overline{\phi}) ds + \int_{\Gamma_D} \overline{d}\phi ds \right\} \quad (1)$$

where ρ, f_i, ρ_e, and γ_{ij} denote mass density, force per unit volume, electric charge density, and dielectric permittivity, respectively. The quantities u_i and ε_{ij} denote mechanical displacement and strain, respectively, and ϕ, E_i, and D_i denote the electric potential, electric field, and electric displacement, respectively. The quantities $\overline{u_i}$, $\overline{t_i}$, $\overline{\phi}$, and $\overline{d}$ denote the prescribed deformation on the displacement boundary (Γ_u), the traction on the stress boundary (Γ_σ), the voltage on the potential boundary (Γ_ϕ) and the surface charge on the charge boundary (Γ_D), respectively, and n_i denotes the unit outward normal of the surface (Γ). The variation of the functional π with independent variation in δu_i, $\delta\sigma_{ij}$, $\delta\phi$, δE_i, δD_i, and δP_i can be written as follows:

$$\sigma_{ij,j} + f_i - \rho\ddot{u}_i = 0, \quad (2)$$

$$D_{i,i} - \rho_e = 0 \quad (3)$$

$$\frac{\partial G(\sigma_{ij},P_i)}{\partial \sigma_{ij}} + \frac{1}{2}(u_{i,j}+u_{j,i}) = 0, \quad (4)$$

$$\frac{\partial G(\sigma_{ij},P_i)}{\partial P_i} - E_i = 0 \quad (5)$$

$$E_i + \phi_{,i} = 0, \quad (6)$$

$$D_i - \gamma_{ij}E_j - P_i = 0 \quad (7)$$

$$u_i - \overline{u_i} = 0 \ \text{ on } \ S_u, \quad (8)$$

$$\overline{t_i} - \sigma_{ij}n_j = 0 \ \text{ on } \ S_\sigma \quad (9)$$

$$\varphi - \overline{\varphi} = 0 \ \text{ on } \ S_\phi, \quad (10)$$

$$\overline{d} - D_i n_i = 0 \ \text{ on } \ S_D. \quad (11)$$

Equations (2)–(11) govern the equilibrium of mechanical and electrical motions, the constitutive relations of the electromechanical field, the electrical characteristics of dielectrics, the mechanical boundary conditions, and the electrical boundary conditions, respectively. The constitutive relation will now be obtained through appropriate form of the elastic Gibbs energy $G(\sigma_{ij},P_i)$ which can address the evolution of polarization and, furthermore, illustrate the hysteresis behavior of piezoceramics. The hysteresis characteristics between polarization and electric field without stress effects will be investigated first. Then, the elastic Gibbs energy will be used to investigate the electromechanical coupling effects.

Polarization and Electric Field Hysteresis

For a piezoelectric body, let E_c, P_r, and P_s denote the coercive electric field, the remnant polarization, and the saturation polarization, respectively. The quantities E and P denote the magnitudes of the electric field vector (**E**) and the polarization vector (**P**), respectively. During the process of repolarization, polarization orientation of any crystal cell is prone to be aligned along the direction of the electric field, denoted **e**, to the best of its ability. This is achieved through either 180 deg switching or 90 deg switching with respect to the polarization orientation of the previous step. The polarization components, if they are perpendicular to the direction of the summed polarization, counteract and cancel each other. As a result, the polarization will be aligned along the direction of the electric field.

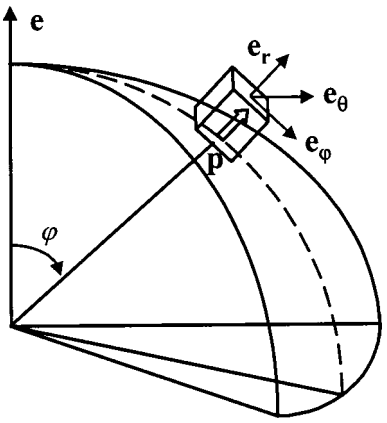


Fig. 1 Illustration of surface mapping

The tetragonal crystallites with dipole moments and inclusions such as voids, both governing the evolution of polarization, exist in piezoceramic materials. The total polarization can be partitioned into two parts, polarization P_d, which addresses the energy required to align the dipole orientation with the direction of electric field through 180 deg and 90 deg domain switching, and polarization P_{in}, which addresses the energy required to overcome the inclusions in piezoelectric materials. That is to say,

$$P = P_d + P_{in}. \quad (12)$$

Both of these polarizations, P_d and P_{in}, will be investigated next and the governing evolution equations will be derived.

In a piezoceramic body of volume V with zero initial polarization, a dipole is modeled by a volume V_p possessing a moment **p** with a fixed magnitude p. Random distribution of dipoles implies that each dipole can be mapped to a unit spherical surface with the moment direction along the direction $\mathbf{e}_r$ (Fig. 1) which guarantees that the summation of dipole moments vanishes. Therefore, a continuous distribution function ρ_s representing dipole moment density on the surface and polarization P_d can be defined as follows:

$$\rho_s = \frac{pV}{4\pi V_p}, \quad (13)$$

$$P_d = \oiint_S \rho_s \mathbf{e}_p \cdot \mathbf{e}\, ds \quad (14)$$

where $\mathbf{e}_p$ denotes the dipole moment direction during repolarization.

A critical electrical energy is required to reorientate the moment direction of the dipole through 180 deg or 90 deg domain switching. The density of this critical energy, ρ_c, can be approximated by $\rho_s E_c$. For the dipole at a location described by the angle φ (Fig. 1), the accumulated electrical energy density that makes a 180 deg flip possible is $\rho_s E \cos\varphi$. The accumulated electrical energy density that makes a 90 deg flip possible is $\rho_s E \sin\varphi$. If the electrical energy density associated with a 180 deg flip reaches the critical value ρ_c, the 180 deg flip will take place first. The same is assumed for the 90 deg flip.

Figure 2(*a*) illustrates the initial state of zero polarization. In the absence of an external electric field, the moment direction of each dipole is along $\mathbf{e}_r$. Figure 2(*b*) illustrates dipole direction variations due to the external electric field E applied along the direction **e**. The dipoles at the bottom surface of the sphere (π

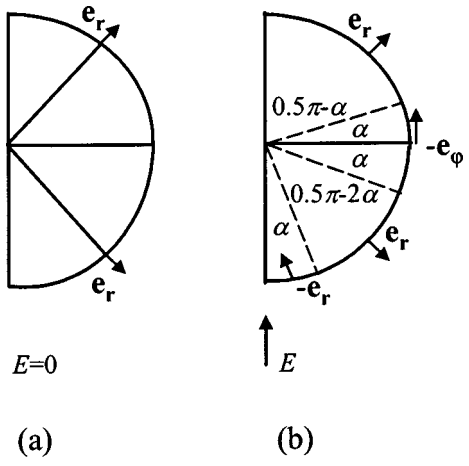


Fig. 2 Illustration of dipole reorientation

$-\alpha \leqslant \varphi \leqslant \pi$) change their direction from $\mathbf{e}_r$ to $-\mathbf{e}_r$ due to the fact that the 180 deg switching energy density, $\rho_s E\cos\varphi$, is greater than the critical value ρ_c. The dipoles at the middle of the sphere ($\pi/2-\alpha \leqslant \varphi \leqslant \pi/2+\alpha$) change their direction from $\mathbf{e}_r$ to $-\mathbf{e}_\varphi$ due to the fact that the 90 deg flip energy density, $\rho_s E\sin\varphi$, is greater than the critical value ρ_c, where α is used to denote the region of dipole switching. By using Eq. (14), the polarization P_d can be calculated as follows:

$$P_d=\rho_s\int_0^{2\pi}d\theta\left[\int_0^{\pi/2-\alpha}\mathbf{e}_r\cdot\mathbf{e}+\int_{\pi/2-\alpha}^{\pi/2+\alpha}-\mathbf{e}_\varphi\cdot\mathbf{e}+\int_{\pi/2+\alpha}^{\pi-\alpha}\mathbf{e}_r\cdot\mathbf{e}+\int_{\pi-\alpha}^{\pi}-\mathbf{e}_r\cdot\mathbf{e}\right]\sin\varphi\,d\varphi. \tag{15}$$

The complete process of repolarization due to a cyclic loading of electric field can be determined following the above procedure. The result is illustrated in Fig. 3. It can be observed that the polarization P_d reverses in the narrow regions around the coercive electric field. The variation of polarization can capture the basic characteristics of the hysteresis loop. Finally, the relation between the mapping values (ρ_s and α) and the piezoceramic characteristics (P_s and E_c) can also be obtained as follows:

$$\rho_s=\frac{P_s}{\pi(2+\pi/2)}, \tag{16}$$

$$\alpha=\arccos\left(\frac{E_c}{|E|}\right). \tag{17}$$

The energy required to overcome the inclusions in piezoceramic materials, addressed by the polarization P_{in}, is assumed proportional to the change in energy required to align the dipoles. The process of repolarization can be modeled such that a dipole is reorientated from the initial position (moment $\mathbf{p}'$ with an angle θ between the directions of the dipole moment and the electric field) to the aligned position (moment $\mathbf{p}$ in the direction of the electric field). The variation in energy density during this process, denoted $\Delta\rho_d$, can be written as follows:

$$\Delta\rho_d=\mathbf{p}\cdot\mathbf{E}-\mathbf{P}'\cdot\mathbf{E}=pE(1-\cos\theta). \tag{18}$$

Therefore, the energy density to overcome the inclusions, denoted ρ_{in}, can be written as follows:

$$\rho_{in}=cpE(1-\cos\theta) \tag{19}$$

where c is a proportionality constant.

Considering the volume of dV with dipole density n and using the average of statistical mechanics ([11]), the energy required to overcome the inclusions and the corresponding change in polarization, denoted ΔR_{in} and ΔP, can be written as follows:

$$\Delta R_{in}=\frac{1}{2}\langle\rho_\pi\rangle(1-\cos\theta)ndV, \tag{20}$$

$$\Delta P=pn(1-\cos\theta) \tag{21}$$

where $\langle\rho_\pi\rangle$ denotes the value of ρ_{in} when $\theta=\pi$ (180 deg flip), that is, $\langle 2cpE\rangle$. Therefore, the energy in the volume dV in terms of the polarization change can be written as follows:

$$\Delta R_{in}=\frac{\langle\rho_\pi\rangle}{2p}\Delta P\,dv. \tag{22}$$

Thus, during the process of repolarization starting from zero to the value P, the energy required to overcome inclusions, denoted R_{in}, can be written as follows:

$$R_{in}=\frac{\langle\rho_\pi\rangle}{2p}dv\int_0^P dP. \tag{23}$$

On the other hand, the quantity R_{in}, representing the energy associated with the polarization P_{in} in the volume dv, can be written in another form as follows:

$$R_{in}=-\int_0^E P_{in}dE\,dv. \tag{24}$$

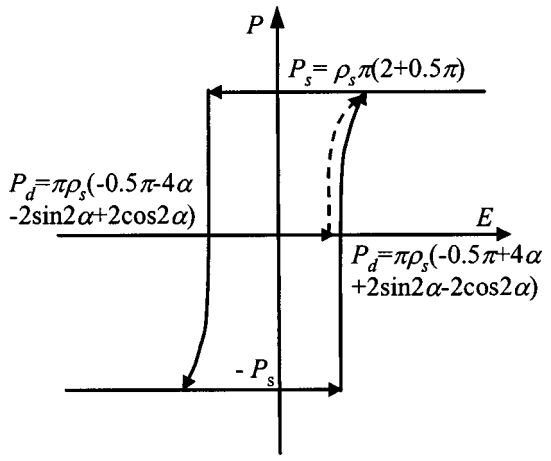


Fig. 3 Illustration of polarization P_d

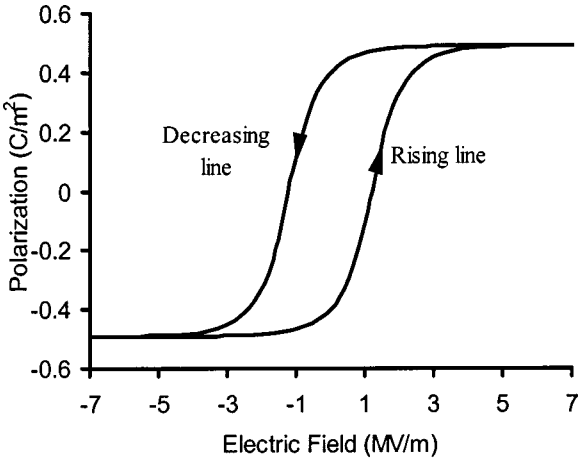


Fig. 4 Illustration of hysteresis loop (PZT 5A)

By equating and differentiating Eqs. (23) and (24) with respect to the electric field E and substituting Eq. (12) into it, the following incremental law governing the polarization development is obtained:

$$k\frac{dP}{dE}=P_d-P \tag{25}$$

where k is a material constant and is denoted $\langle\rho_\pi\rangle/2p$.

The polarization P_d (Fig. 3) shows that the flips take place in the small regions around the coercive value $\pm E_c$. Therefore, P_d can be approximated by two straight lines $\pm P_s$ with jumps at the points of the coercive values. By using this approximation, Eq. (25) can be solved with two explicit solutions, which represent the rising line and the decreasing line of the hysteresis loop (Fig. 4), respectively. The hysteresis loop can now be written as follows:

$$\begin{cases} P(E)=\text{sign}(E-E_c)P_s\left[1-\exp\left(\left|\dfrac{E}{E_c}-1\right|\ln\left(1-\dfrac{P_r}{P_s}\right)\right)\right] & \text{rising line} \\ P(E)=\text{sign}(E+E_c)P_s\left[1-\exp\left(\left|\dfrac{E}{E_c}+1\right|\ln\left(1-\dfrac{P_r}{P_s}\right)\right)\right] & \text{decreasing line} \end{cases} \tag{26}$$

or

$$E(P)=E_c\left[\delta+\text{sign}(P)\frac{\ln(1-|P/P_s|)}{\ln(1-|P_r/P_s|)}\right] \tag{27}$$

where $\delta=1$ represents the rising line and $\delta=-1$ represents the decreasing line.

Generally, piezoceramic material will not be driven to full saturation, which implies that the maximum polarization (P_s') during the process of repolarization is smaller than the saturation value of P_s. The magnitude of P_s' can be obtained through Eq. (26) by substituting the maximum applied electric field into it and the hysteresis relation can be written as follows:

$$\begin{cases} P(E)=P_s\left[1-\exp\left(\left|\dfrac{E}{E_c}-1\right|\ln\left(1-\dfrac{P_r}{P_s}\right)\right)\right] & E_c\leqslant E\leqslant E_{\max} \\ P(E)=-P_s'\left[1-\exp\left(\left|\dfrac{E}{E_c}-1\right|\ln\left(1-\dfrac{P_r}{P_s}\right)\right)\right] & -E_{\max}\leqslant E<E_c \end{cases} \quad \text{rising line} \tag{28}$$

and

$$\begin{cases} P(E)=P_s'\left[1-\exp\left(\left|\dfrac{E}{E_c}+1\right|\ln\left(1-\dfrac{P_r}{P_s}\right)\right)\right] & -E_c\leqslant E\leqslant E_{\max} \\ P(E)=-P_s\left[1-\exp\left(\left|\dfrac{E}{E_c}+1\right|\ln\left(1-\dfrac{P_r}{P_s}\right)\right)\right] & -E_{\max}\leqslant E<-E_c \end{cases} \quad \text{decreasing line} \tag{29}$$

where $E_{\max}$ is the maximum value of the applied electric field. For both rising and decreasing lines, if the directions of the polarization vector and the electric field remain the same ($E_c\leqslant E\leqslant E_{\max}$ for the rising line and $-E_{\max}\leqslant E<-E_c$ for the decreasing line), the equations or the corresponding paths in the loop will be the same as those driven to the saturation point P_s. The repolarization stops when the electric field reaches its maximum value $\pm E_{\max}$ and the pseudo-saturation points $\pm P_s'$ are reached. If the polarization and the electric field have opposite directions ($-E_{\max}\leqslant E\leqslant -E_c$ for the rising line and $-E_c\leqslant E<E_{\max}$ for the decreasing line), the paths will start from the pseudo-saturation points $\pm P_s'$ instead of the saturation points $\pm P_s$.

Elastic Gibbs Free Energy

The elastic Gibbs energy is assumed as follows:

$$G(\sigma_{ij},P_i)=-\tfrac{1}{2}S_{ijkl}\sigma_{ij}\sigma_{kl}-Q_{ijkl}\sigma_{ij}P_kP_l+F(P_i) \tag{30}$$

where the quantity S_{ijkl} denotes the elastic compliance at constant polarization. The quantity Q_{ijkl} denotes the material constants representing piezoelectric effects. The first term on the right-hand side of Eq. (30) represents the mechanical strain energy and the second term results in the electrical-mechanical coupling. The last term, $F(P_i)$, is used to address the hysteresis effects.

Substitution of Eq. (30) into Eqs. (4) and (5) yields the following constitutive relations:

$$\begin{cases} \varepsilon_{ij}=S_{ijkl}\sigma_{kl}+Q_{ijkl}P_kP_l \\ E_i=-2Q_{klil}\sigma_{kl}P_l+l_iE(P) \end{cases} \tag{31}$$

where $l_iE(P)$, representing the hysteresis relation between polarization and electric field in the absence of stresses, results from the derivative of $F(P_i)$ with respect to the polarization P_i. The quantity l_i denotes the direction of electric field and $E(P)$, the variation of electric field due to the change in polarization, has been defined in Eq. (27).

For piezoelectric materials, Q_{ijkl} can be expressed by two material constants, T_1 and T_2. Equation (31) can then be simplified to matrix form as follows:

$$\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6\end{bmatrix}=\begin{bmatrix}S_{11} & S_{12} & S_{13} & 0 & 0 & 0\\ S_{12} & S_{22} & S_{13} & 0 & 0 & 0\\ S_{13} & S_{13} & S_{33} & 0 & 0 & 0\\ 0 & 0 & 0 & S_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & S_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & S_{66}\end{bmatrix}\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{bmatrix}+\begin{bmatrix}T_1+T_2 & T_1 & T_1 & 0 & 0 & 0\\ T_1 & T_1+T_2 & T_1 & 0 & 0 & 0\\ T_1 & T_1 & T_1+T_2 & 0 & 0 & 0\\ 0 & 0 & 0 & 2T_2 & 0 & 0\\ 0 & 0 & 0 & 0 & 2T_2 & 0\\ 0 & 0 & 0 & 0 & 0 & 2T_2\end{bmatrix}\begin{bmatrix}P_1^2\\ P_2^2\\ P_3^2\\ P_2P_3\\ P_3P_1\\ P_1P_2\end{bmatrix} \tag{32}$$

and

$$\begin{bmatrix}E_1\\ E_2\\ E_3\end{bmatrix}=-2\begin{bmatrix}(T_1+T_2)\sigma_1P_1+T_1\sigma_2P_1+T_1\sigma_3P_1+T_2\sigma_6P_2+T_2\sigma_5P_3\\ T_2\sigma_6P_1+T_1\sigma_1P_2+(T_1+T_2)\sigma_2P_2+T_1\sigma_3P_2+T_2\sigma_4P_3\\ T_2\sigma_5P_1+T_2\sigma_4P_2+T_1\sigma_1P_3+T_1\sigma_2P_3+(T_1+T_2)\sigma_3P_3\end{bmatrix}+l_iE(P). \tag{33}$$

Note that induced strains from the linear constitutive relation can be written as follows:

$$\varepsilon_{ij}=S_{ijkl}\sigma_{kl}+d_{ijk}E_k. \tag{34}$$

The hysteresis relation and the linear constitutive relation, addressed by Eqs. (32) and (34), should predict the same induced strains if the applied electric field varies in the vicinity of zero value where the linear constitutive relation is effective. Therefore, the following relation is obtained:

$$\begin{bmatrix}0 & 0 & d_{113}\\ 0 & 0 & d_{113}\\ 0 & 0 & d_{333}\end{bmatrix}\begin{bmatrix}0\\ 0\\ E\end{bmatrix}=\begin{bmatrix}T_1+T_2 & T_1 & T_1\\ T_1 & T_1+T_2 & T_1\\ T_1 & T_1 & T_1+T_2\end{bmatrix}\begin{bmatrix}0\\ 0\\ P^2-P_r^2\end{bmatrix}. \tag{35}$$

Using Tylor expansion to approximate the square term in Eq. (35), the material electrical-mechanical coupling constants T_1 and T_2 can be expressed in terms of d_{ijk}, the piezoelectric coefficients in the linear constitutive relation.

$$\begin{cases}T_1=\dfrac{d_{113}E_c\lambda}{-2P_r(P_s-P_r)\ln(1-P_r/P_s)}\\ T_2=\dfrac{(d_{333}-d_{113})E_c\lambda}{-2P_r(P_s-P_r)\ln(1-P_r/P_s)}\end{cases} \tag{36}$$

where a scalar λ is introduced to obtain the best simulation for the entire process rather than limiting to the vicinity of zero electric field. This quantity should be determined specially for different materials.

Results and Discussions

The developed constitutive relation is used to model the hysteresis behavior of piezoceramics due to significant variations of electric field. The results obtained using the current model are correlated with available experimental results.

First, the hysteresis relation between polarization and electric field is investigated in Figs. 4–6. The prediction obtained from the current model is validated with experimental results ([11]). The solid line and dash line denote the results obtained from the current model and experiments, respectively. A cylinder PZT 5A wafer of diameter one inch and thickness ten mils is considered. The repolarization is generated under quasi-static operating conditions (200 mHz). The hysteresis characteristics are such that the saturation polarization $P_s=0.49\,\mathrm{C/m^2}$ and the coercive electric field $E_c=1.2\,\mathrm{MV/m}$. The remnant polarization of piezoceramics is approximated by the relation of $P_r=0.82P_s$ ([12]). Figure 4 presents the hysteresis loop between the polarization and the applied electric field which is high enough to make the material saturated. It is observed that the current model can capture the characteristics of hysteresis loop in piezoceramics. With increase in the electric field, the polarization direction varies from negative to positive along the rising line due to dipole reorientation. Once the saturation point is reached, the polarization decreases along the decreasing line with decrease in the electric field. The hysteresis loop is also investigated if the material is not fully saturated. The hysteresis loops due to electric potentials with the maximum values of 1600 V and 1000 V, that is, electric fields of 6.3 MV/m and 3.93 MV/m, are presented in Figs. 5 and 6, respectively. The maximum polarizations are smaller compared to the fully satu-

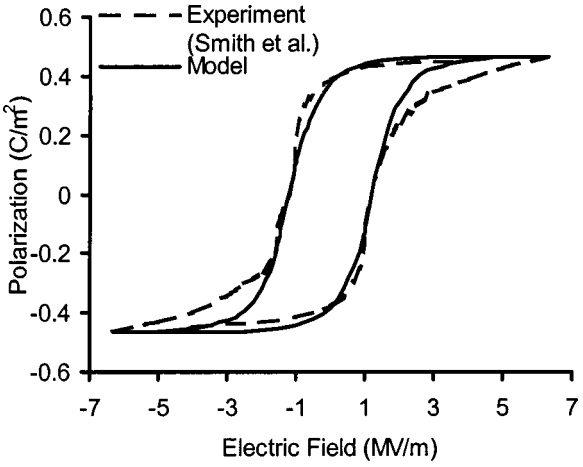


Fig. 5 Comparison of electrical hysteresis loops (PZT 5A, 1600 V)

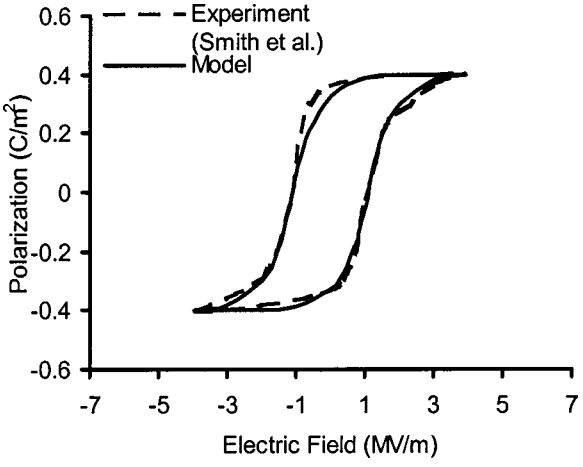


Fig. 6 Comparison of electrical hysteresis loops (PZT 5A, 1000 V)

Table 1 Material properties of soft PZT-51 ceramics

Elastic moduli	$S_{3333}=3.03\times10^{-11}$ m^2/N
	$S_{3311}=2.9\times10^{-11}$ m^2/N
Piezoelectric coefficient	$d_{333}=1520$ pC/N
	$d_{311}=570$ pC/N
Relative dielectric permittivity	$\gamma=11300$
Coercive electric field	$E_c=0.676$ MV/m
Remnant polarization	$P_r=0.1938$ C/m^2

rated case. A good agreement between predictions using the current model and experimental data is observed. This indicates that the physical phenomena with different scales of driving electric fields are modeled accurately by the present analytical model.

Further investigation is performed with soft PZT-51 ceramics including the hysteresis loop between the electric displacement and electric field and the butterfly loop between the induced strain and electric field. The material properties of soft PZT-51 ceramics are listed in Table 1. The bulk specimen is used with the dimensions $10\times10\times16$ mm. The piezoceramic is driven by an electric field up to a value of 1.125 MV/m along the longitudinal direction. The correlations between the results obtained using the current model and from available experimental data ([12]), with zero and a constant compressive stress along the direction of the electric field, are presented in Figs. 7–10. Figure 7 presents the hysteresis loop between the electric displacement and electric field under a zero stress condition ($\sigma_3=0$ Mpa). Linear behavior is observed with a small value of the electric field. This leads to the linear relation between the electric displacement and the electric field addressed by the permittivity constant, representing the slope of the zero electric field point in the curve. With increase in the applied electric field, the nonlinear effect due to the variation of the polarization resulting from significant 90 deg domain switching is observed. If the electric field is applied opposite to the direction of the polarization, both 180 deg and 90 deg domain switching will occur when the applied electric field approaches the coercive value. This leads to the reversal of the polarization direction. Good agreement is observed between the developed model and the experimental results. The hysteresis loop with compressive stress $\sigma_3=-20$ Mpa is presented in Fig. 8. Compared to the case with zero stress, there is a significant decrease in both the remnant value and the saturated value of the electric displacement. This is due to the fact that only depolarization (90 deg domain switching) results from compressive stress, which leads to the decrease in the polarization value. Again, good agreement is observed between the theory and the experimental results. The butterfly loops representing variations of induced longitudinal strain with change in the electric field, under zero stress and a compressive stress $\sigma_3=-20$ Mpa, are presented in Figs. 9 and 10, respectively. The material constant λ is assumed to be 2.4 to provide the best approximation of the entire loop. Note that a zero strain is assumed for the initial state of the zero electric field. Figure 9 shows that the current model is capable of capturing the charac-

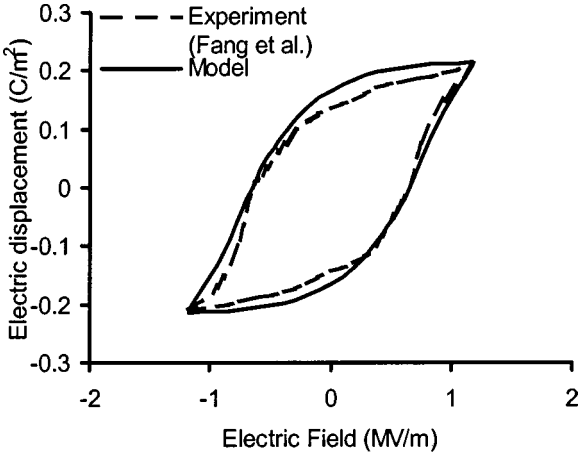


Fig. 8 Comparison of electrical hysteresis loop (PZT-51, $\sigma_3=-20$ Mpa)

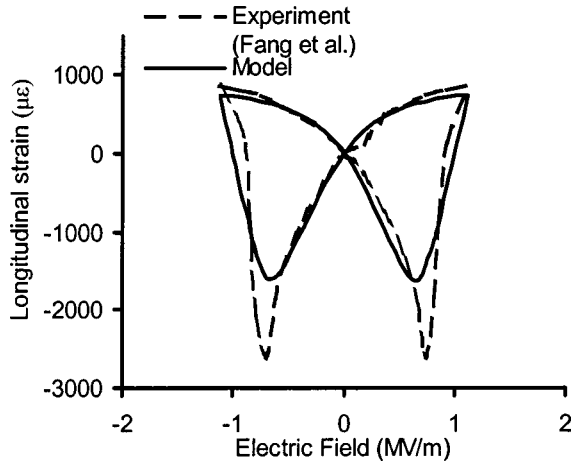


Fig. 9 Comparison of longitudinal strain (PZT-51, $\sigma_3=0$ Mpa)

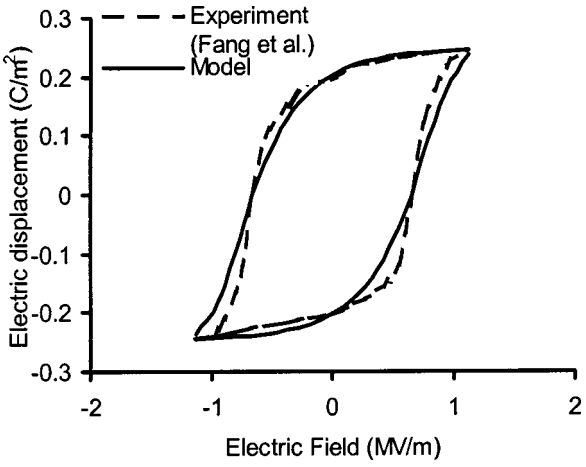


Fig. 7 Comparison of electrical hysteresis loop (PZT-51, $\sigma_3=0$ Mpa)

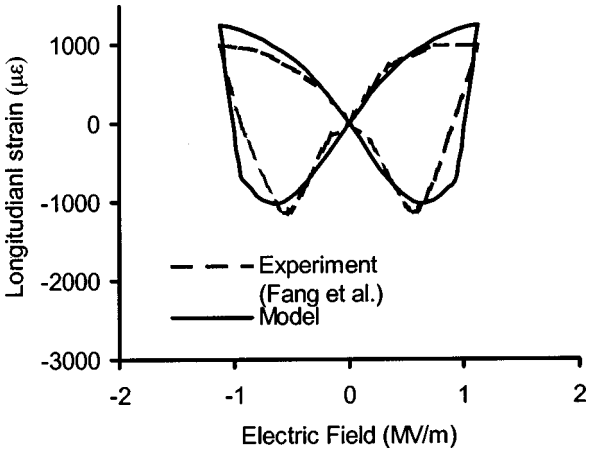


Fig. 10 Comparison of longitudinal strain (PZT-51, $\sigma_3=-20$ Mpa)

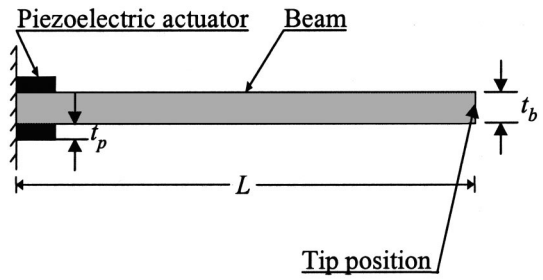


Fig. 11 Illustration of cantilever beam, surface-bonded actuators

teristics of the butterfly hysteresis of piezoceramics. The linear relation between the induced strain and the electric field is valid for small values of the applied electric field. This is predicted by the material coupling constant d_{333} in the linear constitutive relation. With increase in the applied electric field, the induced strain increases and nonlinearity is observed due to variation in the magnitude of the polarization. With further increase in the electric field up to the coercive value, the trend reverses due to the change in the polarization direction. It can also be observed from Fig. 9 that in the case of zero compressive stress, hysteresis is accurately modeled except in the small region where polarization approaches zero value, that is, the region where polarization reverses its direction. One reason for the misprediction is the shift in coercive electric field during the process of electrical loading, which is not considered in the current model. Figure 10 presents the hysteresis under a constant compressive stress ($\sigma_3 = -20$ Mpa). It can be observed that the results obtained from the current model and experiments show good agreement. The range of the strain variation is smaller compared to the case without stress. This results from the effect of depolarization due to the compressive stress.

As shown in Fig. 11, a thin aluminum cantilever beam, with top and bottom surface-bonded piezoelectric actuators at the root, is considered to investigate the nonlinear actuation effects. The beam dimensions are length $L = 0.61$ m, width $b = 0.061$ m and thickness $t_b = 8 \times 10^{-4}$ m. The Young's modulus is 7.235×10^{10} N/m^2. The structure is modeled using the Bernoulli-Euler beam theory. The piezoelectric actuators (PZT-51 ceramics) are of length $L/12$, width b and thickness $t_p = 3.175 \times 10^{-4}$ m. The polarization directions of the top and the bottom actuators are assumed upwards and downwards, respectively. First, actuation effect without change in the polarization direction is investigated in Fig. 12. The electric field, varying from 0 V to 500 V, is applied upwards on only the top actuator. This implies no reversal in

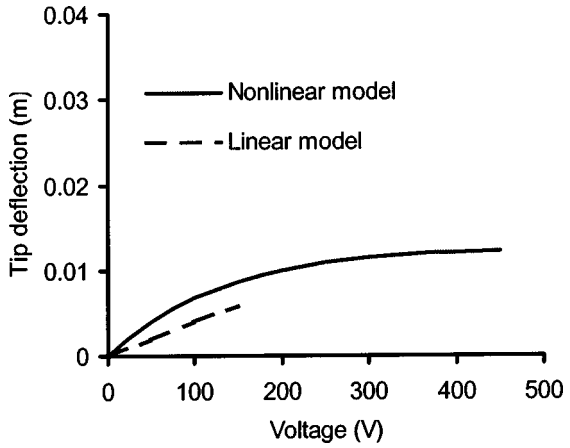


Fig. 12 Comparison of linear and nonlinear induced deflections

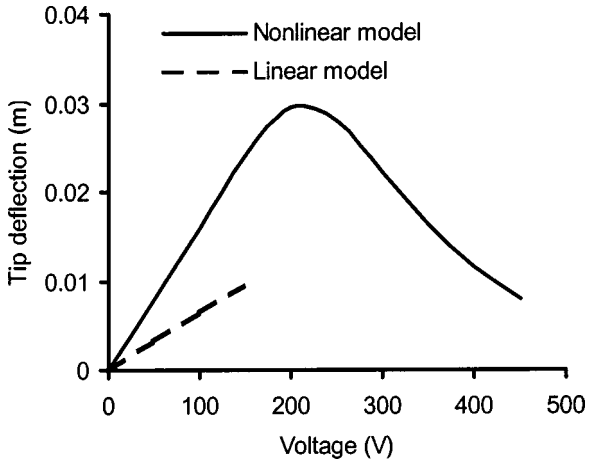


Fig. 13 Comparison of linear and nonlinear induced deflections

polarization. Figure 12 presents comparison of the tip deflections with variation in actuation voltage using the current model and the linear constitutive model. Based on the result obtained using the current model, it can be observed that initially the deformation increases rapidly with increase in the applied voltage due to the large variation in polarization. With further increase in actuation, polarization approaches the saturation value. Therefore, variations in both polarization and deformation are small. The significant difference between the linear model and the present nonlinear model is due to the fact that the linear model neglects the evolution of polarization. It can also be observed that the linear model underpredicts the authority of shape correction ($0 \sim 150$ V), which was pointed out by Crawley et al. [13]. Next, actuation effect with a change in the polarization direction is investigated in Fig. 13. The electric field is applied upwards on both actuators to produce the bending moment. Again, variation of the tip deflection due to different actuation, using the current model and the linear model, is presented. It can be observed that initially the current model predicts the linear relation between the induced deflection and the applied electric field. This is due to the fact that before the electric field approaches the coercive value, polarization in the top actuator increases and polarization in the bottom actuator decreases. The combination of these two effects results in linear variation between deformation and actuation. With further increase in the electric field, up to the value approaching a coercive electric field (200 V), the induced deflection produces significant nonlinearity due to the large variation in polarization of the bottom actuator. The polarization of the bottom actuator reverses, aligning with that of the top actuator. Therefore, the bending deformation becomes insignificant, which leads to zero deflection if both actuators reach the point of saturation.

Concluding Remarks

A new theory is developed to model the hysteresis relation between polarization and the electric field of piezoceramics. An explicit approximate solution is obtained to illustrate the hysteresis loop by using saturation polarization, remnant polarization, and coercive electric field. The elastic Gibbs energy is proposed to address the coupling relations between an electrical field and a mechanical field. The nonlinear constitutive relations are derived from the elastic Gibbs energy. The hysteresis relations between polarization, mechanical strain, and an electric field using the current model are correlated with those obtained from available experiments. The constitutive relation derived from the current model is implemented in the analysis of a cantilever beam with surface-bonded piezoelectric actuators. The following important observations are made from this study.

1 The developed model provides a good physical means to accurately capture the hysteresis behavior of piezoceramic repolarization through the use of material parameters of saturation polarization, remnant polarization, and coercive electric field.

2 The new form of elastic Gibbs energy, based on the state variables of stress and polarization, can accurately predict the nonlinearity of butterfly hysteresis. This leads to the constitutive relation and the solution of the coupling coefficients in terms of linear piezoelectric constants.

3 For the case of a cantilever beam with surface-bonded actuators the linear constitutive relation mispredicts actuation authority. Significant nonlinearity is observed if the applied electric field approaches the coercive value. This phenomenon can be addressed by the developed hysteresis model.

Acknowledgment

The research is supported by NASA Langley Research Center, grant number NAG-1-1988, technical monitor Carol D. Wieseman.

References

[1] Crawley, E. F., and Deluis, J., 1987, ''Use of Piezoelectric Actuators as Elements of Intelligent Structures,'' AIAA J., **25**, No. 10, pp. 1373–1385.

[2] Eringen, A. C., 1963, ''On the Foundations of Electroelastrostatics,'' Int. J. Eng. Sci., **1**, pp. 127–153.

[3] Tiersten, H. F., 1969, *Linear Piezoelectric Plate Vibrations*, Plenum Press, New York.

[4] Chan, K., and Hagood, N., 1994, ''Modeling of Nonlinear Piezoceramics for Structural Actuation,'' *Proc. of SPIE's Symposium on Smart Structures and Materials*, Vol. 2190, pp. 194–205.

[5] Sirohi, J., and Chopra, I., 1998, ''Fundamental Behavior of Piezoceramic Sheet Actuators,'' *Proc. of SPIE's Symposium on Smart Structures and Integrated Systems*, Vol. 3329, pp. 626–646.

[6] Devonshire, A. F., 1992, ''Theory of Barium Titanate (Part I),'' *Key Papers in Physics: Piezoelectricity*, American Institute of Physics, Melvilla, NY, pp. 116–139.

[7] Devonshire, A. F., 1992, ''Theory of Barium Titanate (Part II),'' *Key Papers in Physics: Piezoelectricity*, American Institute of Physics, Melville, NY, pp. 140–154.

[8] Chowdhury, K. L., Epstein, M., and Glockner, P. G., 1979, ''On the Thermodynamics of Non-Linear Elastic Dielectrics,'' Int. J. Non-Linear Mech., **13**, pp. 311–322.

[9] Chen, P. J., and Montgomery, S. P., 1980, ''A Macroscopic Theory for the Existence of the Hysteresis and Butterfly Loops in Ferroelectricity,'' Ferroelectrics, **23**, pp. 199–208.

[10] Ghandi, K., and Hagood, N., 1996, ''Nonlinear Finite Element Modeling of Phase Transition in Electro-Mechanically Coupled Material,'' *Proc. of SPIE's Symposium on Smart Structures and Materials*, Vol. 2715, pp. 121–140.

[11] Smith, R. C., and Ounaies, Z., 1999, ''A Hysteresis Model for Piezoceramic Materials,'' CRSC Technical Reports, CRSC-TR99-21.

[12] Fang, D. N., and Li, C. Q., 1999, ''Nonlinear Electric-Mechanical Behavior of a Soft PZT-51 Ferroelectric Ceramic,'' J. Mater. Sci., **34**, No. 16, pp. 4001–4010.

[13] Crawley, E. F., and Anderson, E. H., 1989, ''Detailed Models of Piezoceramic Actuation of Beams,'' J. Intell. Mater. Syst. Struct., **1**, pp. 4–25.

Rheological Behavior of Confined Fluids in Thin Lubricated Contacts

J. Tichy
Department of Mechanical Engineering,
Aeronautical Engineering and Mechanics,
Rensselaer Polytechnic Institute,
Troy, NY 12180-3580

Continuum based methods are traditionally thought to be of little value in describing boundary lubrication, or the mode of lubrication in molecular scale films that may occur at asperity interactions during the sliding of nominally flat surfaces. There is considerable experimental evidence, which suggests that the classical theory may be valid with modification to films as thin as several nanometers. In addition, lubricants, which exhibit viscous liquid properties in bulk, may form attached solid-like elastic layers when confined between solid surfaces. In the present paper, the simple ''elastic foundation'' concept is used to model the elastic layers, in contact with a viscous fluid film. Several typical bearing contact flow problems are solved, giving hope that boundary lubrication may eventually be modeled in the same manner as hydrodynamic lubrication in thicker films.
[DOI: 10.1115/1.1354204]

Introduction

It has become fairly well accepted based on wide-ranging experimental and theoretical evidence that confined liquids in tribological contacts exhibit complex rheological behavior. In particular, the presence of a solid-like confined layer adhering to the confining surfaces and roughly one molecular length in extent is postulated. Surface force experiments by Israelachvili [1] and colleagues, Gee et al. [2], Homola et al. [3] on a similar apparatus of Tonck et al. [4] and Georges et al. [5] all point to the existence of solid-like structures in confined films of materials that exhibit purely viscous liquid properties in bulk. Such properties are also thought to contribute to an intrinsic (i.e., not caused by roughness, capillarity at asperity junctions, and other factors) stick-slip behavior of some lubricated contacts (Yoshizawa and Israelachvili [6]).

The most convincing experiment as to the presence of solid layers is that of Chan and Horn [7]. The slow squeezing of a thin film between two crossed molecularly smooth mica cylinders is performed. The authors find that the squeeze rate is well predicted by the classical Reynolds equation of hydrodynamic lubrication theory (see, e.g., Hamrock [8]) down to about 30 nm. Reynolds equation is an integrated form of the Navier-Stokes equation for thin films of low curvature, without the presence of inertia or body forces. At thinner gaps, good correlation with experiment is obtained by simply adding a fictitious rigid layer of 0.7 nm to the mica surfaces. Such a layer is really a curve fit parameter in their analysis, but can be thought of as representing a solid layer due to lubricant microstructure. This approach predicts the trajectory remarkably well (± 2 percent) down to films of about 2 nm, after which further squeezing occurs in steps. The molecular size is about 1 nm, and the continuum Reynolds theory works well on a scale much smaller than one might anticipate.

Molecular dynamics simulations have added further credence to the existence of solid-like layers in thin fluid films, and to predictive ability of continuum models in films spanned by relatively few molecular lengths. Such studies are due to Thompson and Robbins [9], Thompson et al. [10], Hu et al. [11], Landman et al. [12], and many others.

The present author has modeled the surface structures as a porous medium [13] and as a highly viscous layer [14]. Auslander and Sidoroff [15] have modeled the layer as an elastic layer, using what they called a ''thin film approximation,'' much like such fluid mechanical assumptions, but nothing like the thin film approximation of solid coatings. (They use a hydrostatic stress as the normal components of the stress tensor.) In any case, they predict an effective viscoelastic material response for the film, consistent with the experiments referred to above of Tonck et al. [4] and Georges et al. [5]. They only apply their model to the simple squeezing case. In the present case, we extend the Auslender and Sidoroff work in the sense of using an elastic film, but we use the ''elastic foundation'' or Winkler or ''mattress'' model discussed by Johnson [16]. We are able to arrive at a much more general formulation, a modified Reynolds equation, and solve a number of ''classical'' lubricated contact problems.

Analysis

Consider two rigid solid surfaces separated by a thin film of thickness $H(x,y,t)$. To model the effect of solid-lake layers we assume that elastic layers of thickness $\delta(x,y,t)$ adhere to the rigid surfaces, and the fluid region is given by

$$h(x,y,t) = H - 2\,\delta. \tag{1}$$

The rigid surface separation H and the undeformed elastic layer thickness δ_i are assumed to be imposed on the problem, while the elastic layer δ may be compressed according to the pressure in the film (see Fig. 1). The elastic deformation is assumed to obey the simplified Winkler or ''mattress'' model, discussed by Johnson [16], in which the displacement u_z at any point depends only on the pressure p at that point:

$$u_z(x,y,t) = \delta_i - \delta = \delta_i \frac{p(x,y,t)}{K} \tag{2}$$

and K is the elastic modulus. Reynolds equation holds for the fluid region:

$$\frac{\partial}{\partial x}\left(\frac{h^3}{12\mu}\frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{h^3}{12\mu}\frac{\partial p}{\partial y}\right) = \frac{1}{2}U\frac{\partial h}{\partial x} + \frac{\partial h}{\partial t} \tag{3}$$

assuming sliding at velocity U occurs only in the x-direction. However, this equation is now strongly nonlinear due to the coupling between film thickness h and p:

$$h = h_i + 2\,\delta_i \frac{p}{K}, \quad h_i(x,y,t) = H - 2\,\delta_i, \tag{4}$$

where h_i is also treated as known or imposed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 2, 1997; final revision, May 3, 1999. Associate Technical Editor: D. A. Siginer. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

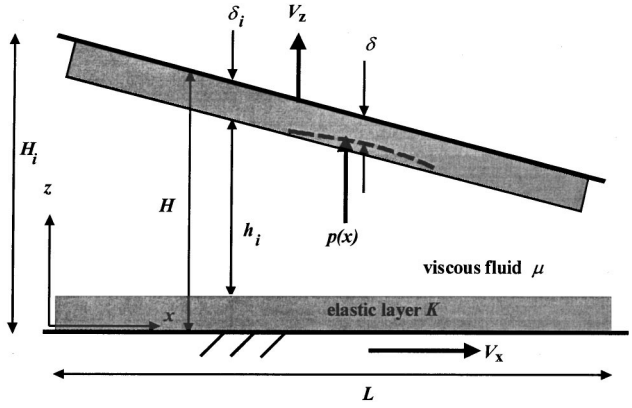


Fig. 1 Schematic of contact

The velocity and shear stress in the flow region ($\delta \leqslant z \leqslant H-\delta$) are given by

$$v_x = \frac{1}{2\mu}\frac{\partial p}{\partial x}[(z-\delta)^2-(z-\delta)h]+U\left[1-\frac{z-\delta}{h}\right], \tag{5}$$

$$v_y = \frac{1}{2\mu}\frac{\partial p}{\partial y}[(z-\delta)^2-(z-\delta)h]$$

$$\tau_{zx}(z=\delta)=\mu\left.\frac{\partial v_x}{\partial z}\right|_{z=\delta}=-\frac{h}{2}\frac{\partial p}{\partial x}-\frac{U}{h}, \tag{6}$$

$$\tau_{zy}(z=\delta)=\mu\left.\frac{\partial v_y}{\partial z}\right|_{z=\delta}=-\frac{h}{2}\frac{\partial p}{\partial x}.$$

The forces on the bearing surfaces are given by

$$F_x=\int_0^L\int_0^W \tau_{zx}(z=\delta)dydx, \quad F_y=\int_0^L\int_0^W \tau_{zy}(z=\delta)dydx, \tag{7}$$

$$F_z=\int_0^L\int_0^W (p-p_a)dydx,$$

(p_a is the ambient pressure) where F_x and F_y are friction or traction forces, and F_z is the supported load.

We nondimensionalize the above equations using

$$x^*=\frac{x}{L} \quad y^*=\frac{y}{L} \quad \delta^*=\frac{\delta}{H_i} \quad h^*=\frac{h}{H_i}, \quad h_i^*=H^*-2\delta_i^* \tag{8}$$

$$p^*=\frac{p-p_a}{\mu V_{xi}L^2/H_i^3} \quad p^*=\frac{p-p_a}{\mu V_{xi}L/H_i^2} \quad V_x^*=\frac{V_x}{V_{xi}} \quad t^*=\frac{t}{L/V_{xi}}$$

where V_{xi} and H_i are a reference sliding velocity and film thickness, respectively. We assume a perturbation solution of the form

$$h^*=h_i^*+2\kappa^*\delta_i^*p^* \quad p^*=p_0^*+\kappa^*\delta_i^*p_i^*+\ldots \quad \kappa^*=\frac{\mu V_{xi}L}{KH_i^2} \tag{9}$$

We obtain a zero-order Reynolds equation (the conventional case)

$$\frac{\partial}{\partial x^*}\left(\frac{h_i^{*3}}{12}\frac{\partial p_0^*}{\partial x^*}\right)+\frac{\partial}{\partial y^*}\left(\frac{h_i^{*3}}{12}\frac{\partial p_0^*}{\partial y^*}\right)=\frac{1}{2}V_x^*\frac{\partial h_i^*}{\partial x^*}+\frac{\partial h_i^*}{\partial t^*}, \tag{10}$$

and a first-order perturbation modified Reynolds equation

$$\frac{\partial}{\partial x^*}\left(\frac{h_i^{*3}}{12}\frac{\partial p_i^*}{\partial x^*}\right)+\frac{\partial}{\partial y^*}\left(\frac{h_i^{*3}}{12}\frac{\partial p_i^*}{\partial y^*}\right)$$

$$=V_x^*\frac{\partial p_0^*}{\partial x^*}+2\frac{\partial p_0^*}{\partial t^*}-h_i^*p_0^*\frac{\partial h_i^*}{\partial x^*}\frac{\partial p_0^*}{\partial x^*}$$

$$-\frac{1}{2}h_i^{*2}\left[\left(\frac{\partial p_0^*}{\partial x^*}\right)^2+\left(\frac{\partial p_0^*}{\partial y^*}\right)^2+p_0^*\frac{\partial^2 p_0^*}{\partial x^{*2}}+p_0^*\frac{\partial^2 p_0^*}{\partial y^{*2}}\right]. \tag{11}$$

The parameter κ^* represents a dimensionless elastic compliance of the layer, the ratio of the order of magnitude of the lubrication theory pressure $\mu V_{xi}L/H_i^2$ to the layer elastic modulus. In physical terms, the zero-order perturbation solution p_0^* with $h_i^*=H^*$ represents the pressure for a Newtonian fluid film. The zero-order solution with $h_i^*=H^*-2\delta_i^*$ represents the solution for a film consisting of a Newtonian fluid and the surfaces covered by rigid layers of thickness δ. The first-order perturbation solution p_1^* is a correction to the rigid layer solution for the elasticity of the layers.

The One-Dimensional Parallel Surface Squeeze Film. Consider the case of two infinite parallel plates undergoing pure squeezing: $H^*=H^*(t^*)$ and $V_x^*=0$. See Fig. 1 for the case when the bearing slope is zero and the origin $x=0$ is at the midpoint. The reference velocity in the nondimensionalization of Eq. (8) is given by $V_{xi}=V_z(L/H)$.

In this case, Eq. (10) becomes

$$\frac{\partial^2 p_0^*}{\partial x^{*2}}=12\frac{1}{h_i^{*3}}\frac{dh_i^*}{dt^*}, \quad x^*=\pm\frac{1}{2}:p_0^*=0. \tag{12}$$

Integrating twice and applying boundary conditions at the film ends, we obtain

$$p_0^*=-6\frac{\dot{H}^*}{h_i^{*3}}\left(\frac{1}{4}-x^{*2}\right), \quad \dot{H}^*=\frac{\partial H^*}{\partial t^*}. \tag{13}$$

The first-order case is as follows:

$$\frac{\partial^2 p_1^*}{\partial x^{*2}}=12\frac{1}{h_i^{*3}}\left[2\frac{\partial p_0^*}{\partial t^*}-\frac{1}{2}h_i^{*2}\left(\frac{\partial p_0^*}{\partial x^*}\right)^2-\frac{1}{2}h_i^{*2}p_0^*\frac{\partial^2 p_0^*}{\partial x^{*2}}\right], \tag{14}$$

$$x^*=\pm\frac{1}{2}:p_1^*=0.$$

Again, integrating twice and applying the boundary conditions, we obtain for the first-order solution

$$p_1^*=36\frac{1}{h_i^{*7}}\left[\dot{H}^{*2}\left(-\frac{1}{2}+3x^{*2}-4x^{*4}\right)\right.$$

$$\left.+\ddot{H}^*h_i^*\left(\frac{5}{48}-\frac{1}{2}x^{*2}+\frac{1}{3}x^{*4}\right)\right]. \tag{15}$$

The dimensionless load, from Eq. (7), can be expressed in the following form:

$$F_z^*=\frac{F_z/W}{\mu V_{xi}L^2/H_i^2}$$

$$=2\int_0^{1/2}p^*dx^*=-\frac{\dot{H}^*}{h_i^{*3}}+\kappa^*\delta_i^*\frac{1}{h_i^{*7}}\left(-\frac{54}{5}\dot{H}^{*2}+\frac{12}{5}\ddot{H}^*h_i^*\right). \tag{16}$$

Let us consider now a case similar to that considered by Auslander and Sideroff [15] with small rapid oscillations ω of the surface superimposed on a slow steady rate of descent:

$$\omega^*=\frac{\omega L}{V_{xi}},\quad H^*=1-V_z^*\frac{L}{H_i}t^*+\alpha^*\sin(\omega^* t^*)=h_i^*+2\delta_i^*,\qquad \alpha^*=\frac{\alpha}{H_i}. \tag{17}$$

Performing the indicated steps and taking the limit for small $V_z^* L/H_i$ and then small α, the zero-order solution for the load is

$$F_{z-01}^*=-\alpha^*\omega^*\left(\cos(\omega^* t^*)+\frac{12}{5}\mu^*\delta_i^*\omega^*\sin(\omega^* t^*)\right). \tag{18}$$

The 01 subscript denotes zero-order solution for $V_z^* L/H_i$ and the first-order solution for α^*. If the film was considered to be a homogeneous linear viscoelastic material, the components of complex viscosity would be

$$\eta'=\mu \quad \text{and} \quad \eta''=\mu\omega^*\kappa^*\delta_i^*=\mu\frac{\delta_i}{H_i}\frac{L^3}{H_i^3}\frac{\omega\mu}{K}. \tag{19}$$

Thus, we see classical linear viscoelasticity exhibited. The equivalent relaxation time λ and Deborah number De would be

$$\lambda=\frac{12}{5}\frac{\delta_i}{H_i}\frac{L^3}{H_i^3}\frac{\mu}{K},\quad \mathrm{De}=\lambda\omega. \tag{20}$$

Due to the symmetry about the film center ($x=0$), the global friction force F_x equals zero.

The One-Dimensional Steady Wedge Contact. The dimensionless film thickness is, again referring to Fig. 1,

$$H^*=1-mx^*=h_i^*+2\delta_i^*,\quad m=\frac{H(0)-H(L)}{H(0)} \tag{21}$$

and for $p^*(x^*)$ the zero-order solution is

$$p_0^*=\frac{6}{m}\left\{h_i^*(x^*)^{-1}-h_i^*(0)^{-1}-\frac{h_i^*(0)h_i^*(1)}{h_i^*(0)+h_i^*(1)}[h_i^*(x^*)^{-2}-h_i^*(0)^{-2}]\right\}. \tag{22}$$

The first-order modified Reynolds equation takes the form

$$\frac{\partial}{\partial x^*}\left(\frac{h_i^{*3}}{12}\frac{\partial p_1^*}{\partial x^*}\right)=\frac{\partial p_0^*}{\partial x^*}-h_i^* p_0^*\frac{\partial h_i^*}{\partial x^*}\frac{\partial p_0^*}{\partial x^*}-\frac{1}{2}h_i^{*2}\left[\left(\frac{\partial p_0^*}{\partial x^*}\right)^2+p_0^*\frac{\partial^2 p_0^*}{\partial x^{*2}}\right]=\frac{144m}{h_i^{*4}(h_i^*(0)+h_i^*(1))^2}[a_0+a_1x^*+a_2x^{*2}], \tag{23}$$

$$x^*=0,\ 1{:}p_i^*=0$$

$$a_0=h_i^*(0)^2h_i^*(1)\quad a_1=-2h_i^*(0)[3m^2-5mh_i^*(0)+h_i^*(0)^2] \tag{24}$$

$$a_3=m[-m^2+7mh_i^*(0)-7h_i^*(0)^2].$$

Again, integrating twice and applying the boundary conditions, we obtain for the first-order solution

$$p_i^*=b_2[h_i^*(x^*)^{-2}-h_i^*(0)^{-2}]+b_3[h_i^*(x^*)^{-3}-h_i^*(0)^{-3}]+b_4[h_i^*(x^*)^{-4}-h_i^*(0)^{-4}]+b_5[h_i^*(x^*)^{-5}-h_i^*(0)^{-5}] \tag{25}$$

$$b_2=\frac{48[m^4+5m^3h_i^*(0)-35m^2h_i^*(0)^2+60mh_i^*(0)^3-30h_i^*(0)^4]}{5m^2h_i^*(1)h_i^*(1)(h_i^*(0)+h_i^*(1))^3}$$

$$b_3=-\frac{48[7h_i^*(0)^2-7mh_i^*(0)+m^2]}{m^2(h_i^*(0)+h_i^*(1))^2},\qquad b_4=\frac{-144h_i^*(1)h_i^*(0)}{m^2(h_i^*(0)+h_i^*(1))},\quad b_5=\frac{-432h_i^*(1)^2h_i^*(0)^2}{5m^2(h_i^*(0)+h_i^*(1))^2}. \tag{26}$$

Dimensionless load and friction are computed from the following formulas:

$$F_z^*=\int_0^1 p^*dx^*,\quad F_x^*=\int_0^1\left(h^*\frac{\partial p^*}{\partial x^*}+\frac{1}{h^*}\right)dx^*. \tag{27}$$

These integrals can be evaluated in closed form for the wedge case, but the expressions are not shown here.

The Journal Bearing Contact. Film thickness expressions are given below.

$$h=c-2\delta_i+e\cos\theta,\quad h_i^*=1+\epsilon_i\cos\theta,\qquad \epsilon_i=\frac{\epsilon}{1-2\delta_i^*},\quad \epsilon=\frac{e}{c}. \tag{27a}$$

The radial clearance c equals $R_{\text{outer}}-R_{\text{inner}}\ll R_{\text{outer}}$, and the circumferential coordinate θ plays the role of x^* where $L=2\pi R_{\text{outer}}$. In dimensionless form, the layer thickness has been folded into the eccentricity ratio, to take advantage of the conventional journal bearing integrals, which are expressed in terms of $1+\epsilon\cos\theta$. The zero-order solution is

$$p_0^*=\frac{6\epsilon_i(2+\epsilon_i\cos\theta)\sin\theta}{(2+\epsilon_i^2)h_i^{*2}} \tag{28}$$

and, as above, the first-order differential equation looks like

$$\frac{\partial}{\partial x^*}\left(h_i^{*3}\frac{\partial p_1^*}{\partial x^*}\right)=\frac{1}{(h_i^*(0)+h_i^*(1))h_i^{*4}}\times(c_0+c_1\cos\theta+c_2\cos 2\theta+c_3\cos 3\theta) \tag{29}$$

$$c_0=-6\epsilon_i^2(-8+34\epsilon_i^2+\epsilon_i^4),$$

$$c_1=-3\epsilon_i(-8+49\epsilon_i^2+31\epsilon_i^4), \tag{30}$$

$$c_2=-6\epsilon_i^2(10-2\epsilon_i^2+\epsilon_i^4),\quad c_3=9\epsilon_i^3(-1+\epsilon_i^2).$$

Integrating twice and applying the boundary conditions that $p_1^*(0)=0$ and $p_1^*(\pi^+)=p_1^*(\pi^-)$ gives

$$p_1^*=\frac{12}{5(2+\epsilon_i^2)^2}\times\left(\frac{d_0+d_1+d_2}{h_i^{*5}}-\frac{11-76\epsilon_i-80\epsilon_i^2-20\epsilon_i^3}{(1+\epsilon_i)^4}\right),$$

$$d_0=11-136\epsilon_i^2-10\epsilon_i^4,\quad d_1=-5\epsilon_i(13+20\epsilon_i^2),\quad d_2=10\epsilon_i^2(2+\epsilon_i^2). \tag{30a}$$

Results and Discussion

Results for the case of simple squeezing flow between parallel plates are shown in Figs. 2–5. The solution consists of the zero-order solution, which treats the confined layers as a rigid solid. The first-order solution, shown in Figs. 2 and 3, shows a reduction of pressure (negative correction) due to the elastic softening of the layers. The pressure correction for steady squeezing ($\dot{H}^*=-1$, $\ddot{H}^*=0$) is reduced by acceleration ($\dot{H}^*=-1$, $\ddot{H}^*=1$) of the surface, as shown in Fig. 2. The effect of the relative thickness of the layers is strong. A change from only $\delta_i^*=0.1$ to $\delta_i^*=0.15$ causes the compliance effect to increase manifold (see Fig. 3).

The complete pressure field is shown in Fig. 4. Three cases are portrayed. The case of a purely Newtonian film with no layer yields the lowest pressure ($\delta_i^*=0$) in which case the acceleration H^* doesn't matter and the compliance parameter κ^* is not defined. The case of rigid layers ($\kappa^*=0$, $\delta_i^*=0.1$) produces much larger pressure. Compliant layers reduce the pressure considerably, with and without acceleration. Compliance also adds a time-dependent viscoelastic effect (due to the acceleration term) as exhibited in the oscillating film case (Eqs. (17)–(20)). Boundary lubrication would be the case as $h_i^*=H^*-2\delta_i^*\ll H^*$. The present perturbation analysis is valid for $\kappa^*\delta_i^*\ll 1$ and thus could be thought of as applying in conditions that approach boundary lubrication. Under these conditions, the model produces much larger load carrying capacity than predicted by pure hydrodynamic theory, consistent with the physical picture.

Figure 5 shows the ratio of normal force with the layers present to the force without the layers. The effect of compliance is to reduce effect of the confined layers, and the reduction is stronger for thinner films.

Results for the case of a simple wedge contact (plane slider bearing) are shown in Figs. 6 and 7. This case is presented as representative of behavior in a converging contact, such as may occur between certain asperities in a micro-EHL contact, rather than as plane slider performance per se.

The first-order solution, shown in Fig. 6, exhibits a reduction of pressure (negative correction) in the first part of the contact, and then an increase is exhibited near the exit. This differs from the parallel squeeze case, where the only effect of compliance is to decrease. Recall that the displacement of the layer is simply proportional to pressure, see Eq. (2). However, in the perturbation modified Reynolds equation, Eq. (11), it is the slope of the zero-order pressure, $\partial p_0^*/\partial x^*$ (proportional to the slope of the film) which drives the correction as well as the zero-order pressure itself. Thus the effect of the elastic layer may be to increase or decrease the correction depending on the sign of the slope. The effect of the relative thickness of the layers is strong, as a small

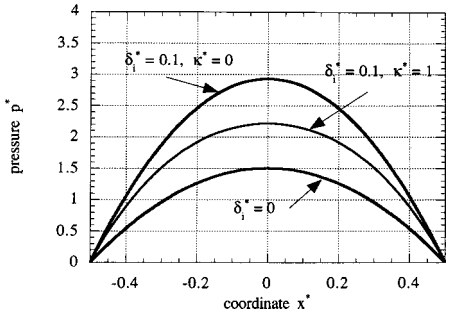


Fig. 4 Dimensionless pressure with and without layers. Parallel squeeze film, effect of layer compliance, with parameter values: $H^*=1$, $\dot{H}^*=-1$, $\ddot{H}=0$.

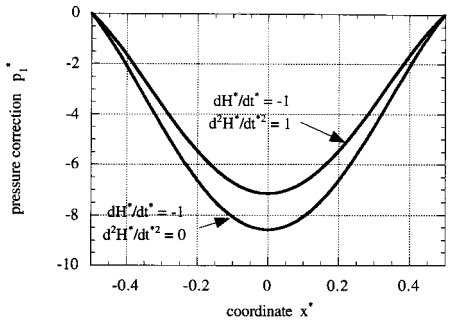


Fig. 2 Pressure correction for layer compliance. Parallel squeeze film, effect of acceleration, with parameter values: $H^*=1$, $\delta_i^*=0.1$.

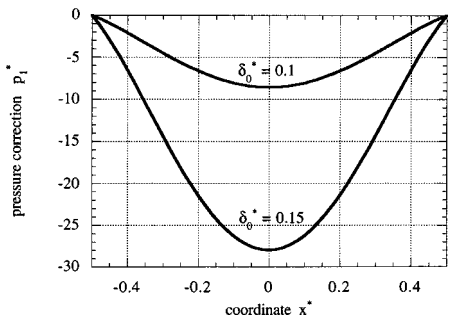


Fig. 3 Pressure correction for layer compliance. Parallel squeeze film, effect of layer thickness, with parameter values: $H^*=1$, $\dot{H}^*=-1$, $\ddot{H}^*=1$.

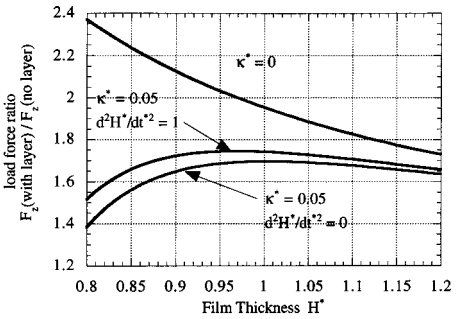


Fig. 5 Load ratio of normal force (with and without elastic layers) versus film thickness. Parallel squeeze film, effect of acceleration and compliance.

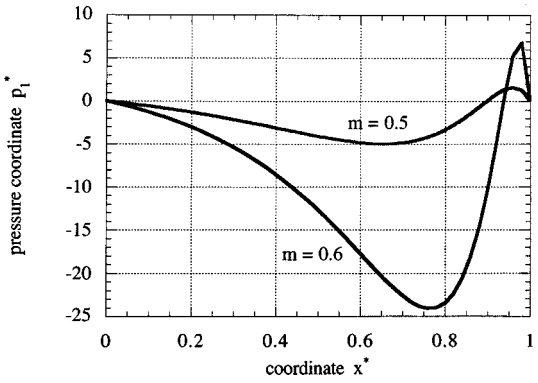


Fig. 6 Pressure correction for layer compliance. Wedge contact, effect of wedge incline, with parameter value: $\delta_i^*=0.1$.

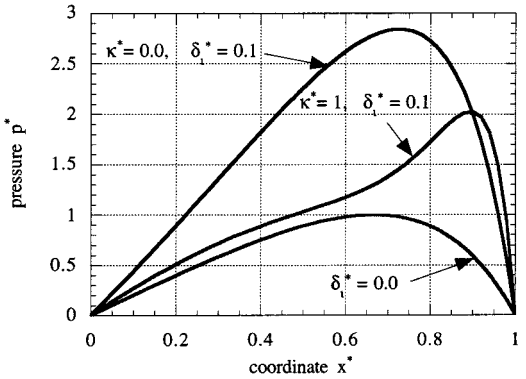


Fig. 7 Dimensionless pressure with and without layers. Wedge contact, effect of layer compliance, with parameter value: $m=0.5$.

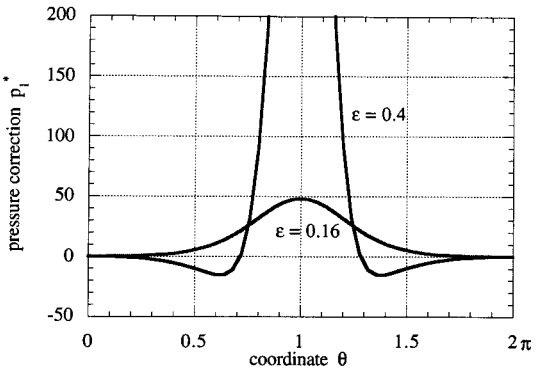


Fig. 9 Pressure correction for layer compliance. Journal bearing contact, effect of eccentricity ratio, with parameter value: $\delta_i=0.1$.

change of the slope parameter m causes a large change in pressure correction. The exit film $h_i^*(1)$ changes from 0.2 to 0.3 (50 percent) for a change in m from 0.6 to 0.5 (17 percent).

The complete pressure field is shown in Fig. 7. Again, three cases are portrayed. The case of a purely Newtonian film with no layer yields the lowest pressure ($\delta_i^*=0$) in which case the compliance parameter κ^* is not defined. The case of rigid layers ($\kappa^*=0$, $\delta_i^*=0.1$) produces much larger pressure everywhere. Compliant layers ($\kappa^*=1$, $\delta_i^*=0.1$) generally reduce the pressure, but increase it slightly at the exit.

Load and friction behavior for the wedge contact is shown in Figs. 8(a) and (b). Both parameters increase with the slope of the wedge. Recall that for all predictions generated by hydrodynamic theory, friction is smaller than load by order of magnitude factor

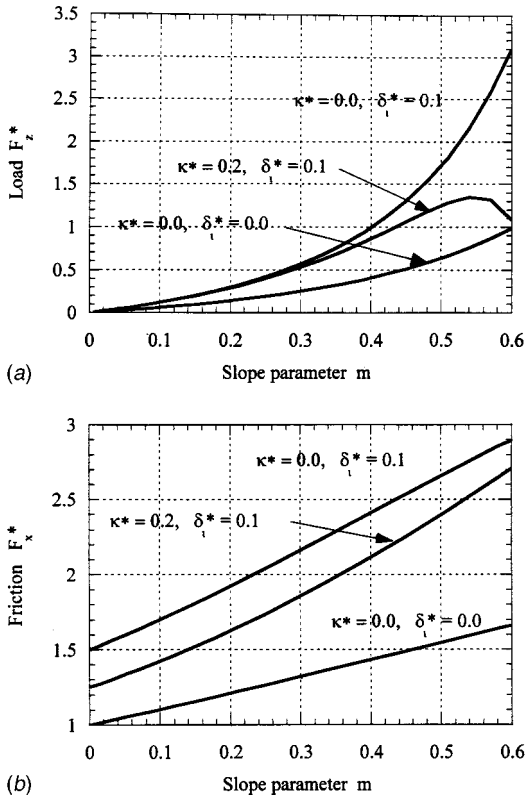


Fig. 8 Wedge contact load and friction behavior. The cases shown are (I) no layer, (II) rigid layer, and (III) compliant layer. (a) Dimensionless load; (b) dimensionless friction.

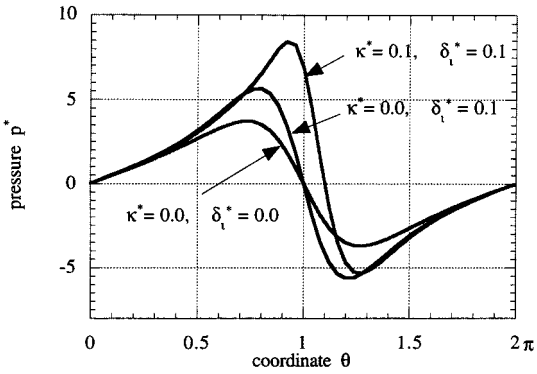


Fig. 10 Dimensionless pressure with and without layers. Journal bearing contact, effect of layer compliance, with parameter value: $e=0.4$.

H/L, which is reflected in the dimensionless scaling. Note that the layers increase both load and friction while compliance tends to reduce these forces. As to modeling of boundary lubrication, the present conceptual model would better represent global load than friction for a contact consisting of an ensemble of microcontacts such as the representative wedge. Global load is likely the sum of the load of the microcontacts. However, global sliding friction is more probably influenced by additional factors such as elastic deformation, ploughing action, etc.

Similarly, results for the case of a journal bearing contact (plane slider bearing) are shown in Figs. 9 and 10. As above, this case is presented as representative of behavior in a converging-diverging contact, rather than as journal bearing behavior.

The first-order solution, shown in Fig. 9 shows a small reduction of pressure (negative correction) in the region of maximum film, then a sharp increase is found at the minimum film point. The same kind of arguments apply here as for the wedge case, concerning the effect of the slope of the pressure and displacement, and the strong effect of the relative local layer thickness.

The complete pressure field is shown in Fig. 10. Again, the same three cases are portrayed and the same trends are exhibited.

Conclusions

Continuum based methods are traditionally thought to be of little value in describing boundary lubrication, or the mode of lubrication in molecular scale films such as occur at asperity interactions during the sliding of nominally flat surfaces. Pure hydrodynamic theory predicts no load carrying action in these conditions. In fact, a definition of boundary lubrication is lubrication where lubricant viscosity in bulk appears to play a negligible role in contact behavior.

A rheological model has been developed which is appropriate to study boundary lubrication in very thin films. The lubricant microstructure in confined contacts is described through the use of a hypothetical compliant elastic layer that adheres to the solid surfaces. The model is applicable to conditions in which the molecular length scale is the same order of magnitude as the film thickness.

The model contains three property parameters: (1) the conventional viscosity, (2) the thickness of the elastic layer, and (3) an effective elastic modulus of the solid-like layer. In principle, the effective properties may be found by molecular dynamics simulations or by curve fitting results to surface force apparatus experiments.

An appropriate modified Reynolds equation is developed which is not much more difficult to solve than the classical equation. Behavior in several typical contacts is found: the simple squeeze film, the converging wedge, and the converging/diverging journal bearing. The strongest effect of the solid-like layer (zero-order perturbation) is to greatly increase the lubricating action relative to predictions of the pure hydrodynamic theory, i.e., for a given set of conditions (film thickness, surface speed, viscosity in bulk) the predicted pressure is larger. Classical hydrodynamics does not predict sufficient load carrying to take place in boundary lubrication conditions. The layer compliance may further increase or decrease the pressure through a first-order correction.

In many otherwise sophisticated computer programs to analyze manufacturing processes, engine dynamics and the like, friction, a crucial component arising from boundary lubrication conditions, is entered as simply a free parameter. Current studies seek to bring, incrementally, the study of boundary lubrication to a level appropriate to its use in modern engineering predictive tools.

References

[1] Israelachvili, J., 1992, *Intermolecular & Surface Forces*, Academic Press, London.

[2] Gee, M. L., McGuiggan, P. M., and Israelachvili, J., 1990, ''Liquid to Solid-Like Transitions of Molecularly Thin Films Under Shear,'' J. Chem. Phys., **93**, No. 3, pp. 1895–1906.

[3] Homola, A., Israelachvili, J., Gee, M., and McGuiggan, P., 1989, ''Measurements of and Relation Between the Adhesion and Friction of Two Surfaces Separated by Molecularly Thin Liquid Films,'' ASME J. Tribol., **111**, pp. 675–682.

[4] Tonck, A., Georges, J. M., and Loubet, J. L., 1988, ''Measurements of Intermolecular Forces and the Rheology of Duodecane Between Alumina Surfaces,'' J. Colloid Interface Sci., **126**, No. 1, pp. 150–163.

[5] Georges, J. M., Millot, S., Loubet, J. L., and Tonck, A., 1993, ''Drainage of Thin Films Between Relatively Smooth Surfaces,'' J. Chem. Phys., **98**, pp. 7345–7360.

[6] Yoshizawa, H., and Israelachvili, J., 1993, ''Fundamental Mechanisms of Interfacial Friction,'' J. Phys. Chem., **97**, No. 43, pp. 11,300–11,313.

[7] Chan, D. Y. C., and Horn, R. G., 1985, ''The Drainage of Thin Liquid Films Between Solid Surfaces,'' J. Chem. Phys., **83**, No. 10, pp. 5311–5324.

[8] Hamrock, B. J., *Fundamentals of Fluid Film Lubrication*, McGraw-Hill, New York, pp. 141–165.

[9] Thompson, P. A., and Robbins, M. O., 1990, ''Shear Flow Near Solids: Epitaxial Order and Flow Boundary Conditions,'' Phys. Rev. A, **41**, No. 12, pp. 6830–6837.

[10] Thompson, P. A., Robbins, M. O., and Grest, G. S., 1993, ''Simulations of Lubricant Behavior at the Interface With Bearing Solids,'' *Thin Films in Tribology*, D. Dowson et al., eds., Elsevier, New York, pp. 347–360.

[11] Hu, Y. Z., Wang, H., and Guo, Y., 1993, ''Molecular Dynamics Simulation of Poiseuille Flow in Ultra Thin Films,'' *Proceedings of the International Symposium on Tribology*, Beijing, China, pp. 58–65.

[12] Landman, U., Luedtke, W. D., Burnham, N. A., and Colton, R. J., 1990, ''Atomistic Mechanisms and Dynamics of Adhesion, Nanoindentation, and Fracture,'' Science, **248**, pp. 454–461.

[13] Tichy, J. A., 1995, ''A Surface Layer Model for Thin Film Lubrication,'' STLE Tribol. Trans., **38**, No. 1, pp. 108–118.

[14] Tichy, J. A., 1995, ''A Porous Media Model for Thin Film Lubrication,'' ASME J. Tribol., **117**, No. 1, pp. 16–21.

[15] Auslander, F., and Sideroff, F., 1996, ''Nanorheological Behavior of Confined Liquid Layers for Normal Contact,'' *The Third Body Concept*, D. Dowson et al., eds., Elsevier, New York, pp. 195–203.

[16] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, New York, pp. 104–107.

Viscoelastic Functionally Graded Materials Subjected to Antiplane Shear Fracture

G. H. Paulino[1]
e-mail: paulino@uiuc.edu
Mem. ASME

Z.-H. Jin
Mem. ASME

Department of Civil and Environmental Engineering,
University of Illinois at Urbana-Champaign,
Newmark Laboratory,
205 North Mathews Avenue,
Urbana, IL 61801

In this paper, a crack in a strip of a viscoelastic functionally graded material is studied under antiplane shear conditions. The shear relaxation function of the material is assumed as $\mu=\mu_0 \exp(\beta y/h) f(t)$, *where h is a length scale and f(t) is a nondimensional function of time t having either the form* $f(t)=\mu_\infty/\mu_0+(1-\mu_\infty/\mu_0)\exp(-t/t_0)$ *for a linear standard solid, or* $f(t)=(t_0/t)^q$ *for a power-law material model. We also consider the shear relaxation function* $\mu=\mu_0 \exp(\beta y/h)[t_0 \exp(\delta y/h)/t]^q$ *in which the relaxation time depends on the Cartesian coordinate y exponentially. Thus this latter model represents a power-law material with position-dependent relaxation time. In the above expressions, the parameters* β, μ_0, μ_∞, t_0; δ, q *are material constants. An elastic crack problem is first solved and the correspondence principle (revisited) is used to obtain stress intensity factors for the viscoelastic functionally graded material. Formulas for stress intensity factors and crack displacement profiles are derived. Results for these quantities are discussed considering various material models and loading conditions.* [DOI: 10.1115/1.1354205]

1 Introduction

Functionally graded materials are the outcome of the need to accommodate materials exposure to nonuniform service requirements. Those materials are characterized by continuously varying properties due to continuous change in *microstructural details* over defined geometrical orientations and distances, such as composition, morphology, and crystal structure. The material gradation may be either continuous or layered comprised, for example, of gradients of ceramics and metals. In applications involving severe thermal gradients (e.g., thermal protection systems), functionally graded material systems take advantage of heat and corrosion resistance typical of ceramics, and mechanical strength and toughness typical of metals. Other relevant applications of functionally graded materials involve polymers ([1]), biomedical systems ([2]), natural composites ([3]), and thermoelectric devices for energy conversion ([4]). Various thermomechanical problems associated to functionally graded materials have been studied, for example, constitutive modeling ([5–7]), higher order theory ([8]), thermal stresses ([9,10]), static and dynamic response of plates ([11]), yield stress gradient effect ([12]), strain gradient theory ([13]), fracture behavior ([14–16]), and statistical model for brittle fracture ([17]).

The antiplane shear crack problem has been extensively studied in the literature as it provides the basis for understanding the opening mode crack problem. Several numerical and analytical/semi-analytical solutions have been presented considering homogeneous materials (e.g., [18,19]), nonhomogeneous materials (e.g., [20,21]), and bonded homogeneous viscoelastic layers ([22]). However, to the best of the authors' knowledge, there is no published analytical/semi-analytical type solution for the problem of an antiplane shear crack in viscoelastic functionally graded materials. This is the subject of this paper, which consists of applying Paulino and Jin's ([23]) revisited correspondence principle for viscoelastic functionally graded materials to fracture mechanics.

One of the primary application areas of functionally graded materials is high-temperature technology. Materials will exhibit creep and stress relaxation behavior at high temperatures. Viscoelasticity offers a basis for the study of phenomenological behavior of creep and stress relaxation. In this paper, viscoelastic fracture (stationary crack) of functionally graded materials is studied under antiplane shear conditions. Specifically, an infinitely long strip containing a crack parallel to the strip boundaries is investigated. The shear relaxation function of the material is assumed to take separable forms in space and time, i.e.,

$$\mu=\mu_0 \exp(\beta y/h) f(t),$$

where h is a length scale and $f(t)$ is a nondimensional function of time t having either the form

$$f(t)=\mu_\infty/\mu_0+(1-\mu_\infty/\mu_0)\exp(-t/t_0): \quad \text{linear standard solid}$$

or

$$f(t)=(t_0/t)^q: \quad \text{power-law material.}$$

We also consider the following variant form of the power-law material model

$$\mu=\mu_0 \exp(\beta y/h)[t_0 \exp(\delta y/h)/t]^q,$$

in which the relaxation time depends on the Cartesian coordinate y exponentially. In the above expressions, the parameters β, μ_0, μ_∞, t_0; δ, q are material constants. An elastic crack problem is first solved and the ''correspondence principle'' is used to obtain the stress intensity factor for the viscoelastic functionally graded material.

This manuscript is organized as follows. The next section presents the basic equations of viscoelasticity theory of functionally graded materials, which are the basis for this study. Then the correspondence principle is revisited and recast in the form recently given by Paulino and Jin [23], followed by a discussion of relaxation functions with separable forms. Next, the antiplane shear problem is formulated together with an integral equation solution approach for a crack in a viscoelastic functionally graded material strip. Formulas for stress intensity factors (as a function

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 24, 2000; final revision, July 13, 2000. Associate Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

of geometry, material constants, and loading) are derived considering both Heaviside step function loading and exponentially decaying or increasing loading. Afterwards, the recovery of the displacement field is carried out and applied to obtain the actual crack profile. Several results for the above problem are presented and discussed. Finally, conclusions are inferred and extensions of this work are pointed out. An Appendix, showing the integral equation kernel derivation, supplements the paper.

2 Basic Equations

The basic equations of quasi-static viscoelasticity of functionally graded materials are the equilibrium equation

$$\sigma_{ij,j}=0, \tag{1}$$

the strain-displacement relationship

$$\varepsilon_{ij}=\tfrac{1}{2}(u_{i,j}+u_{j,i}), \tag{2}$$

and the viscoelastic constitutive law

$$s_{ij}=2\int_0^t \mu(\mathbf{x};t-\tau)\frac{de_{ij}}{d\tau}d\tau, \quad \sigma_{kk}=3\int_0^t K(\mathbf{x};t-\tau)\frac{d\varepsilon_{kk}}{d\tau}d\tau, \tag{3}$$

with

$$s_{ij}=\sigma_{ij}-\tfrac{1}{3}\sigma_{kk}\delta_{ij}, \quad e_{ij}=\varepsilon_{ij}-\tfrac{1}{3}\varepsilon_{kk}\delta_{ij}, \tag{4}$$

where σ_{ij} are stresses, ε_{ij} are strains, s_{ij} and e_{ij} are deviatoric components of the stress and strain tensors, respectively, u_i are displacements, δ_{ij} is the Kronecker delta, $\mathbf{x}=(x_1,x_2,x_3)$, $\mu(\mathbf{x},t)$ and $K(\mathbf{x},t)$ are the relaxation functions in shear and dilatation, respectively, t denotes the time, and the Latin indices have the range 1, 2, 3 with repeated indices implying the summation convention. *Note that for functionally graded materials the relaxation functions also depend on spatial positions, whereas in homogeneous viscoelasticity, they are only functions of time, i.e.* $\mu\equiv\mu(t)$ and $K\equiv K(t)$ ([24]).

3 Correspondence Principle Revisited

In general, the correspondence principle of homogeneous viscoelasticity theory does not hold for functionally graded materials. However, for a class of functionally graded materials with relaxation functions of the form

$$\begin{aligned}\mu(\mathbf{x},t)&=\mu_0\tilde{\mu}(\mathbf{x})f(t),\\ K(\mathbf{x},t)&=K_0\tilde{K}(\mathbf{x})g(t),\end{aligned} \tag{5}$$

where μ_0 and K_0 are material constants, and $\tilde{\mu}(\mathbf{x})$, $\tilde{K}(\mathbf{x})$, $f(t)$, and $g(t)$ are nondimensional functions, Paulino and Jin [23] showed that the correspondence principle still holds. In this case, *the Laplace transformed nonhomogeneous viscoelastic solution can be obtained directly from the solution of the corresponding nonhomogeneous elastic problem by replacing* μ_0 *and* K_0 *with* $\mu_0 p\bar{f}(p)$ *and* $K_0 p\bar{g}(p)$, *respectively, where* $\bar{f}(p)$ *and* $\bar{g}(p)$ *are the Laplace transforms of f(t) and g(t), respectively, and p is the transform variable. The final solution is realized upon inverting the transformed solution*.

Among the various models for graded viscoelastic materials are the *standard linear solid* defined by

$$\begin{aligned}\mu(\mathbf{x},t)&=\mu_\infty(\mathbf{x})+[\mu_e(\mathbf{x})-\mu_\infty(\mathbf{x})]\exp\left[-\frac{t}{t_\mu(\mathbf{x})}\right],\\ K(\mathbf{x},t)&=K_\infty(\mathbf{x})+[K_e(\mathbf{x})-K_\infty(\mathbf{x})]\exp\left[-\frac{t}{t_K(\mathbf{x})}\right],\end{aligned} \tag{6}$$

and the *power-law model* given by

$$\mu(\mathbf{x},t)=\mu_e(\mathbf{x})\left[\frac{t_\mu(\mathbf{x})}{t}\right]^q, \quad K(\mathbf{x},t)=K_e(\mathbf{x})\left[\frac{t_K(\mathbf{x})}{t}\right]^q, \quad 0<q<1, \tag{7}$$

where $t_\mu(\mathbf{x})$ and $t_K(\mathbf{x})$ are the relaxation times in shear and bulk muduli, respectively, and q is a material constant. Particular instances of the above models for graded viscoelastic materials may be obtained such that assumption (5) is satisfied. Thus the discussion below indicates the type of revision needed in the general viscoelastic models so that the correspondence principle still holds.

- *Standard linear solid* (6). If the relaxation times t_μ and t_K are constants, if $\mu_e(\mathbf{x})$ and $\mu_\infty(\mathbf{x})$ have the same functional form, and if $K_e(\mathbf{x})$ and $K_\infty(\mathbf{x})$ also have the same functional form, then the model (6) satisfies assumption (5).
- *Power-law model* (7). If the relaxation times t_μ and t_K are independent of spatial position in model (7), then assumption (5) is readily satisfied. Moreover, even if the relaxation times depend on the spatial position in model (7), then the corresponding nonhomogeneous elastic material has the properties

$$\mu=\mu_e(\mathbf{x})[t_\mu(\mathbf{x})]^q, \quad K=K_e(\mathbf{x})[t_K(\mathbf{x})]^q, \tag{8}$$

rather than $\mu=\mu_e(\mathbf{x})$ and $K=K_e(\mathbf{x})$. Thus assumption (5) is satisfied again.

4 Viscoelastic Antiplane Shear Problem

Under antiplane shear conditions, the only nonvanishing field variables are

$$\begin{aligned}u_3(\mathbf{x},t)&=w(x,y;t),\\ \sigma_{31}(\mathbf{x},t)&=\tau_x(x,y;t), \quad \sigma_{32}(\mathbf{x},t)=\tau_y(x,y;t),\\ 2\varepsilon_{31}(\mathbf{x},t)&=\gamma_x(x,y;t), \quad 2\varepsilon_{32}(\mathbf{x},t)=\gamma_y(x,y;t),\end{aligned}$$

with $\mathbf{x}=(x_1,x_2)=(x,y)$. Here new notations for the nonvanishing displacement, stresses, and strains are used for the sake of simplicity. The basic equations of mechanics satisfied by these variables are

$$\frac{\partial\tau_x}{\partial x}+\frac{\partial\tau_y}{\partial y}=0, \tag{9}$$

$$\gamma_x=\frac{\partial w}{\partial x}, \quad \gamma_y=\frac{\partial w}{\partial y}, \tag{10}$$

$$\tau_x=\int_0^t \mu(x,y;t-\tau)\frac{d\gamma_x}{d\tau}d\tau, \quad \tau_y=\int_0^t \mu(x,y;t-\tau)\frac{d\gamma_y}{d\tau}d\tau. \tag{11}$$

In the present study, the following three material models are employed. The first is the *standard linear solid* (see (6)) with constant relaxation time

$$\mu=\mu_0\exp(\beta y/h)\left[\frac{\mu_\infty}{\mu_0}+\left(1-\frac{\mu_\infty}{\mu_0}\right)\exp\left(-\frac{t}{t_0}\right)\right], \tag{12}$$

where β, μ_0, μ_∞, and t_0 are material constants and h is a length scale. The second model is a *power-law material* (see (7)) with *constant relaxation time*

$$\mu=\mu_0\exp(\beta y/h)\left(\frac{t_0}{t}\right)^q. \tag{13}$$

The third model is also a *power-law material* (see (7)), but with *position-dependent relaxation time*

$$\mu=\mu_0\exp(\beta y/h)\left[\frac{t_0\exp(\delta y/h)}{t}\right]^q=\mu_0\exp[(\beta+\delta q)y/h]\left(\frac{t_0}{t}\right)^q, \tag{14}$$

where δ and q are material constants.

5 Relaxation Functions With Separable Forms

The present discussion is based on the main argument that the functional form of the chosen relaxation function(s) is appropriate if the basic constitutents of the functionally graded material have

approximately the same relaxation pattern. Thus this argument indicates the need for an approach integrating mechanics modeling, material properties experiments, and synthesis (see [25] for a review of fabrication processes for functionally graded materials). This point is elaborated upon below.

It can be seen in (12), (13), and (14) that the relaxation moduli are separable functions in space and time. This is necessary for the revisited correspondence principle (see Section 3) to be applied ([23]). This kind of relaxation functions may be appropriate for a functionally graded material with its constituent materials having the same time-dependence of shear modulus. For model (12), this means that the constituents should have the same ratio μ_∞/μ_0 and relaxation time t_0. For model (13), this implies that the constituents should have the same relaxation time t_0 and parameter q. For model (14), however, it is only required that the constituents have the same parameter q. The constituents may have different relaxation times. Potentially, this kind of functionally graded materials may include some polymeric/polymeric materials such as Propylene-homopolymer/Acetal-copolymer. The relaxation behavior of Propylene homopolymer and Acetal copolymer are found to be similar—see Figs. 7.5 and 10.3, respectively, of Ogorkiewicz [26].

Another argument potentially in favor of the selected class of relaxation functions (5) is the technique developed by Lambros et al. [27] for fabricating large scale polymeric functionally graded materials. The technique consists of generating a continuously inhomogeneous property variation by selective ultraviolet irradiation of a polyethylene carbon monoxide copolymer. Due to the fact that the functionally graded material is obtained by controlling ultraviolet irradiation time of the same base polymer, we conjecture that the viscoelastic behavior of such material may be predicted by (5). However, further experimental research needs to be done in order to validate or invalidate the present conjecture.

6 Mode III Crack in a Functionally Graded Material Strip

Consider an infinite nonhomogeneous viscoelastic strip containing a central crack of length $2a$, as shown in Fig. 1. The strip is fixed along the lower boundary ($y=-h$) and is displaced $w(t)=w_0W(t)$ along the upper boundary ($y=h$), where w_0 is a constant and $W(t)$ is a nondimensional function of time t. It is supposed that the crack lies on the x-axis, from $-a$ to a, and is of infinite extent in the z-direction (normal to the x-y plane). The crack surfaces remain traction-free. The boundary conditions of the crack problem, therefore, are

$$w=0, \quad y=-h, \quad |x|<\infty, \tag{15}$$

$$w=w_0W(t), \quad y=h, \quad |x|<\infty, \tag{16}$$

$$\tau_y=0, \quad y=0, \quad |x|\leqslant a, \tag{17}$$

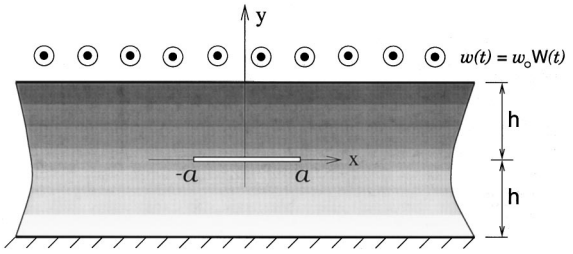


Fig. 1 A viscoelastic functionally graded material strip occupying the region $|x|<\infty$ and $|y|\leqslant h$ with a crack at $|x|\leqslant a$ and $y=0$. The lower boundary of the strip ($y=-h$) is fixed and the upper boundary ($y=h$) is subjected to uniform antiplane displacement $w_0W(t)$. The symbol ⊙ indicates an arrow perpendicular to the strip plane and pointing toward the viewer.

$$\tau_y(x,0^+)=\tau_y(x,0^-), \quad a<|x|<\infty, \tag{18}$$

$$w(x,0^+)=w(x,0^-), \quad a<|x|<\infty. \tag{19}$$

According to the correspondence principle (see Section 3), one can first consider a nonhomogeneous elastic material with the shear modulus

$$\mu=\mu_0\exp(\beta y/h), \tag{20}$$

and the viscoelastic solutions for models (12) and (13) may be obtained by the correspondence principle. For the material model (14) the viscoelastic solution can still be obtained by the correspondence principle provided that the corresponding elastic material has the shear modulus $\mu=\mu_0\exp[(\beta+q\delta)y/h]$ (cf. (5) and (14)).

For the elastic crack problem, the solution consists of a regular solution (for an uncracked strip)

$$w=w(y)=\frac{\exp(\beta)-\exp(-\beta y/h)}{\exp(\beta)-\exp(-\beta)}w_0, \tag{21}$$

$$\tau_x=0, \quad \tau_y=\frac{\beta\mu_0 w_0/h}{\exp(\beta)-\exp(-\beta)} \tag{22}$$

and a perturbed solution (by the crack) satisfying the following boundary conditions:

$$w=0, \quad y=\pm h, \quad |x|<\infty, \tag{23}$$

$$\tau_y=-\frac{\beta\mu_0 w_0/h}{\exp(\beta)-\exp(-\beta)}, \quad y=0, \quad |x|\leqslant a, \tag{24}$$

$$\tau_y(x,0^+)=\tau_y(x,0^-), \quad a<|x|<\infty, \tag{25}$$

$$w(x,0^+)=w(x,0^-), \quad a<|x|<\infty. \tag{26}$$

The governing differential equation of $w(x,y)$ for the nonhomogeneous elastic material (20) is

$$\nabla^2 w+\frac{\beta}{h}\frac{\partial w}{\partial y}=0. \tag{27}$$

By using the Fourier transform method (see, for example, [28]), the boundary value problem described by Eqs. (23) to (27) can be reduced to the following singular integral equation (see Appendix):

$$\int_{-1}^{1}\left[\frac{1}{s-r}+k(r,s,\beta)\right]\varphi(s)ds=-\frac{2\pi\beta w_0/h}{\exp(\beta)-\exp(-\beta)}, \quad |r|\leqslant 1, \tag{28}$$

where the unknown density function $\varphi(r)$ is given by

$$\varphi(x)=\frac{\partial}{\partial x}[w(x,0^+)-w(x,0^-)], \tag{29}$$

the nondimensional coordinates r and s are

$$r=x/a, \quad s=x'/a, \tag{30}$$

respectively, and the Fredholm kernel $k(r,s,\beta)$ is

$$k(x,x',\beta)=a\int_0^\infty P(x,x',\xi,\beta)d\xi \tag{31}$$

with $P(x,x',\beta)$ being given by

$$\begin{aligned}P(x,x',\xi,\beta)=[&\xi(\sqrt{(\beta/h)^2+4\xi^2}-2\xi)-2(\beta^2/h^2+2\xi^2)\\&\times\exp(-\sqrt{\beta^2+4h^2\xi^2})\\&-\xi(2\xi+\sqrt{(\beta/h)^2+4\xi^2})\\&\times\exp(-2\sqrt{\beta^2+4h^2\xi^2})]\end{aligned}$$

$$\times\frac{\sin[(x-x')\xi]}{\xi[1-\exp(-2\sqrt{\beta^2+4h^2\xi^2})]\sqrt{(\beta/h)^2+4\xi^2}}. \quad (32)$$

As expected, in the limit of $h\to\infty$ (free space) and $\beta\to 0$ (homogeneous material case), we obtain that $P(x,x',\xi,\beta)\to 0$. *Moreover, the kernel $k(x,x',\beta)$ is symmetric with respect to β.* Such symmetry will be addressed later in the paper. The function $\varphi(r)$ can be further expressed as

$$\varphi(r)=\psi(r)/\sqrt{1-r^2}, \quad (33)$$

where $\psi(r)$ is continuous for $r\in[-1,1]$. When $\varphi(r)$ is normalized by w_0/h, the elastic Mode III stress intensity factor, K^e_{III}, is obtained as

$$K^e_{\mathrm{III}}=-\mu_0\left(\frac{w_0}{2h}\right)\sqrt{\pi a}\,\psi(1,\beta). \quad (34)$$

Here, the notation $\psi(1,\beta)$ is adopted to emphasize the dependence of $\psi(1)$ on β.

7 Stress Intensity Factors

The stress intensity factors for viscoelastic functionally graded materials satisfying (5) can be obtained using the correspondence principle between the elastic and the Laplace transformed viscoelastic equations. Thus, formulas for stress intensity factors are derived first for general time-dependent loading, and then the results obtained are particularized for exponentially decaying or increasing loading and Heaviside step function loading.

As stated above, for nonhomogeneous viscoelastic materials, the Mode III stress intensity factor, K_{III}, can be obtained by means of the correspondence principle (see Section 3). The upper boundary $y=h$ of the strip is subjected to an antiplane displacement $w_0W(t)$, as illustrated by Fig. 1. In this case, the stress intensity factors for material models (12), (13), and (14) will be

$$K_{\mathrm{III}}=-\mu_0\left(\frac{w_0}{2h}\right)\sqrt{\pi a}\,\psi(1,\beta)\mathcal{L}^{-1}\times\left\{\left[\frac{\mu_\infty}{\mu_0}+\left(1-\frac{\mu_\infty}{\mu_0}\right)\frac{p}{p+1/t_0}\right]\overline{W}(p)\right\}, \quad (35)$$

$$K_{\mathrm{III}}=-\mu_0\left(\frac{w_0}{2h}\right)\sqrt{\pi a}\,\psi(1,\beta)\mathcal{L}^{-1}[t_0^q\Gamma(1-q)p^q\overline{W}(p)], \quad (36)$$

and

$$K_{\mathrm{III}}=-\mu_0\left(\frac{w_0}{2h}\right)\sqrt{\pi a}\,\psi(1,\beta+q\delta)\mathcal{L}^{-1}[t_0^q\Gamma(1-q)p^q\overline{W}(p)], \quad (37)$$

respectively, where p is the Laplace transform variable, $\mathcal{L}^{-1}$ represents the inverse Laplace transform, $\overline{W}(p)$ is the Laplace transform of $W(t)$, and $\Gamma(\cdot)$ is the Gamma function.

7.1 Stress Intensity Factors for Exponentially Decaying or Rising Loading. Consider as an example

$$W(t)=\exp(-t/t_L)\to\overline{W}(p)=1/(p+1/t_L) \quad (38)$$

where t_L is a constant measuring the time variation of the load. Note that $t_L>0$ represents an exponentially decaying load, while $t_L<0$ corresponds to an exponentially rising load. This kind of time-dependent loading has been used by Broberg [29]. The stress intensity factors under the loading condition (38) then become

$$K_{\mathrm{III}}=-\mu_0\left(\frac{w_0}{2h}\right)\sqrt{\pi a}\,\psi(1,\beta)F(t), \quad (39)$$

where

$$F(t)=\frac{\mu_\infty}{\mu_0}\exp\left(-\frac{t}{t_L}\right)+\left(1-\frac{\mu_\infty}{\mu_0}\right)\frac{1}{t_0-t_L}\times\left[t_0\exp\left(-\frac{t}{t_L}\right)-t_L\exp\left(-\frac{t}{t_0}\right)\right], \quad (40)$$

for the *standard linear solid* (12), and

$$F(t)=\left(\frac{t_0}{t}\right)^q-\frac{1}{t_L}\int_0^t\left(\frac{t_0}{\tau}\right)^q\exp\left(-\frac{t-\tau}{t_L}\right)d\tau, \quad (41)$$

for the *power-law model* (13).

For the *power-law material with position-dependent relaxation time* (14), the stress intensity factor is

$$K_{\mathrm{III}}=-\mu_0\left(\frac{w_0}{2h}\right)\sqrt{\pi a}\,\psi(1,\beta+q\delta)F(t), \quad (42)$$

where $F(t)$ is given in (41).

7.2 Stress Intensity Factors for Heaviside Step Function Loading. For the Heaviside loading conditions,

$$W(t)=H(t)\to\overline{W}(p)=1/p, \quad (43)$$

where $H(t)$ is the Heaviside step function. The stress intensity factors then become (cf. (39))

$$K_{\mathrm{III}}=-\mu_0\left(\frac{w_0}{2h}\right)\sqrt{\pi a}\,\psi(1,\beta)F(t),$$

where $F(t)$ is given by

$$F(t)=\frac{\mu_\infty}{\mu_0}+\left(1-\frac{\mu_\infty}{\mu_0}\right)\exp\left(-\frac{t}{t_0}\right), \quad (44)$$

for the *linear standard solid* (12), and

$$F(t)=\left(\frac{t_0}{t}\right)^q \quad (45)$$

for the *power-law model* (13).

Finally, the stress intensity factor for the *power-law material with position-dependent relaxation time* (14) is given by (cf. (42))

$$K_{\mathrm{III}}=-\mu_0\left(\frac{w_0}{2h}\right)\sqrt{\pi a}\,\psi(1,\beta+q\delta)F(t),$$

where $F(t)$ is provided in (45). It is seen that q and δ (parameters describing the position dependence of the relaxation time) affect the stress intensity factor only through the combined parameter $(\beta+q\delta)$.

8 Crack Displacement Profile

Accurate representation of the crack profile is relevant in fracture mechanics, especially when the crack-surface displacement is measured experimentally and correlated with results obtained by numerical methods. Thus the crack displacement profile for the problem illustrated in Fig. 1 is recovered in this section. First, general time-dependent loading is considered, and then the formulation is particularized for the Heaviside step function loading and the exponentially decaying or rising loading.

It follows from Eqs. (29) and (33), and the correspondence principle, that the crack-sliding displacement under the time-dependent loading, $w_0 W(t)$, can be expressed by the density function $\varphi(x)$ or $\psi(r)$ (normalized by w_0/h) as follows:

$$[w]=w(x,0^+)-w(x,0^-) =\frac{w_0 W(t)}{h}\int_{-a}^{x}\varphi(x')dx'=w_0 W(t)\left(\frac{a}{h}\right)\int_{-1}^{r}\frac{\psi(s)}{\sqrt{1-s^2}}ds. \tag{46}$$

The displacement at the upper surface of the crack is given by

$$w(x,0^+)=\frac{1}{2}[w]+\frac{1}{2\pi}\left[\frac{w_0 W(t)}{h}\right]\int_{-a}^{a}k_d(x,x')\varphi(x')dx' =\frac{1}{2}[w]+\frac{w_0 W(t)}{2\pi}\left(\frac{a}{h}\right)\int_{-1}^{1}k_d(r,s)\frac{\psi(s)}{\sqrt{1-s^2}}ds, \tag{47}$$

where $k_d(x,x')$ is

$$k_d(x,x')=\int_0^\infty\frac{[-\beta+2\beta\exp(-\sqrt{\beta^2+4h^2\xi^2})-\beta\exp(-2\sqrt{\beta^2+4h^2\xi^2})]}{\sqrt{(\beta/h)^2+4\xi^2}[1-\exp(-2\sqrt{\beta^2+4h^2\xi^2})]}\times\frac{\sin[(x-x')\xi]}{\xi}d\xi. \tag{48}$$

Note that the displacements are not symmetric with respect to β (see Fig. 1), however, the stress intensity factors are (cf. (28) and (32)). The displacement at the lower crack surface is then given by

$$w(x,0^-)=w(x,0^+)-[w]. \tag{49}$$

In expressions (46) and (47), $W(t)$ is given in (38) for the exponentially decaying or rising load. For the Heaviside step function load, $W(t)$ is given by (43).

9 Numerical Aspects

To obtain the numerical solution of integral Eq. (28), $\psi(r)$ is expanded into a series of Chebyshev polynomials of the first kind. By noting the relationship (33) between $\varphi(r)$ and $\psi(r)$, $\varphi(r)$ is expressed as follows (Erdogan et al. [28]):

$$\varphi(r)=\frac{1}{\sqrt{1-r^2}}\sum_{n=1}^{\infty}a_n T_n(r),\quad |r|\leqslant 1, \tag{50}$$

where $T_n(r)$ are Chebyshev polynomials of the first kind and a_n are unknown constants. By substituting the above equation into integral Eq. (28), we have

$$\sum_{n=1}^{\infty}\{\pi U_{n-1}(r)+H_n(r)\}a_n=-\frac{\pi\beta w_0/h}{\exp(\beta)-\exp(-\beta)},\quad |r|\leqslant 1, \tag{51}$$

where $U_{n-1}(r)$ are Chebyshev polynomials of the second kind and $H_n(r)$ are given by

$$H_n(r)=\int_{-1}^{1}ak(r,s,\beta)\frac{T_n(s)}{\sqrt{1-s^2}}ds. \tag{52}$$

To solve the functional Eq. (51), the series on the left side is first truncated at the Nth term. A collocation technique is then used and the collocation points, r_i, are chosen as the roots of the Chebyshev polynomials of the first kind

$$r_i=\cos\frac{(2i-1)\pi}{2N},\quad i=1,2,\ldots,N. \tag{53}$$

The functional Eq. (51) is then reduced to a linear algebraic equation system

$$\sum_{n=1}^{N}\{\pi U_{n-1}(r_i)+H_n(r_i)\}a_n=-\frac{\pi\beta w_0/h}{\exp(\beta)-\exp(-\beta)},\quad i=1,2,\ldots,N. \tag{54}$$

After $a_n(n=1,2,\ldots,N)$ are determined, the nondimensional stress intensity factor, $-\psi(1,\beta)$, is computed as follows:

$$-\psi(1,\beta)=-\sum_{n=1}^{N}a_n. \tag{55}$$

In the following numerical calculations, it is found that 20 collocation points lead to a convergent stress intensity factor result.

It is known from (39) that the stress intensity factor is a multiplification of three parts. The first term is a dimensional base, $\mu_0(w_0/2h)\sqrt{\pi a}$, the second term is a geometrical and material nonhomogeneity correction factor, $-\psi(1,\beta)$, which can be obtained from the numerical solution of singular integral Eq. (28), and the third term is the time evolution of stress intensity factor, F(t), which is obtained analytically from the inverse Laplace transform.

Note that, according to Fig. 1, the crack is located at midheight of the material strip and the origin of the coordinate system (x,y) is located at the center of the crack. Such choice of reference system introduces certain symmetries in the solution, which are discussed in the examples below.

10 Results

Numerical results for stress intensity factors are first obtained for a homogeneous elastic strip (see Fig. 1). According to Table 1, the stress intensity factors are found in good agreement with those reported in the literature, e.g., the handbook by Tada et al. [30]. Furthermore, for a homogeneous viscoelastic strip, it is evident that the stress intensity factor is given by (39) with $\beta=0$ and $F(t)$ is given by (40) and (41) for the exponentially decaying or rising loading, and by (44) and (45) for the Heaviside step function loading. Note that the function $F(t)$ is not related to the nonhomogeneous material parameter β.

Figure 2 shows normalized stress intensity factor (see (39)), $-\psi(1,\beta)$, versus the nonhomogeneous parameter β considering various strip thicknesses h/a for the linear standard solid (12) and the power-law model with constant relaxation time (13). Note that although the relaxation times are different for both models

Table 1 Mode III stress intensity factors (SIF) for a homogeneous strip

h/a	SIF (this study)	SIF ([30])
0.5	0.5360	0.5631
1.0	0.7598	0.7641
1.5	0.8626	0.8634
2.0	0.9136	0.9138
5.0	0.9840	0.9840
10.0	0.9959	0.9959

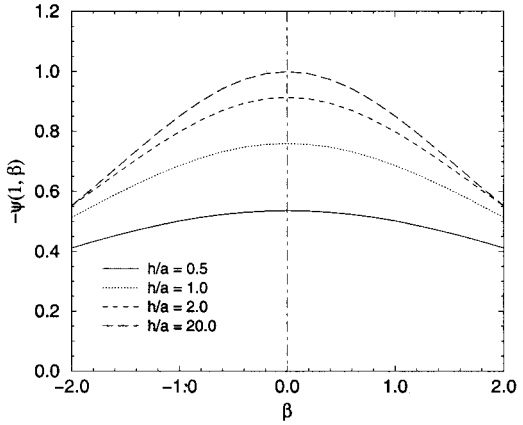


Fig. 2 Normalized Mode III stress intensity factor versus nonhomogeneous material parameter β for various strip thicknesses considering the linear standard solid and the power-law material with constant relaxation time

(cf. Eqs. (40) and (41), or Eqs. (44) and (45)), they do have the same solution $-\psi(1,\beta)$ (see Section 5). The stress intensity factor decreases with increasing $|\beta|$ for all thickness cases (h/a) considered. The stress intensity factor is lower than that of the corresponding homogeneous material ($\beta=0$). It is noted that the stress intensity factor is an even function of β. However, this symmetry is valid only for the crack located in the center of the strip.

Figure 3 shows normalized stress intensity factor (see (42)), $-\psi(1,\beta+q\delta)$, versus the nonhomogeneous parameter β for various strip thicknesses h/a and three δ values for the power-law model with position-dependent relaxation time (14). The effect of spatial position dependence of the relaxation time on the stress intensity factor is reflected through the parameter δ. The parameter q is taken as 0.4 in all calculations. Thus the curves for $\delta=\pm1$ may be obtained from the curve $\delta=0$ by shifting this curve by $\beta=\mp0.4$. It is clear from Fig. 3 that with respect to the corresponding model with constant relaxation time (i.e., $\delta=0$), a positive δ lowers the stress intensity factor when $\beta>0$ and increases the stress intensity factor for β less than $-0.5q\delta$. A negative δ lowers the stress intensity factor when $\beta<0$ and increases the stress intensity factor for β larger than $0.5q\delta$.

Figure 4 illustrates the time evolution of normalized stress intensity factors, $F(t)$, considering both Heaviside step function loading and exponentially decaying or rising loading for the standard linear solid (see (40) and (44)) and the power-law material (see (41) and (45)). The ratio μ_∞/μ_0 is taken as 0.5 in all subsequent calculations for the standard linear solid. It is evident that under the fixed displacement condition, stress intensity factor decreases monotonically with increasing time for Heaviside step function loading and exponentially decaying loading (Figs. 4(*a*) and 4(*b*)). For exponentially rising loading, however, the stress intensity factors will increase with time for longer times (Figs. 4(*c*) and 4(*d*)). By observing the plots in Figs. 4(*a*) and 4(*b*), one notices that, for exponentially decaying loading, the stress intensity factor can become negative as the ratio t_L/t_0 decreases, which occurs, for example, for $t_L/t_0=1.0$. This happens because of stress relaxation for long-time behavior. *Note that a negative stress intensity factor does not imply crack closure because it is*

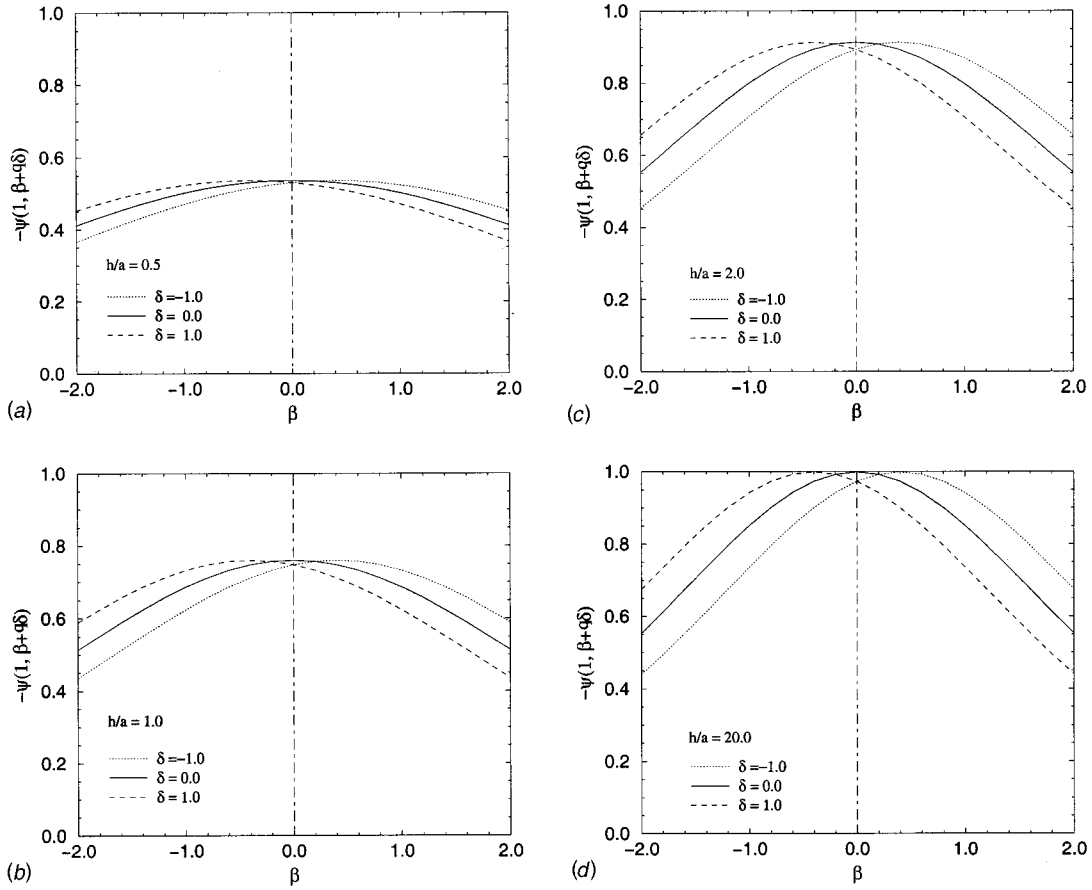


Fig. 3 Normalized Mode III stress intensity factors versus nonhomogeneous parameter β for three δ values and $q=0.4$, (*a*) $h/a=0.5$; (*b*) $h/a=1.0$; (*c*) $h/a=2.0$; (*d*) $h/a=20.0$

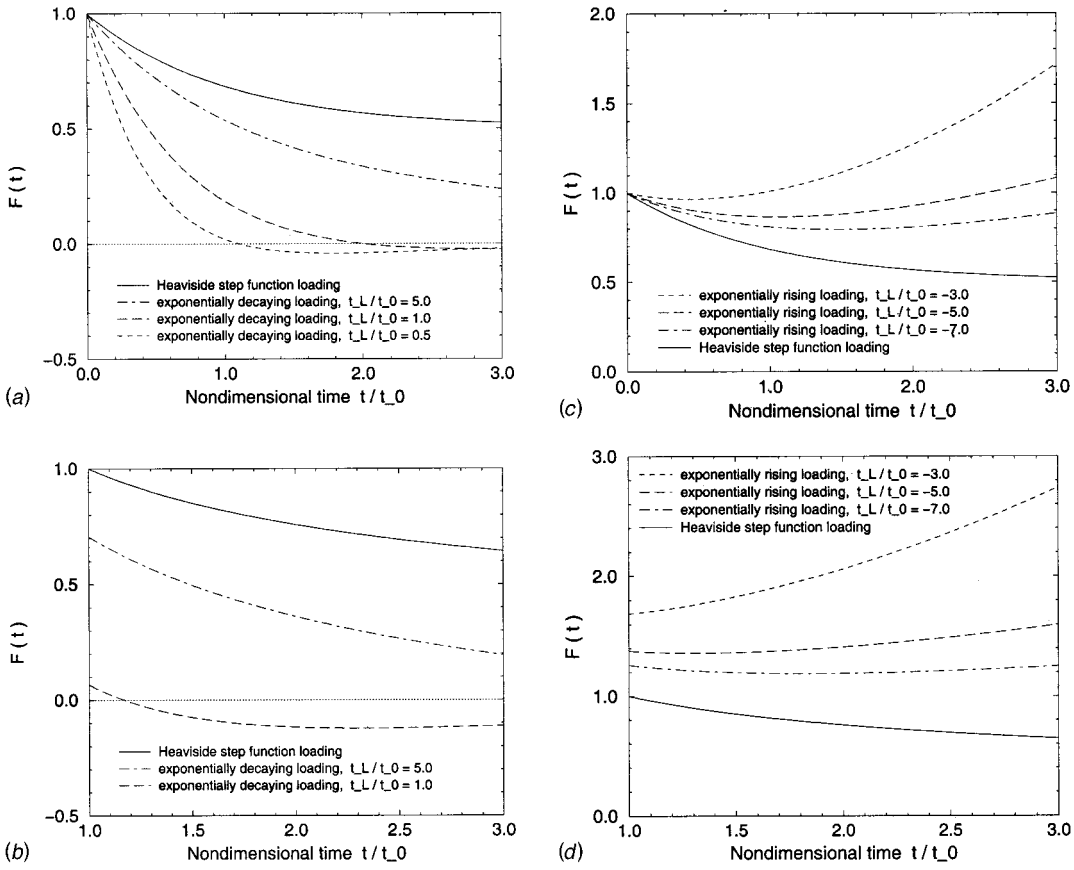


Fig. 4 Time variation of normalized Mode III stress intensity factor (*a*) standard linear solid (decaying loading); (*b*) power-law material (decaying loading); (*c*) standard linear solid (rising loading); (*d*) power-law material (rising load)

associated to a Mode III crack, and not a Mode I (or mixed mode) crack. Thus, in the present study, a negative stress intensity factor is allowed without violating the crack face traction free condition. The crack faces do not close, they just slide in the opposite direction.

Figure 5 illustrates the normalized stress intensity factors (normalized by $\mu_0(w_0/2h)\sqrt{\pi a}$) versus time for Heaviside step function loading considering the following viscoelastic material models: standard linear solid (see (39) and (44)), power law material (see (39) and (45)), and power-law material with position-dependent relaxation time (see (42) and (45)). A finite thickness strip with $h/a=2.0$ (Fig. 1) is considered. Note that, for all the models, the stress intensity factors decrease monotonically with increasing time. The first two models are investigated for the nonhomogeneous parameter $\beta=0$, 1, 2 with $\beta=0$ corresponding to the homogeneous material case. Due to the symmetry of stress intensity factor about β, the stress intensity factor for $\beta=-1$, -2 are identical to those for $\beta=1$, 2, respectively. Moreover, the stress intensity factor decreases with increasing $|\beta|$. The last model is investigated for $\beta=2$ and $\delta=-1,0,1$ with $\delta=0$ corresponding to position-independent relaxation time. For this special case, the stress intensity factor decreases with increasing δ. This is because $\beta+q\delta$ is positive for the β and δ values considered.

Figure 6 illustrates the normalized stress intensity factors (normalized by $\mu_0(w_0/2h)\sqrt{\pi a}$) versus time for exponentially decaying loading considering the following models: standard linear solid (see (39) and (40)), power-law material (see (39) and (41)), and power-law material with position-dependent relaxation time (see (42) and (41)). The same qualitative observations for Fig. 5 also hold for Fig. 6.

Figure 7 presents the normalized stress intensity factors (normalized by $\mu_0(w_0/2h)\sqrt{\pi a}$) versus time for exponentially rising loading for the standard linear solid (see (39) and (40)). Comparing this figure with Figs. 5(*a*) and 6(*a*) (Heaviside step function loading and exponentially decaying loading), one observes that the time variation of stress intensity factors show a convex shape in Fig. 7 while it shows a monotonically decreasing trend in Figs. 5(*a*) and 6(*a*).

Figure 8 shows crack profiles for the Heaviside step function loading considering the standard linear solid and the power law material with position-dependent relaxation time (see (46), (47), and (49)). A finite thickness strip geometry with $h/a=2$ (Fig. 1) is considered. The former model (Fig. 8(*a*)) is investigated for the nonhomogeneity parameter $\beta=0$, 1, 2 with $\beta=0$ corresponding to the homogeneous material case. The latter model (Fig. 8(*b*)) is investigated for $\beta=2$ and $\delta=-1$, 0, 1 with $\delta=0$ corresponding to position-independent relaxation time.

Figure 9 shows crack profiles for the exponentially decaying loading considering the standard linear solid and the power-law material with position-dependent relaxation time (see (46), (47), and (49)). As before, a finite thickness strip geometry with $h/a=2$ (Fig. 1) is considered. The former model (Fig. 9(*a*)) is investigated for the nonhomogeneity parameter $\beta=2$ and $t/t_0=1$, 2, 3. The latter model (Fig. 9(*b*)) is investigated for $\beta=2$, $\delta=1$, and $t/t_0=1$, 2, 3. A comparison of all the plots in Figs. 8 and 9 permits to evaluate the corresponding crack profiles for various material models and various parameters. This information is potentially valuable when correlated with fracture experiments, e.g., crack-sliding displacement measurements.

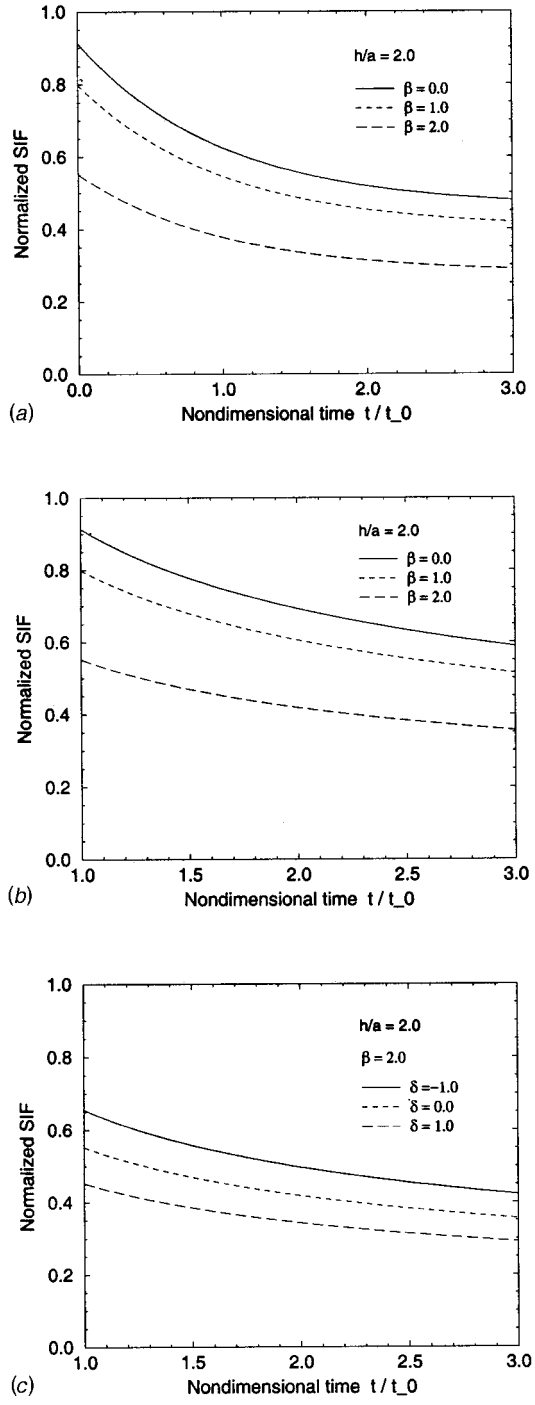


Fig. 5 Normalized Mode III stress intensity factor versus time: Heaviside step function loading, (*a*) standard linear solid; (*b*) power-law material; (*c*) power-law material with position-dependent relaxation time

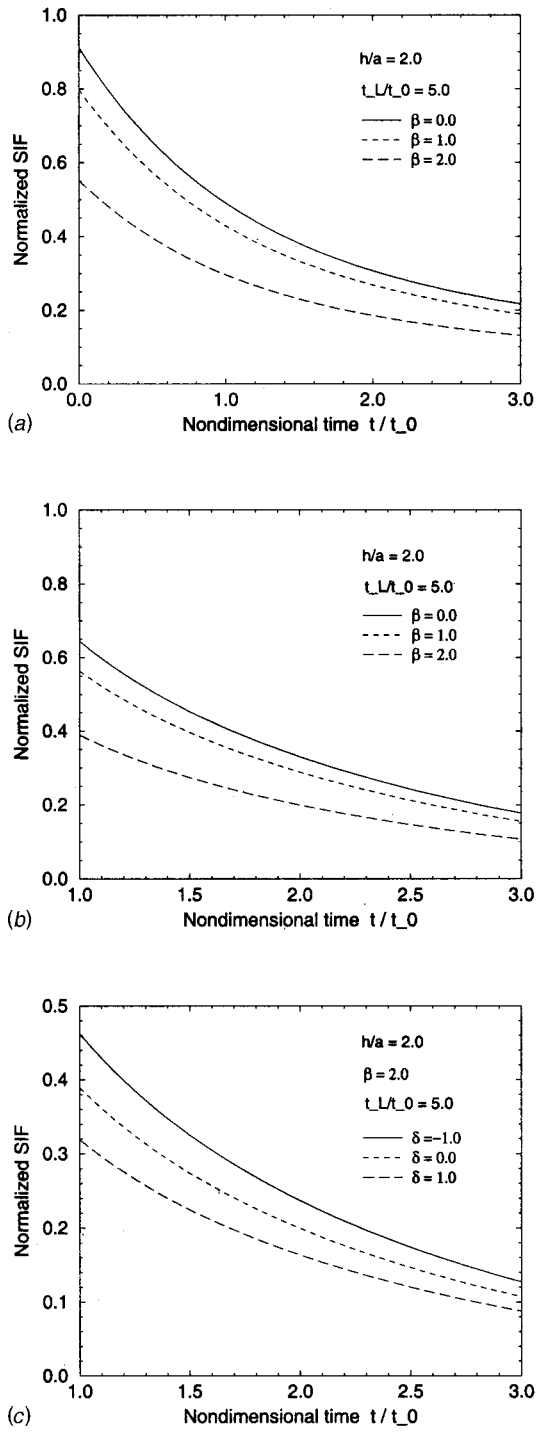


Fig. 6 Normalized mode III stress intensity factor versus time: exponentially decaying loading, (*a*) standard linear solid; (*b*) power-law material; (*c*) power-law material with position-dependent relaxation time

11 Concluding Remarks and Extensions

This paper illustrates an application of Paulino and Jin's [23] revisited correspondence principle to fracture mechanics of viscoelastic functionally graded materials. An effective integral equation method for antiplane shear cracking in viscoelastic functionally graded materials is presented. The elastic functionally graded material crack problem is solved and the correspondence principle between the elastic and the Laplace transformed viscoelastic equations is used to obtain stress intensity factors for viscoelastic functionally graded materials. Formulas for stress intensity factors and crack displacement profiles are provided. Several numerical results for these quantities are presented considering various viscoelastic material models (e.g., standard linear solid, power-law model with both position-independent and position-dependent relaxation time) and loading conditions. It is important to remark that the solution of the fracture mechanics problem with prescribed displacement (see Fig. 1) is different from the solution with prescribed traction (cf. Erdogan [15,20]).

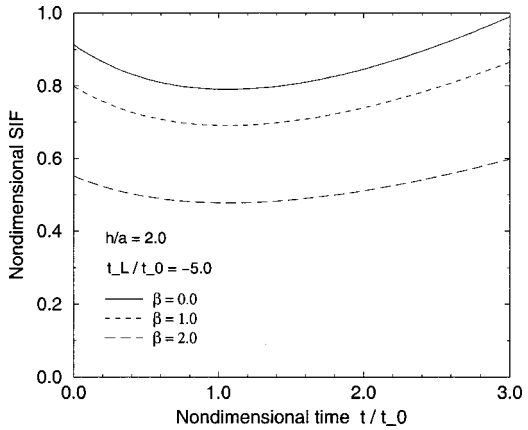


Fig. 7 Normalized Mode III stress intensity factor versus time: exponentially rising loading (standard linear solid)

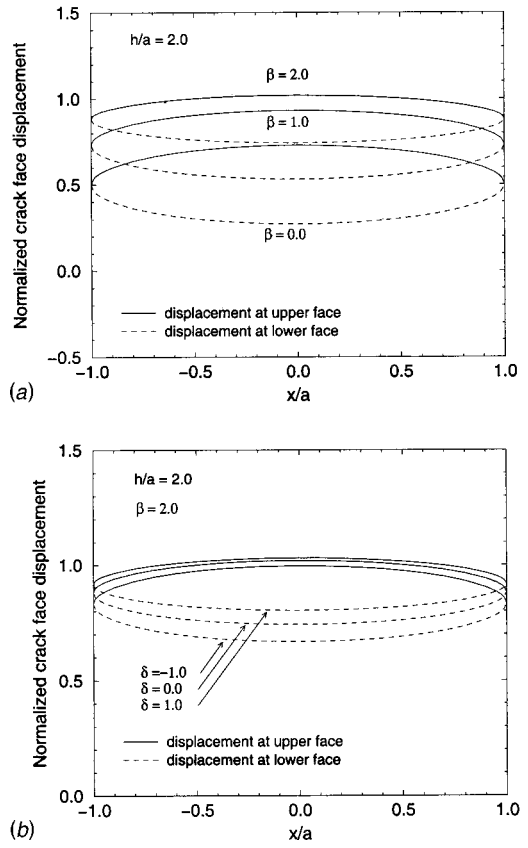


Fig. 8 Crack face displacements: Heaviside step function loading, (*a*) standard linear solid; (*b*) power-law material with position-dependent relaxation time

This work offers promising avenues for further extension. For instance, it may be used to calibrate numerical methods (e.g., finite element method and boundary element method) for viscoelastic functionally graded materials. Moreover, the discussion on relaxation functions of separable forms in space and time (Section 5) indicates the need for micromechanics models for viscoelastic behavior. Other relevant topics associated with this work also deserve further investigation. Such topics include: (a) investigation of antiplane shear cracking in bonded viscoelastic layers where one of the layers is a functionally graded material; (b) extension of the antiplane shear crack model to Mode I cracks. These topics are presently being pursued by the authors.

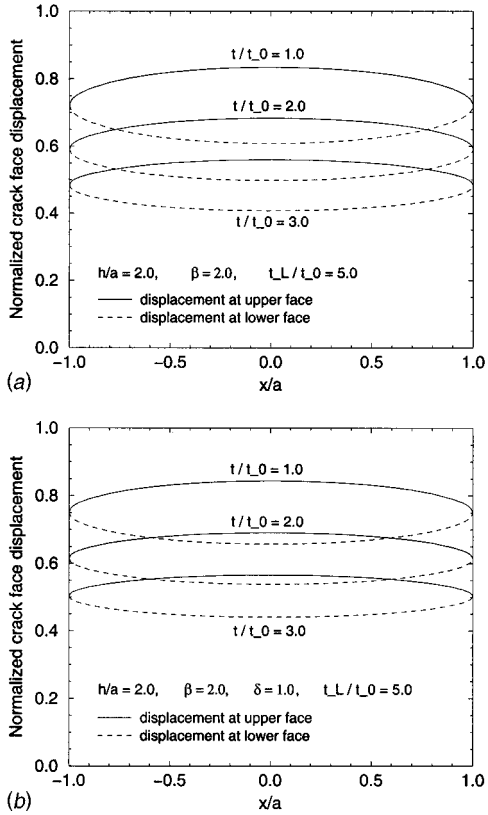


Fig. 9 Crack face displacements: exponentially decaying loading, (*a*) standard linear solid; (*b*) power-law material with position-dependent relaxation time

Acknowledgments

We would like to acknowledge Professor Pindera and four anonymous reviewers for their valuable comments regarding our work. We also acknowledge the support from the National Science Foundation (NSF) under grant No. CMS-9996378 (Mechanics & Materials Program).

Appendix

In the following, a relatively detailed derivation of integral Eq. (28) is given, which refers to the Mode III fracture mechanics problem illustrated by Fig. 1. By using Fourier transform, the solution of the basic Eq. (27) can be expressed as follows:

$$w=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{A_1\exp\left[\frac{-\beta+m}{2}\frac{y}{h}\right]+A_2\exp\left[\frac{-\beta-m}{2}\frac{y}{h}\right]\right\}\times\exp(-ix\xi)d\xi,\quad y>0,$$

$$w=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{B_1\exp\left[\frac{-\beta+m}{2}\frac{y}{h}\right]+B_2\exp\left[\frac{-\beta-m}{2}\frac{y}{h}\right]\right\}\times\exp(-ix\xi)d\xi,\quad y<0, \tag{56}$$

where A_1, A_2, B_1, and B_2 are unknowns, and m is

$$m\equiv m(\xi)=\sqrt{\beta^2+4h^2\xi^2}. \tag{57}$$

The stress τ_y is obtained from (56) by Hooke's law,

$$\tau_y=\mu_0\exp(\beta y/h)\frac{\partial w}{\partial y}$$
$$=\frac{\mu_0\exp(\beta y/h)}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\frac{-\beta+m}{2h}A_1\exp\left[\frac{-\beta+m}{2}\frac{y}{h}\right]+\frac{-\beta-m}{2h}A_2\exp\left[\frac{-\beta-m}{2}\frac{y}{h}\right]\right\}\exp(-ix\xi)d\xi,\quad y>0,$$
$$\tau_y=\mu_0\exp(\beta y/h)\frac{\partial w}{\partial y}$$
$$=\frac{\mu_0\exp(\beta y/h)}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\frac{-\beta+m}{2h}B_1\exp\left[\frac{-\beta+m}{2}\frac{y}{h}\right]+\frac{-\beta-m}{2h}B_2\exp\left[\frac{-\beta-m}{2}\frac{y}{h}\right]\right\}\exp(-ix\xi)d\xi,\quad y<0. \tag{58}$$

By using the boundary conditions (23) to (26), the unknowns A_2, B_1, and B_2 can be expressed by A_1 which is given by

$$A_1=\frac{1}{i\xi}\left\{1-\exp(m)-[1-\exp(-m)]\times\frac{-\beta+m+(\beta+m)\exp(m)}{-\beta+m+(\beta+m)\exp(-m)}\right\}\times\frac{1}{\sqrt{2\pi}}\int_{-a}^{a}\varphi(x')\exp(ix'\xi)dx', \tag{59}$$

where $\varphi(x)$ is the density function defined by

$$\varphi(x)=\frac{\partial}{\partial x}[w(x,0^+)-w(x,0^-)]. \tag{60}$$

Further, the stress τ_y at $y=0$ is expressed by $\varphi(x)$ as

$$\tau_{y|y=0}=\frac{\mu_0}{2\pi}\int_{-a}^{a}\left[\frac{1}{x'-x}+k(x,x',\beta)\right]\varphi(x')dx' \tag{61}$$

where $k(x,x',\beta)$ is given in (31). By substituting the above expression into the boundary condition (24), the integral Eq. (28) is deduced.

References

[1] Orihara, K., 1999, ''Self-Assembly of Functionally Graded Plastics by a Particular Phase Separation of Polymer Blend and Its Applications,'' *Fifth U.S. National Congress on Computational Mechanics, Book of Abstracts*, pp. 399–400.

[2] Pompe, W., Lampenscherf, S., Rössler, S., Scharnweber, D., Weis, K., Worch, H., and Hofinger, J., 1999, ''Functionally Graded Bioceramics,'' Mater. Sci. Forum, **308–311**, pp. 325–330.

[3] Nogata, F., and Takahashi, H., 1995, ''Intelligent Functionally Graded Material: Bamboo,'' Composites Eng., **5**, pp. 743–751.

[4] Hirano, T., Teraki, J., and Nishio, Y., 1999, ''Computational Design for Functionally Graded Thermoelectric Materials,'' Mater. Sci. Forum, **308–311**, pp. 641–646.

[5] Reiter, T., Dvorak, G. J., and Tvergaard, V., 1997, ''Micromechanical Models for Graded Composite Materials,'' J. Mech. Phys. Solids, **45**, pp. 1281–1302.

[6] Reiter, T., and Dvorak, G. J., 1998, ''Micromechanical Models for Graded Composite Materials: II. Thermomechanical Loading,'' J. Mech. Phys. Solids, **46**, pp. 1655–1673.

[7] Gasik, M. M., 1998, ''Micromechanical Modelling of Functionally Graded Materials,'' Comput. Mater. Sci., **13**, pp. 42–55.

[8] Aboudi, J., Pindera, M. J., and Arnold, S. M., 1999, ''Higher-Oder Theory for Functionally Graded Materials,'' Composites, Part B, **30B**, pp. 777–832.

[9] Kawasaki, A., and Watanabe, R., 1987, ''Finite Element Analysis of Thermal Stress of the Metals/Ceramics Multi-Layer Composites with Controlled Compositional Gradients,'' J. Jpn. Inst. Met., **51**, pp. 525–529.

[10] Noda, N., 1999, ''Thermal Stresses in Functionally Graded Materials,'' J. Therm. Stresses, **22**, pp. 477–512.

[11] Praveen, G. N., and Reddy, J. N., 1998, ''Nonlinear Transient Thermoelastic Analysis of Functionally Graded Ceramic-Metal Plates,'' Int. J. Solids Struct., **35**, pp. 4457–4476.

[12] Kolednik, O., 2000, ''The Yield Stress Gradient Effect in Inhomogeneous Materials,'' Int. J. Solids Struct., **37**, pp. 781–808.

[13] Paulino, G. H., Fannjiang, A. C., and Chan, Y. S., 1999, ''Gradient Elasticity Theory for a Mode III Crack in a Functionally Graded Material,'' Mater. Sci. Forum, **308–311**, pp. 971–976.

[14] Cai, H., and Bao, G., 1998, ''Crack Bridging in Functionally Graded Coatings,'' Int. J. Solids Struct., **35**, pp. 701–717.

[15] Erdogan, F., 1995, ''Fracture Mechanics of Functionally Graded Materials,'' Composites Eng., **5**, pp. 753–770.

[16] Jin, Z.-H., and Batra, R. C., 1996, ''Some Basic Fracture Mechanics Concepts in Functionally Graded Materials,'' J. Mech. Phys. Solids, **44**, pp. 1221–1235.

[17] Becker, Jr., T. L., Cannon, R. M., and Ritchie, R. O., 1999, ''A Statistical RKR Fracture Model for the Brittle Fracture of Functionally Graded Materials,'' Mater. Sci. Forum, **308–311**, pp. 957–962.

[18] Rice, J. R., 1968, ''Mathematical Analysis in the Mechanics of Fracture,'' *Fracture—An Advanced Treatise*, Vol. II, H. Liebowitz, ed., Pergamon Press, Oxford, pp. 191–311.

[19] Paulino, G. H., Saif, M. T. A., and Mukherjee, S., 1993, ''A Finite Elastic Body With a Curved Crack Loaded in Anti-Plane Shear,'' Int. J. Solids Struct., **30**, pp. 1015–1037.

[20] Erdogan, F., 1985, ''The Crack Problem for Bonded Nonhomogeneous Materials Under Antiplane Shear Loading,'' ASME J. Appl. Mech., **52**, pp. 823–828.

[21] Ang, W. T., Clements, D. L., and Cooke, T., 1999, ''A Hypersingular Boundary Integral Equation for Antiplane Crack Problems for a Class of Inhomogeneous Anisotropic Elastic Materials,'' Eng. Anal. Boundary Elem., **23**, pp. 567–572.

[22] Atkinson, C., and Chen, C. Y., 1996, ''The Influence of Layer Thickness on the Stress Intensity Factor of a Crack Lying in an Elastic (Viscoelastic) Layer Embedded in a Different Elastic (Viscoelastic) Medium (Mode III Analysis),'' Int. J. Eng. Sci., **34**, pp. 639–658.

[23] Paulino, G. H., and Jin, Z.-H., 2001, ''Correspondence Principle in Viscoelastic Functionally Graded Materials,'' ASME J. Appl. Mech., **68**, pp. 129–132.

[24] Christensen, R. M., 1971, *Theory of Viscoelasticity*, Academic Press, New York.

[25] Hirai, T., 1996, ''Functionally Gradient Materials,'' *Materials Science and Technology, 17B: Processing of Ceramics, Part 2* R. J. Brook, ed., VCH Verlagsgesellschaft mbH, Weinheim, Germany, pp. 292–341.

[26] Ogorkiewicz, R. M., 1970, *Engineering Properties of Thermoplastics*, John Wiley and Sons, London.

[27] Lambros, J., Santare, M. H., Li, H., and Sapna, III, G. H., 1999, ''A Novel Technique for the Fabrication of Laboratory Scale Model Functionally Graded Materials,'' Exp. Mech., **39**, pp. 184–190.

[28] Erdogan, F., Gupta, G. D., and Cook, T. S., 1973, ''Numerical Solution of Singular Integral Equations,'' *Mechanics of Fracture*, Vol. 1, G. C. Sih, ed., Noordhoff, Leyden, pp. 368–425.

[29] Broberg, K. B., 1999, *Cracks and Fracture*, Academic Press, London.

[30] Tada, H., Paris, P., and Irwin, G., 1973, *The Stress Analysis of Cracks Handbook*, Del Research Corporation, Helletown, PA.

Rupture of Thin Power-Law Liquid Film on a Cylinder

Rama Subba Reddy Gorla
Mem. ASME
Department of Mechanical Engineering,
Cleveland State University,
Cleveland, OH 44115
e-mail: gorla@csaxp.csuohio.edu

The dynamic rupture process of a thin power-law type non-Newtonian liquid film on a cylinder has been analyzed by investigating the stability to finite amplitude disturbances. The dynamics of the liquid film is formulated using the balance equations including a body force term due to van der Waals attractions. The governing equation for the film thickness was solved by finite difference method as part of an initial value problem for spatial periodic boundary conditions. A decrease in the cylinder radius will induce a stronger lateral capillary force and thus will accelerate the rupture process. The influence of the power-law exponent on rupture is discussed. [DOI: 10.1115/1.1355033]

Introduction

Study of the rupture of thin liquid films has been motivated by the industrial applications in many disperse and colloid systems in chemical, mechanical and biomedical engineering fields. When a liquid layer becomes ultra thin (100–1000 Å), it becomes unstable. The instability is due to the van der Waals potential and results in the rupture of the layer. Ruckenstein and Jain [1] studied the spontaneous rupture of a liquid film on a planar solid wall. The liquid film was modeled by them as a Navier-Stokes continuum with a potential due to the van der Waals interactions. They used lubrication approximation to obtain the linear dynamic instability results. From this analysis, one can obtain rough estimates for the rupture time, namely, the time needed for the thin film to attain zero thickness at some point. William and Davis [2] examined the nonlinear evolution equation and numerically treated it as an initial value problem with periodic boundary conditions. Their results indicated that the nonlinearities of the system would accelerate the rupture phenomenon.

Cheng and Chang [3] considered the stability of thin liquid film on a cylindrical surface. Brochard [4] discussed the spreading of liquids on thin cylinders and found that for volatile liquids, the vapor would adhere to the cylinders and showed two different ways to prevent the Rayleigh instability from developing on fibers coated with a liquid film.

All the previous studies were concerned with Newtonian fluids. There exist relatively few studies concerning the non-Newtonian fluids, which are important in connection with plastics manufacturing, lubricant performance, applications of paints and movement of biological fluids. Non-Newtonian fluids generally exhibit a nonlinear relationship between shear stress and shear rate. These fluids may be classified as inelastic and viscoelastic. The inelastic fluids may be subdivided as time-dependent fluids and in time-independent fluids. The time-dependent fluids, in turn, are subdivided into two groups: thixotropic and rheopectic. The time-independent fluids can be subdivided into four groups: (1) pseudoplastic, (2) dilatant behavior, (3) Bingham plastic, and (4) pseudoplastic with yield stress. Inelastic time-independent non-Newtonian fluids have received the greatest attention from rheologists which has resulted in the development of a number of equations or models proposed to represent their flow behavior. The Ostwald-de Waele power-law model represents several inelastic time-independent non-Newtonian fluids of practical interest and therefore has been used in this paper. When $n<1$, the model describes pseudoplastic behavior whereas $n>1$ represents dilatant behavior. More details on the classifications of non-Newtonian fluids may be found in Skelland [5]. Sathyagal [6] used a linearized analysis to predict the critical thickness of rupture for a power-law liquid film. Hwang and Chang [7] investigated the nonlinear aspects of the problem.

In practical applications, there are many superslender cylinders such as optical fibers with radii less than 10 μm. If the thickness of a liquid coating film is ultrathin (100–1000 Å), then the van der Waals potential may affect obviously the stability of this film flow system. A review of the literature indicated that no one until now has addressed the question of how the thin power-law type non-Newtonian liquid film on a cylinder ruptures in the presence of van der Waals forces. The present work has been undertaken in order to investigate this problem. We are interested in the specific working regimes of the parameters, where it will be possible to predict rupture or dry out of the thin film. This will be accomplished by solving the equations of thin film motion. A long wave theory is formulated for the nonlinear dynamic instabilities of the thin film. Examples of power-law fluids that might be of interest in the thin film applications are polymers, lubricants, paints, and biological fluids.

Analysis

We consider the flow of a thin liquid film along a horizontal cylinder. We choose x and y directions to be parallel and normal to the cylinder, respectively, as shown in Fig. 1. We assume the characteristic thickness of the film to be h_0 and the length scale parallel to the film to be L. The aspect ratio is given by $\xi = h_0/L$. If we assume that $\xi \ll 1$, we have a thin film. Assuming that the liquid is a power-law fluid, we may write

$$\tau_{ij} = m|\dot{\gamma}_{ij}|^{n-1}\dot{\gamma}_{ij} \tag{1}$$

where τ_{ij} is the stress tensor, $\dot{\gamma}_{ij}$ is the rate of strain tensor, n is the power-law exponent, and m is the viscosity index. We now use the following length scales:

time: $[h_0(\rho h_0^n/m)^{1/(n-2)}]$

length: $[h_0]$

velocity (U_0): $[(\rho h_0^n/m)^{1/(n-2)}]$

pressure and stress: $[\rho(\rho h_0^n/m)^{2/(n-2)}]$.

The liquid layer is assumed thin enough that van der Waals forces are effective ($<\approx 0.1$ μm) and thick enough that a continuum theory of the liquid is applicable. The dimensionless conservation equations of mass, momentum and angular momentum equations are given by

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 7, 1999; final revision, Nov. 2, 2000. Associate Editor: J. T. Jenkins. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

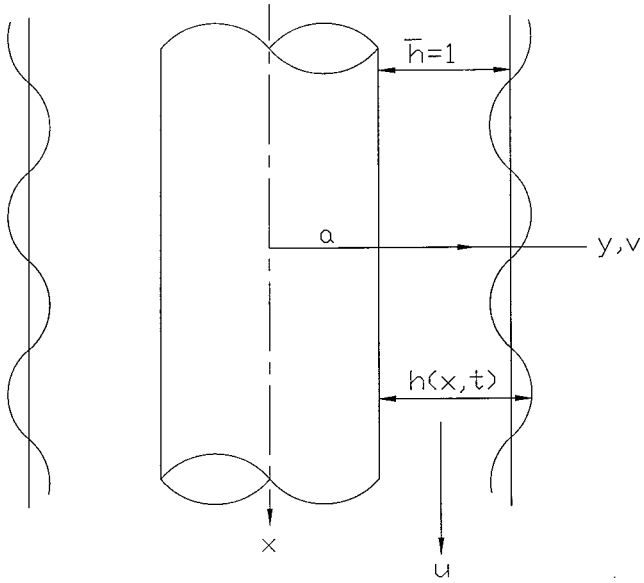


Fig. 1 Flow model for the thin film flow

$$\frac{\partial u}{\partial x}+\left(\frac{1}{a+y}\right)\frac{\partial (a+y)v}{\partial y}=0 \tag{2}$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=-\frac{\partial p}{\partial x}-\frac{\partial \psi}{\partial x}+\frac{1}{(a+y)}\frac{\partial\left[(a+y)\left(\frac{\partial u}{\partial y}\right)^{n}\right]}{\partial y} \tag{3}$$

$$0=-\frac{\partial p}{\partial y}. \tag{4}$$

In the above dimensionless equations, u and v represent the velocity components in x and y directions, respectively, p the pressure and ψ the potential function describing the van der Waals forces. We follow Williams and Davis [2] and write a modified expression for ψ:

$$\psi=Ah^{-3}. \tag{5}$$

In the previous equation, the van der Waals forces are represented through the potential function Ψ and A' is the dimensional Hamaker constant. A is related to the Hamaker constant A' as

$$A=\frac{A'}{6\pi\rho h_0^3 D} \quad \text{where} \quad D=\rho(\rho h_0^n/m)^{2/(n-2)}. \tag{6}$$

The boundary conditions along the solid plane wall are given by

$$y=0:\ u=v=0. \tag{7}$$

At the fluid interface, we have the kinematic condition:

$$y=h(x,t):\quad \frac{\partial h}{\partial t}+u\frac{\partial h}{\partial x}=v. \tag{8}$$

The continuity of tangential stress on the interface requires

$$y=h(x,t):\quad \frac{\partial u}{\partial y}=0. \tag{9}$$

The continuity of normal stress at the interface $y=h(x,t)$ becomes

$$\frac{2\left[\left(\left(\frac{\partial h}{\partial x}\right)^2\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)-\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\frac{\partial h}{\partial x}\right]}{\left[1+\left(\frac{\partial h}{\partial x}\right)^2\right]}$$
$$=-\frac{3S\left[\frac{\partial^2 h}{\partial x^2}-\frac{\left(1+\left(\frac{\partial h}{\partial x}\right)^2\right)}{(a+y)}\right]}{\left[1+\left(\frac{\partial h}{\partial x}\right)^2\right]^{3/2}}-p \tag{10}$$

where

$$S=\sigma/3h_0 D \tag{11}$$

dimensionless surface tension. Our aim here is to solve for the stability of the liquid film while including the effect of van der Waals forces.

We now apply the long-wave theory to study the stability problem. When the layer is thinner than a critical value, small disturbances begin to grow. These waves have wavelengths much larger than the mean thickness of the layer. Defining a small parameter κ that is related to wave number of such disturbances, we may rescale the governing equations by order of magnitude considerations:

$$X=\kappa x;\ \ Y=y;\ \ \tau=\kappa^4 t;\ \ \Psi_0=\kappa^{-2}\psi$$
$$U=\kappa^{-3}u;\ \ V=\kappa^{-4}v;\ \ P=\kappa^{-2}p. \tag{12}$$

We assume that $\partial/\partial X, \partial/\partial Y, \partial/\partial\tau = O(1)$ as $\kappa\to 0$. Given that $U=O(1)$, Eq. (3) indicates that $V=O(\kappa)$. We now let p, $\psi_0=O(1/\kappa)$ as $\kappa\to 0$.

We now assume the following expansions for the flow field:

$$U=U_0+\kappa^2 U_1+\ \ldots$$
$$V=V_0+\kappa^2 V_1+\ \ldots$$
$$P=P_0+\kappa^2 P_1+\ \ldots \tag{13}$$
$$\Psi_0=\kappa^{-2}\psi.$$

The governing equations and the corresponding boundary conditions for the zeroth-order problem may be written as

$$\frac{\partial U_0}{\partial X}+\frac{\partial V_0}{\partial Y}=0 \tag{14}$$

$$n\left(\frac{\partial U_0}{\partial Y}\right)^{n-1}\frac{\partial^2 U_0}{\partial Y^2}=\frac{\partial P_0}{\partial X}+\frac{\partial \Psi_0}{\partial X} \tag{15}$$

$$\frac{\partial P_0}{\partial Y}=0. \tag{16}$$

The boundary conditions are given by

$$Y=0:\ \ U_0=V_0=0 \tag{17}$$

$$Y=h:\ \ \frac{\partial U_0}{\partial Y}=0$$

$$P_0=-3S\frac{\partial^2 h}{\partial X^2}-\frac{3S(a-h)}{a^2}$$

$$\frac{\partial h}{\partial \tau}+U_0\frac{\partial h}{\partial X}=V_0. \tag{18}$$

The solutions for the velocity field are given by

$$U_0=\left[\frac{n}{(n+1)}\frac{\partial P_0'}{\partial X}^{1/n}\right]\{-(h-Y)^{(n+1)/n}+h^{(n+1)/n}\} \tag{19}$$

$$V_0 = -\frac{1}{(n+1)}\left(\frac{\partial P_0'}{\partial X}\right)^{1}{}^{(1-n)/n}\frac{\partial^2 P_0'}{\partial X^2}$$
$$\times\left[\frac{n}{2n+1}(h-Y)^{(2n+1)/n}+h^{(n+1)/n}Y-\frac{nh^{(2n+1)/n}}{(2n+1)}\right]$$
$$+\frac{n}{(n+1)}\left(\frac{\partial P_0'}{\partial X}\right)^{1/n}\frac{\partial h}{\partial X}$$
$$\times\left\{h^{(n+1)/n}-\left((h-Y)^{(n+1)/n}+\frac{n+1}{n}h^{1/n}Y\right)\right\} \quad (20)$$

where

$$P_0' = -(P_0+\psi_0). \quad (21)$$

$$P_0 = \left(-3S\frac{\partial^2 h}{\partial X^2}-\frac{3S(a-h)}{a^2}\right) \quad (22)$$

Similarly, expressions u_1, v_1, and p_1 may be derived. These expressions are not used in the computations and they are very long. Therefore, they are not reproduced here. Using equations (14)–(22) we may show that the leading order evolution equation for the film rupture is given by

$$\frac{\partial h}{\partial \tau}+\left\{3S\frac{\partial^3 h}{\partial X^3}+3Sa^{-2}+\frac{3A}{h^4}\frac{\partial h}{\partial X}\right\}^{1/n}h^{(n+1)/n}\frac{\partial h}{\partial X}$$
$$=-\frac{1}{(2n+1)}\left(3S\frac{\partial^3 h}{\partial X^3}+3Sa^{-2}+\frac{3A}{h^4}\frac{\partial h}{\partial X}\right)^{(1-n)/n}$$
$$\times\left(3S\frac{\partial^4 h}{\partial X^4}+3Sa^{-2}+3A\frac{\partial}{\partial X}\left(\frac{1}{h^4}\frac{\partial h}{\partial X}\right)\right)h^{(2n+1)/n} \quad (23)$$

subject to initial conditions

$$h(X,0)=f(X). \quad (24)$$

Equation (23) governs long wave interfacial disturbances to the static film (having $h=1$) subject to van der Waals attractions.

Results and Discussion

The nonlinear partial differential Eq. (23) was solved numerically using the finite difference method. Central differences were used for space variable and the midpoint rule was used for time. The Newton-Raphson method was used to solved the resulting system of difference equations. The problem was treated as an initial value problem with spatial periodic boundary conditions within the interval $0<X<2\pi/q_m$.

In order to obtain a solution independent of the grid size, several computational runs were performed to obtain the optimum step sizes in X and τ directions. The optimization procedure of the grid size includes computing the spatial film thickness distribution at an arbitrary time, employing a given number of grid points in spatial direction. After that the number of grid points is increased gradually, each time, a computer run was performed to compute the film thickness profile. A residue is defined as the absolute difference in film thickness between the two runs. The procedure is continued until the residue approaches a value less than 1×10^{-4}. At this point the spatial grid size is fixed. A similar procedure was followed to choose the optimum time step. Based on these calculations, we used spatial grid points $N=50$ and time steps $\Delta\tau=0.01$ in all the computations. The initial condition was given by

$$h(X,0)=1+H_0\sin(q_m X), \quad 0\leq X\leq 2\pi/q_m. \quad (25)$$

The following parameters were used in the numerical calculations to describe both pseudoplastic and dilatant materials:

$$H_0=0.1,0.25;\ \ a=80,100,150;\ \ n=2,1,0.9,0.75,0.5;$$

$$A=0.001,0.002,0.003;\ \ S=1,2,3.$$

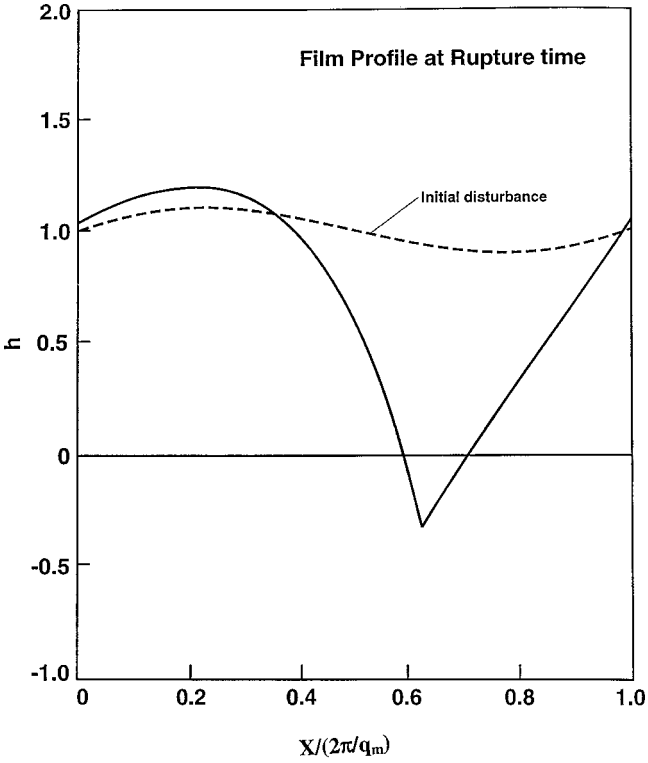


Fig. 2 Film thickness distribution at the rupture time

Figure 2 shows the initial disturbance introduced and the film profile at the time of film rupture. It may be noted that the film thickness is zero at rupture and negative thickness has no physical meaning. Figure 3 shows that the rupture time increases with cylinder radius.

Figure 4 displays the timewise variation of minimum film thickness. We observe that the rupture time for the Newtonian

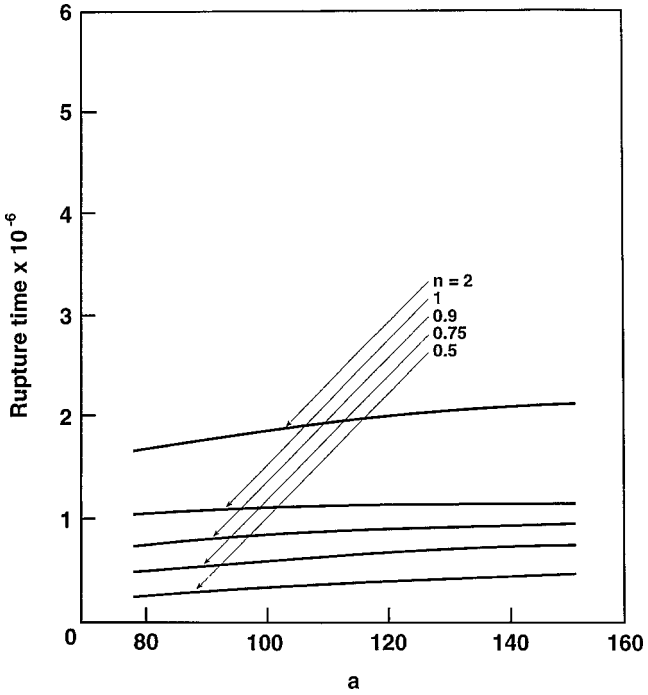


Fig. 3 Rupture time versus cylinder radius

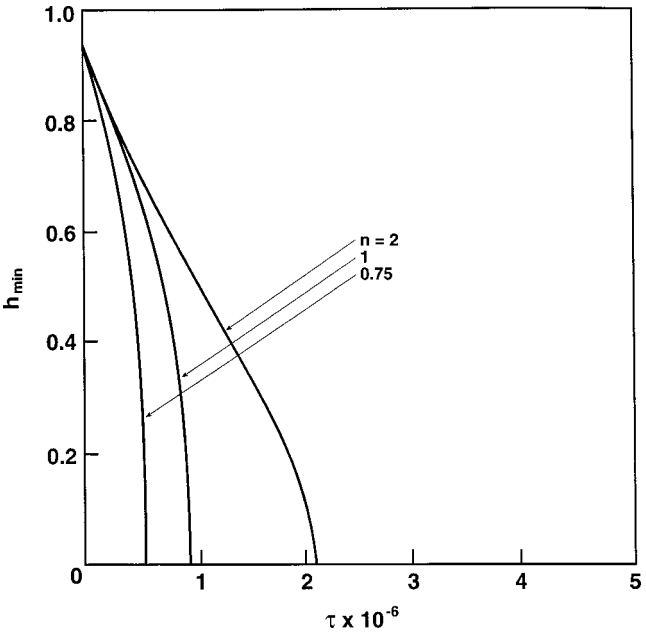


Fig. 4 Minimum film thickness versus time

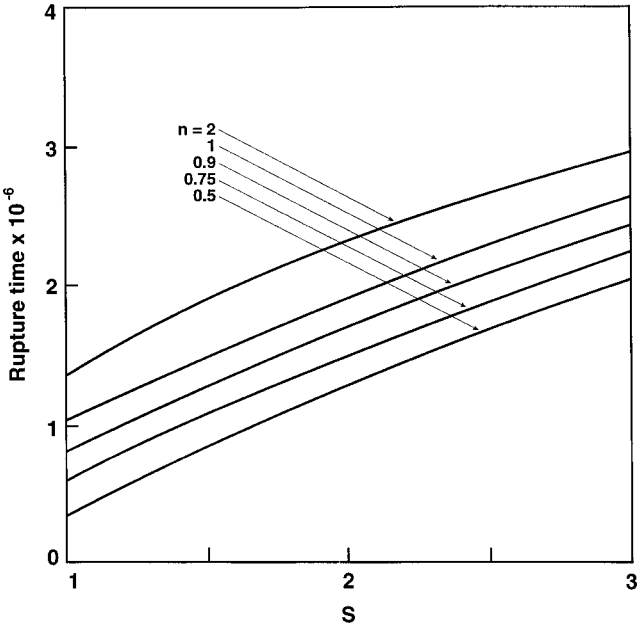


Fig. 6 Rupture time versus *S*

fluid ($n=1$) is smaller than that for the dilatant fluids ($n>1$) but greater than the corresponding value for the pseudoplastic fluids ($n<1$). As n increases, the rupture time increases.

Figure 5 shows the variation of the rupture time versus A. The rupture time decreases as A increases or the van der Waals potential increases. As n increases, the rupture time increases.

Figure 6 shows that as the surface tension increases, the rupture time increases. The strong lateral capillary force induces an acceleration effect on the rupture process as the magnitude of the surface tension parameter, S increases. As n increases, the rupture time increases.

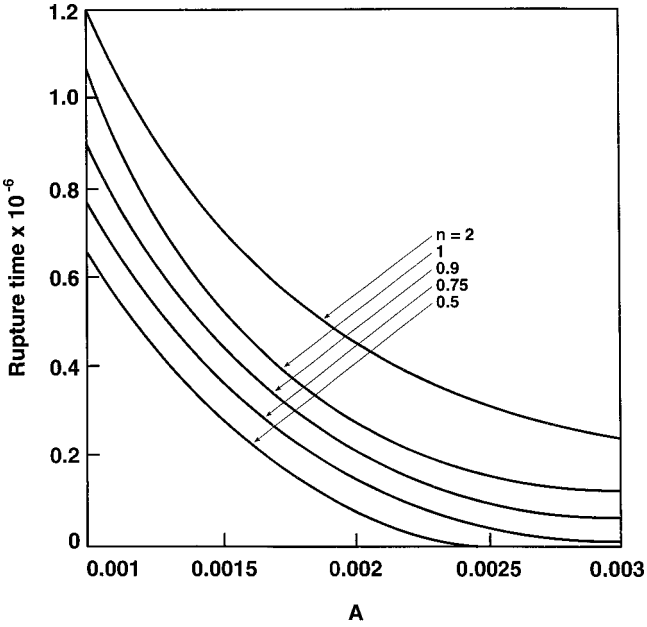


Fig. 5 Rupture time versus *A*

Concluding Remarks

In this paper, we have formulated a long-wave theory for the nonlinear dynamic instabilities of a thin power-law type non-Newtonian liquid film on a cylinder. Numerical solutions are obtained for the simplified form of the equations governing the dynamics of the liquid film. As the radius of the cylinder increases, the rupture time increases. The rupture time increases with surface tension parameter and decreases with increasing van der Waals force potential. The rupture times for the dilatant fluids ($n>1$) are higher than that of Newtonian fluid where as pseudoplastic fluids ($n<1$) are associated with smaller rupture times. Therefore, non-Newtonian fluids find an important application where it may be desirable to maintain the stability of thin films over longer or shorter periods of time when compared to Newtonian fluids.

Acknowledgments

The author is grateful to the helpful comments of the reviewers.

References

[1] Ruckenstein, E., and Jain, R. K., 1974, ''Spontaneous Rupture of Thin Liquid Films,'' Chemical Society, Faraday Transactions, **270**, pp. 132–146.
[2] William, M. B., and Davis, S. H., 1982, ''Nonlinear Theory of Film Rupture,'' J. Colloid Interface Sci., **90**, pp. 220–228.
[3] Cheng, M., and Chang, H. C., 1992, ''Stability of Axisymmetric Waves on Liquid Films Flowing Down a Vertical Column to Azimuthal and Streamwise Disturbances,'' Chem. Eng. Commun., **118**, pp. 327–340.
[4] Brochard, F., 1986, ''Spreading of Liquid Drops on Thin Cylinders: The Machon/Droplet Transition,'' J. Chem. Phys., **84**, p. 4664.
[5] Skelland, A. H. P., 1967, *Non-Newtonian Flow and Heat Transfer*, John Wiley and Sons, New York.
[6] Sathyagal, A. N., and Narasimhan, G., 1992, ''On Rupture of a Thinning Film of Non-Newtonian Power Law Fluid,'' Chem. Eng. Commun., **111**, pp. 161–166.
[7] Hwang, C. C., and Chang, S. H., 1993, ''Rupture Theory of Thin Power Law Liquid Films,'' J. Appl. Phys., **74**, pp. 2965–2967.

Thermal Deformation of Initially Curved Substrates Coated by Thin Inhomogeneous Layers

A. Wikström

P. Gudmundson
Mem. ASME

Department of Solid Mechanics,
Royal Institute of Technology (KTH),
SE-100 44 Stockholm, Sweden

Thermal curvature changes and membrane strains are analyzed for elastic shallow shell substrates which are coated by thin, generally inelastic, inhomogeneous and anisotropic layers. The analysis is restricted to linear kinematics. It is shown that the deformation is governed by the corresponding solution for a flat substrate and a correction due to the initial curvature. The correction is determined from a shallow shell problem for the bare substrate with a loading expressed by the coefficients of thermal curvature for the substrate/layer system. For constant initial curvatures, certain analytic solutions are presented. For situations when the initial deflection of the substrate is much larger than the substrate thickness, a boundary layer solution is derived. In the particular case of a circular isotropic substrate with a spherical initial curvature and a coating of arbitrary anisotropy, the solution is presented in closed form. For nonflat substrates, measured curvatures can generally not be used to extract layer stresses without a proper compensation for the initial curvature. In the paper, it is explicitly presented how to accurately compensate for a spherical initial curvature. The results are particularly discussed in relation to curvature measurements on Silicon substrates. [DOI: 10.1115/1.1357169]

1 Introduction

Many engineering applications employ initially curved shells which are coated by thin, generally inhomogeneous and anisotropic layers. Most common examples include thermal-barrier or wear-resistant coatings and organic paints deposited on curved surfaces. In addition, nominally flat substrates, such as Si wafers used to fabricate microelectronic chips, may in fact exhibit nonnegligible initial curvatures in response to some processing and manufacturing steps. The subsequent deposition of metal interconnects and passivating material then builds up a shallow shell with a thin coating which belongs to the class of problems that are investigated in the present paper. While there exists considerable information on the mechanics of flat substrates, the deformation of initially curved substrates with anisotropic layers appears to be relatively unexplored.

In the present paper, the attention is directed towards analyses within the context of small deformations of elastic shallow shell substrates coated with thin anisotropic and possibly inelastic layers. First it is shown that the analysis can be reduced to a problem which only involves the initially curved substrate without coating with a loading in terms of the thermal curvature coefficients of the substrate/coating system. All information about the thin layer such as elastic, plastic, and creep behavior are then contained within the coefficients of thermal curvatures. For constant initial curvatures, closed-form solutions are presented for certain parameter combinations. Since one of the most common methods used to perform mechanical testing of thin films (Flinn et al. [1]) and unpassivated lines (Yeo et al. [2]) is to measure curvature changes during thermal cycling, the implications of performing curvature measurements on nonflat substrates under the assumption of initial flatness is thoroughly investigated. Specific applications of the results to initial spherical curvatures are addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 8, 2000; final revision, October 19, 2000. Associate Editor: S. Kyriakides. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2 Theoretical Basis

Elastic shallow shell substrates with thin coatings, are modeled within the context of linear kinematics. The substrate/layer system is assumed to have free boundaries and it is subjected to a temperature load. Due to thermal expansion mismatch, stresses and curvatures will develop. The constitutive relation for a shallow shell substrate with a thin layer deposit can be treated as that of a homogenized anisotropic plate.

2.1 Shallow Shell Equations. A Cartesian coordinate system x_α is introduced. The middle surface of the undeformed shell is expressed by $z=-\tilde{\kappa}_{\alpha\beta}x_\alpha x_\beta/2$, where $\tilde{\kappa}_{\alpha\beta}$ is the initial curvature of the shell and the summation convention is utilized for Greek subscripts which range from 1 to 2. The condition for a shell to be considered sufficiently shallow is $z_{,\alpha}z_{,\alpha}\ll 1$. As a starting point, the differential equations for an anisotropic linear elastic shallow shell are considered (Flügge [3] and Leissa et al. [4]).

The equilibrium equations for vanishing body forces, such as gravity, can be expressed as

$$N_{\alpha\beta,\beta}=0 \quad M_{\alpha\beta,\alpha\beta}-N_{\alpha\beta}\tilde{\kappa}_{\alpha\beta}=0 \tag{1}$$

where $N_{\alpha\beta}$ and $M_{\alpha\beta}$ denote membrane forces and moments per unit length. The homogenized constitutive law may be written as

$$\begin{aligned} N_{\alpha\beta}&=A_{\alpha\beta\gamma\delta}(\varepsilon_{\gamma\delta}-\alpha_{\gamma\delta}T)+B_{\alpha\beta\gamma\delta}(\kappa_{\gamma\delta}-\beta_{\gamma\delta}T)\\ M_{\alpha\beta}&=B_{\alpha\beta\gamma\delta}(\varepsilon_{\gamma\delta}-\alpha_{\gamma\delta}T)+D_{\alpha\beta\gamma\delta}(\kappa_{\gamma\delta}-\beta_{\gamma\delta}T)\end{aligned} \tag{2}$$

where $A_{\alpha\beta\gamma\delta}$, $B_{\alpha\beta\gamma\delta}$, and $D_{\alpha\beta\gamma\delta}$ represent stiffness tensors of the homogenized shell element. The parameters $\alpha_{\alpha\beta}$ and $\beta_{\alpha\beta}$ denote the coefficients of thermal expansion and thermal curvature of the homogenized shell element and T a constant temperature change. The kinematics of the shallow shell are described by

$$\begin{aligned} \varepsilon_{\gamma\delta}&=\tfrac{1}{2}(u_{\gamma,\delta}+u_{\delta,\gamma})+w\tilde{\kappa}_{\gamma\delta}\\ \kappa_{\gamma\delta}&=-w_{,\gamma\delta}\end{aligned} \tag{3}$$

where $\varepsilon_{\gamma\delta}$ represent in-plane strains, $\kappa_{\gamma\delta}$ curvatures, u_δ midsurface in-plane displacements, and w the out-of-surface displacement. Strain compatibility requires that

$$e_{\alpha\gamma}e_{\beta\delta}(\varepsilon_{\gamma\delta}-w\tilde{\kappa}_{\gamma\delta})_{,\alpha\beta}=0 \tag{4}$$

where $e_{\alpha\beta}$ represents the two dimensional permutation tensor.

The boundary conditions for a free edge read

$$N_{\alpha\beta}n_\beta=0, \quad M_{\alpha\beta}n_\alpha n_\beta=0, \quad Q_\alpha n_\alpha+M_{nt,t}=0 \tag{5}$$

where n_α, t_α are normal and tangent directions to the boundary, respectively, and $Q_\alpha=M_{\alpha\beta,\beta}$ is the equivalent transverse shear force. All tensorial quantities in Eqs. (1)–(5) are given in a curvilinear coordinate system defined in Flügge [3] which is specified by the following transformation:

$$d\hat{u}_\alpha=du_\alpha-z_{,\alpha}dw, \quad d\hat{w}=du_{,\alpha}z_{,\alpha}+dw \tag{6}$$

where $d\hat{u}_\alpha$ and $d\hat{w}$ denote components parallel to the fixed Cartesian coordinate system in which the middle surface of the shell is expressed. Note that if $\tilde{\kappa}_{\gamma\delta}$ vanish, the problem reduces to the flat-plate problem.

2.2 Thin Layer Approximation. The thermoelastic properties given in Eq. (2) for a substrate with a thin layer deposit can be expanded as a Taylor expansion in $\delta=t/h\ll 1$ where t is the thickness of the thin layer and h the substrate thickness. Only terms up to first order are considered here. The expanded thermoelastic properties can be written in matrix form as

$$\begin{bmatrix} \mathbf{N}^0+\delta\mathbf{N}^1 \\ \mathbf{M}^0+\delta\mathbf{M}^1 \end{bmatrix}=\begin{bmatrix} \mathbf{A}^0+\delta\mathbf{A}^1 & \mathbf{B}^0+\delta\mathbf{B}^1 \\ \mathbf{B}^0+\delta\mathbf{B}^1 & \mathbf{D}^0+\delta\mathbf{D}^1 \end{bmatrix} \times\left(\begin{bmatrix} \boldsymbol{\varepsilon}^0+\delta\boldsymbol{\varepsilon}^1 \\ \boldsymbol{\kappa}^0+\delta\boldsymbol{\kappa}^1 \end{bmatrix}-T\begin{bmatrix} \boldsymbol{\alpha}^0+\delta\boldsymbol{\alpha}^1 \\ \boldsymbol{\beta}^0+\delta\boldsymbol{\beta}^1 \end{bmatrix}\right) \tag{7}$$

where $\mathbf{A}^0$, $\mathbf{B}^0$, $\mathbf{D}^0$, $\boldsymbol{\alpha}^0$, $\boldsymbol{\beta}^0$ are properties for a bare substrate without layer. For a substrate which is homogeneous in the thickness direction $\mathbf{B}^0$, $\boldsymbol{\beta}^0$ vanish. Hence

$$\varepsilon^0_{\alpha\beta}=\alpha^0_{\alpha\beta}T, \quad w^0=\kappa^0_{\alpha\beta}=M^0_{\alpha\beta}=N^0_{\alpha\beta}=0 \tag{8}$$

fulfill the equilibrium Eq. (1) and compatibility for the kinematic relations stated in Eq. (4) as well as the boundary conditions in Eq. (5). It is noted that the solution to the zeroth-order problem also is the solution to the flat bare substrate problem. Since the zeroth-order solution is a pure in-plane strain, the second-order stresses and strains that develop in the thin layer as a result of a temperature change can be solved as a separate problem represented by a thin coating on an infinite elastic half-space. By computing the volume average stresses in the separate problem, the coefficients of thermal curvature $\boldsymbol{\beta}^1$ ($\boldsymbol{\beta}^0=0$) may be computed through moment equilibrium (Stoney formula), see (Wikström et al. [5])

$$\langle\sigma_{\alpha\beta}\rangle=-\frac{h^2}{6t}C^0_{\alpha\beta\gamma\delta}\beta^1_{\gamma\delta}T \tag{9}$$

where $\langle\sigma_{\alpha\beta}\rangle$ represents volume-averaged stress changes in the layer of thickness t. The tensor $C^0_{\alpha\beta\gamma\delta}$ represents the plane stress stiffness tensor of the bare substrate, and the middle surface z is chosen as the midplane of the substrate. For a flat substrate with a thin coating, the thermal curvature changes $\kappa^1_{\alpha\beta}$ will equal $\beta^1_{\alpha\beta}T$. Hence a measurement of curvatures will through Eq. (9) enable estimates of volume average layer stresses (Flinn et al. [1] and Yeo et al. [2]). Therefore, provided that a sufficiently small representative part of the layer can be chosen, the effects of plasticity, creep, crystal directions and so on may be included in $\boldsymbol{\beta}^1$.

In order to analyze the first-order problem, the following substitutions are made:

$$\varepsilon^1_{\alpha\beta}=T\alpha^1_{\alpha\beta}+\varepsilon^c_{\alpha\beta} \quad N^1_{\alpha\beta}=N^c_{\alpha\beta} \quad \kappa^1_{\alpha\beta}=T\beta^P_{\alpha\beta}+\kappa^c_{\alpha\beta}$$

$$M^1_{\alpha\beta}=-T\frac{h^3}{12}C^0_{\alpha\beta\gamma\delta}(\beta^1_{\gamma\delta}-\beta^P_{\gamma\delta})+M^c_{\alpha\beta} \tag{10}$$

$$w^1=-\frac{T}{2}\beta^P_{\alpha\beta}x_\alpha x_\beta+w^c$$

where $\beta^P_{\alpha\beta}$ represents a prescribed tensor that is independent of x_α and which at this point may be arbitrarily chosen. For a constant initial curvature, the remaining system of equations that is valid in the interior of the shell becomes

$$N^c_{\alpha\beta,\beta}=0 \quad M^c_{\alpha\beta,\alpha\beta}-N^c_{\alpha\beta}\tilde{\kappa}_{\alpha\beta}=0 \tag{11}$$

$$N^c_{\alpha\beta}=hC^0_{\alpha\beta\gamma\delta}\varepsilon^c_{\gamma\delta} \quad M^c_{\alpha\beta}=\frac{h^3}{12}C^0_{\alpha\beta\gamma\delta}\kappa^c_{\gamma\delta} \tag{12}$$

$$e_{\alpha\gamma}e_{\beta\delta}\left(\varepsilon^c_{\gamma\delta,\alpha\beta}-\left(w^c-\frac{T}{2}\beta^P_{\theta\varphi}x_\theta x_\varphi\right)_{,\alpha\beta}\tilde{\kappa}_{\gamma\delta}\right)=0 \tag{13}$$

$$\kappa^c_{\alpha\beta}=-w^c_{,\alpha\beta} \tag{14}$$

and the boundary conditions can be expressed as

$$N^c_{\alpha\beta}n_\beta=0$$

$$Q^c_\alpha n_\alpha+M^c_{nt,t}=T\frac{h^3}{12}C^0_{\alpha\beta\gamma\delta}(n_\alpha t_\beta)_{,\varphi}t_\varphi(\beta^1_{\gamma\delta}-\beta^P_{\gamma\delta}) \tag{15}$$

$$n_\alpha n_\beta M^c_{\alpha\beta}=T\frac{h^3}{12}C^0_{\alpha\beta\gamma\delta}n_\alpha n_\beta(\beta^1_{\gamma\delta}-\beta^P_{\gamma\delta}).$$

Three observations (that would be valid also for a varying initial curvature) can now be made: (a) It is sufficient to consider the stiffness matrix of the bare substrate only. (b) Regardless of how $\beta^P_{\alpha\beta}$ is chosen, the coefficients of thermal curvature will represent an inhomogeneous term somewhere in Eqs. (11)–(15) and since the system is linear, the solution will be linear in the inhomogeneous term. (c) The membrane and bending state decouple for zero initial curvatures (flat-plate problem) and as will be shown subsequently also for very large initial curvatures (a large initial curvature within the shallow shell theory correspond to initial out-of-plane deflections which are much larger than the shell thickness as well as much smaller than in-plane dimensions).

First, let $\beta^P_{\alpha\beta}=\beta^1_{\alpha\beta}$, this means that w^c represents a correction to the flat-plate solution. The boundary conditions become homogeneous and the inhomogeneous term appears in the interior description of the plate (Eq. (13)). A physical interpretation of the correction can be made. It is seen that if a constant C is chosen as $C=(\beta^1_{\alpha\alpha}\tilde{\kappa}_{\beta\beta}-\beta^1_{\alpha\beta}\tilde{\kappa}_{\alpha\beta})$, the following equality holds:

$$e_{\alpha\gamma}e_{\beta\delta}\tilde{\kappa}_{\gamma\delta}(\beta^1_{\theta\varphi}x_o x_\theta x_\varphi/2)_{,\alpha\beta}=e_{\alpha\gamma}e_{\beta\delta}(Cx_\varphi x_\varphi\delta_{\gamma\delta}/4)_{,\alpha\beta}. \tag{16}$$

Hence, the inhomogeneous term in Eq. (13) may be transferred to Eq. (12) by the substitution $\varepsilon^c_{\gamma\delta}=\varepsilon^d_{\gamma\delta}-TCx_\alpha x_\alpha\delta_{\gamma\delta}/4$. The correction problem (save for the resulting strain) is therefore equivalent to a radially symmetric temperature distribution around some arbitrarily chosen origin with a constant isotropic coefficient of thermal expansion which is proportional to C. This analogy makes it very easy to solve the general anisotropic problem with commercially available finite element codes. As a special case it is observed that $w^c=0$ and $\varepsilon^c_{\alpha\beta}=0$ represent the exact solution when

$$e_{\alpha\gamma}e_{\beta\delta}\left(-\frac{T}{2}\beta^1_{\theta\varphi}x_\theta x_\varphi\right)_{,\alpha\beta}\tilde{\kappa}_{\gamma\delta}=T(\beta^1_{\alpha\alpha}\tilde{\kappa}_{\beta\beta}-\beta^1_{\alpha\beta}\tilde{\kappa}_{\alpha\beta})=0. \tag{17}$$

As an example of this situation one could consider an initial cylindrical shape ($\tilde{\kappa}_{11}\neq 0$, $\tilde{\kappa}_{22}=\tilde{\kappa}_{12}=0$) and a thin layer consisting of a periodic pattern of unpassivated parallel lines with a geometry such that ($\beta^1_{11}\neq 0$, $\beta^1_{22}=\beta^1_{12}=0$). This situation can arise if the lines are oriented along the x_1-direction and their cross-sectional shape is chosen appropriately, see Wikström et al. [5].

Secondly, $\beta^P_{\alpha\beta}$ may be chosen such that the deflection $w=-T\beta^P_{\alpha\beta}x_\alpha x_\beta/2$ fully represents the behavior in the interior of the shell for large initial curvatures. The inhomogeneous term

then only appears in the boundary conditions. In this case, w^c and $\varepsilon^c_{\alpha\beta}$ represent a boundary layer correction to the solution for very large curvatures. This boundary layer solution is presented in closed form in the Appendix. However, first the proper choice of $\beta^P_{\alpha\beta}$ must be determined. For a given initial curvature, two constant nonzero linearly independent solutions $\kappa^{C1}_{\alpha\beta}$ and $\kappa^{C2}_{\alpha\beta}$ can always be chosen such that they fulfill the compatibility equation (cf. Eq. (17)):

$$\kappa^{C1,C2}_{\alpha\alpha}\tilde{\kappa}_{\beta\beta}-\kappa^{C1,C2}_{\alpha\beta}\tilde{\kappa}_{\alpha\beta}=0. \tag{18}$$

They can be used to represent the solution that is valid in the inner part of the shell. Once the solutions have been chosen, they can be ortho-normalized such that

$$\frac{h^3}{12}\kappa^{C1}_{\alpha\beta}C^0_{\alpha\beta\gamma\delta}\kappa^{C1}_{\gamma\delta}=1 \quad \frac{h^3}{12}\kappa^{C2}_{\alpha\beta}C^0_{\alpha\beta\gamma\delta}\kappa^{C2}_{\gamma\delta}=1 \quad \kappa^{C1}_{\alpha\beta}C^0_{\alpha\beta\gamma\delta}\kappa^{C2}_{\gamma\delta}=0. \tag{19}$$

The most general compatible curvatures that fulfill Eq. (18) may then be expressed as

$$\kappa^{\infty}_{\alpha\beta}=T(D_1\kappa^{C1}_{\alpha\beta}+D_2\kappa^{C2}_{\alpha\beta}). \tag{20}$$

In order to determine D_1 and D_2, the potential energy is considered. For large initial curvatures, the contribution from the boundary layer can be neglected. The potential energy may then be written as

$$U=\frac{1}{2}\int_S \frac{h^3}{12}\kappa^{\infty}_{\alpha\beta}C^0_{\alpha\beta\gamma\delta}(\kappa^{\infty}_{\gamma\delta}-\beta^1_{\gamma\delta}T)dS \tag{21}$$

where S denotes the arbitrarily shaped shallow midsurface of the shell. Minimizing the potential energy with respect to D_1 and D_2 yields

$$D_1=\frac{h^3}{12}\kappa^{C1}_{\alpha\beta}C^0_{\alpha\beta\gamma\delta}\beta^1_{\gamma\delta} \quad D_2=\frac{h^3}{12}\kappa^{C2}_{\alpha\beta}C^0_{\alpha\beta\gamma\delta}\beta^1_{\gamma\delta}. \tag{22}$$

Thus, in order to divide the solution into an interior part and a boundary layer, the prescribed coefficient of thermal curvature should be chosen as

$$\beta^P_{\alpha\beta}=D_1\kappa^{C1}_{\alpha\beta}+D_2\kappa^{C2}_{\alpha\beta}. \tag{23}$$

The complementary boundary layer solution is presented in the Appendix for the case of an anisotropic shell of arbitrary in-plane shape and large initial curvature.

The total solution can be assembled as a sum of solutions to the problem of zeroth and first-order, respectively. The error in this approximation is of order $O(\delta^2)$. Observe that so far, no restriction is placed on the type of layer used or the shape of the boundary.

2.3 Substrates with Spherical Initial Curvature. An isotropic spherical cap coated by a thin, possibly inelastic, anisotropic layer is now considered. The aim is to determine how an initially curved, stress-free substrate with free boundaries reacts to temperature changes in comparison to an initially flat substrate. The dimensions of the shell are defined in Fig. 1. The solution to the problem of order zero is given by Eq. (8). The objective is therefore to find a solution to the first-order problem defined by Eqs. (10)–(15).

The initial curvature of a spherical cap can be expressed as

$$\tilde{\kappa}_{\alpha\beta}=\tilde{\kappa}\delta_{\alpha\beta}=2\frac{z_0}{a^2}\delta_{\alpha\beta} \tag{24}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta. The previously introduced tensor $\beta^P_{\alpha\beta}$ is chosen according to one of the two limiting cases discussed in Section 2.2. The prescribed coefficient of thermal expansion for the two alternatives can be worked out as

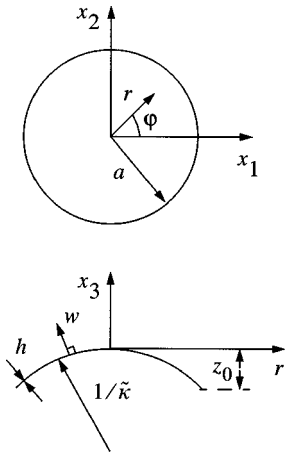


Fig. 1 Geometry of a substrate with spherical initial curvature. The curvilinear coordinate system is indicated.

$$\begin{aligned}\beta^P_{\alpha\beta}&=\beta^1_{\alpha\beta} \quad \text{(zero initial curvature)}\\ \beta^P_{\alpha\beta}&=\beta^1_{\alpha\beta}-\tfrac{1}{2}\beta^1_{\gamma\gamma}\delta_{\alpha\beta} \quad \text{(very large initial curvature)}.\end{aligned} \tag{25}$$

The difference between the two alternatives is just a radially symmetric term that may be added or subtracted afterwards. The choice is arbitrary, however $\beta^P_{\alpha\beta}=\beta^1_{\alpha\beta}$ is selected here since a correction to the flat plate solution may have more physical significance. The boundaries are free and the only inhomogeneous term appears in Eq. (13) and it reduces to

$$T(\beta^1_{\alpha\alpha}\tilde{\kappa}_{\beta\beta}-\beta^1_{\alpha\beta}\tilde{\kappa}_{\alpha\beta})=T\tilde{\kappa}\beta^1_{\alpha\alpha}. \tag{26}$$

It is observed that even if arbitrary coefficients of thermal curvature are permitted, the correction problem exhibits radial symmetry since $\beta^1_{\alpha\alpha}$ is constant and invariant during a change of coordinates. The angular dependence of the solution to this first-order problem is thus entirely contained within the prescribed part of the solution. This makes it possible to find a closed-form solution to the problem.

A stress function is introduced such that

$$N^c_{\alpha\beta}=e_{\alpha\gamma}e_{\beta\delta}\Phi^c_{,\gamma\delta}. \tag{27}$$

For a constant temperature change, the first-order problem given by Eqs. (11)–(14) may for an isotropic substrate be reduced to

$$\begin{aligned}\nabla^2\nabla^2\Phi^c-\tilde{\kappa}Eh\nabla^2w^c&=-\tilde{\kappa}EhT\beta^1_{\alpha\alpha},\\ Eh^3\nabla^2\nabla^2w^c+c^4\tilde{\kappa}\nabla^2\Phi^c&=0\end{aligned} \tag{28}$$

with boundary conditions expressed in polar coordinates

$$w^c(0)=N^c_r(a)=M^c_r(a)=0 \quad M^c_r(0),N^c_r(0) \text{ finite} \tag{29}$$

where ∇^2 is the Laplace operator and the constant c is defined as

$$c^2=\sqrt{12(1-\nu^2)} \tag{30}$$

and where E and ν represent Young's modulus and Poisson's ratio of the bare substrate. Due to the radial symmetry, the Laplace operator reduces to

$$\nabla^2=\frac{d^2}{dr^2}+\frac{1}{r}\frac{d}{dr} \tag{31}$$

and radial moments and membrane forces per unit length become

$$M^c_r=-\frac{Eh^3}{c^4}\left(\frac{d^2w^c}{dr^2}+\nu\frac{1}{r}\frac{dw^c}{dr}\right), \quad N^c_r=\frac{1}{r}\frac{d\Phi^c}{dr}. \tag{32}$$

It can be shown that with radial symmetry and zero external loads, the following conditions are automatically satisfied:

$$\left.\frac{dw^c}{dr}\right|_0 = Q_r^c(a) = N_{r\varphi}^c(a) = M_{r\varphi}^c(a) = 0. \tag{33}$$

It should be noted that in contradiction to Leissa et al. [4], Reissner [6] includes a term $a\tilde{\kappa}N_r^c(a)$ in $Q_r^c(a)$, this term does not matter here because $N_r^c(a)=0$. Following the work of Reissner [6–8] the total solution to Eqs. (28)–(29) subject to the continuity requirement and displacement condition at the center may be expressed as

$$w^c = \tfrac{1}{2}a^2 T\beta_{\alpha\alpha}^1 \overline{f}(\overline{r}), \quad \Phi^c = \tfrac{1}{2}Eh^2a^2T\beta_{\alpha\alpha}^1\overline{s}(\overline{r}) \tag{34}$$

where $\overline{r}=r/a$. The dimensionless functions $\overline{f}(\overline{r})$ and $\overline{s}(\overline{r})$ may be expressed in terms of the Kelvin functions $\mathrm{ber}_n(x)$ and $\mathrm{bei}_n(x)$ as

$$\begin{aligned}\overline{f}(\overline{r}) &= C_1(\mathrm{ber}_0(\zeta\overline{r})-1) + C_2\,\mathrm{bei}_0(\zeta\overline{r}) + \tfrac{1}{2}\overline{r}^2 \\ \overline{s}(\overline{r}) &= \frac{1}{c^2}(C_1\,\mathrm{bei}_0(\zeta\overline{r}) - C_2\,\mathrm{ber}_0(\zeta\overline{r}))\end{aligned} \tag{35}$$

where the dimensionless parameter ζ is defined as

$$\zeta = c\sqrt{\frac{a^2\tilde{\kappa}}{h}} = c\sqrt{2\left(\frac{z_0}{h}\right)}. \tag{36}$$

Applying the boundary conditions given by Eq. (29), the constants C_1 and C_2 can be determined as

$$\begin{aligned}C_1 &= \frac{\sqrt{2}(1+\nu)}{\zeta\lambda}\{\mathrm{ber}_1(\zeta)+\mathrm{bei}_1(\zeta)\} \\ C_2 &= \frac{-\sqrt{2}(1+\nu)}{\zeta\lambda}\{\mathrm{ber}_1(\zeta)-\mathrm{bei}_1(\zeta)\}\end{aligned} \tag{37}$$

$$\begin{aligned}\lambda = {} & 2(1-\nu)\{\mathrm{ber}_1^2(\zeta)+\mathrm{bei}_1^2(\zeta)\} + \sqrt{2}\zeta\{\mathrm{ber}_0(\zeta)(\mathrm{ber}_1(\zeta)-\mathrm{bei}_1(\zeta)) \\ & + \mathrm{bei}_0(\zeta)(\mathrm{ber}_1(\zeta)+\mathrm{bei}_1(\zeta))\}.\end{aligned}$$

Thus $\overline{f}(\overline{r})$, $\overline{s}(\overline{r})$ depend only on $\overline{r}$, z_0/h and ν. In order to compute curvatures and forces per unit length, the following dimensionless functions are introduced:

$$\begin{aligned}\overline{h}_1(\overline{r}) &= -\frac{1}{\overline{r}}\frac{d\overline{f}}{d\overline{r}} = -\frac{\sqrt{2}}{2}\frac{\zeta}{\overline{r}}\{(C_1-C_2)\mathrm{ber}_1(\zeta\overline{r}) \\ &\qquad + (C_1+C_2)\mathrm{bei}_1(\zeta\overline{r})\} - 1 \\ \overline{h}_2(\overline{r}) &= -\frac{d^2\overline{f}}{d\overline{r}^2} = c^2\zeta^2\overline{s}(\overline{r}) - \overline{h}_1(\overline{r}) - 2 \\ \overline{h}_3(\overline{r}) &= \frac{1}{\overline{r}}\frac{d\overline{s}}{d\overline{r}} = -\frac{1}{c^2}\frac{\sqrt{2}}{2}\frac{\zeta}{\overline{r}}\{(C_1+C_2)\mathrm{ber}_1(\zeta\overline{r}) \\ &\qquad - (C_1-C_2)\mathrm{bei}_1(\zeta\overline{r})\} \\ \overline{h}_4(\overline{r}) &= \frac{d^2\overline{s}}{d\overline{r}^2} = \frac{\zeta^2}{c^2}\{C_1\,\mathrm{ber}_0(\zeta\overline{r}) + C_2\,\mathrm{bei}_0(\zeta\overline{r})\} - \overline{h}_3(\overline{r}).\end{aligned} \tag{38}$$

For zero initial curvature, all functions in Eq. (38) vanish and for large initial curvatures $\overline{h}_1(\overline{r})$ and $\overline{h}_2(\overline{r})$ tends to -1 while $\overline{h}_3(\overline{r})$ and $\overline{h}_4(\overline{r})$ again tends to zero for $0\leqslant\overline{r}<1$. The curvatures and normal forces that solves the correction problem can then be expressed as

$$\begin{bmatrix}\kappa_r^c \\ \kappa_\varphi^c \\ 2\kappa_{r\varphi}^c\end{bmatrix} = \frac{1}{2}T\beta_{\alpha\alpha}^1\begin{bmatrix}\overline{h}_2(\overline{r}) \\ \overline{h}_1(\overline{r}) \\ 0\end{bmatrix}, \quad \begin{bmatrix}N_r^c \\ N_\varphi^c \\ N_{r\varphi}^c\end{bmatrix} = \frac{1}{2}Eh^2T\beta_{\alpha\alpha}^1\begin{bmatrix}\overline{h}_3(\overline{r}) \\ \overline{h}_4(\overline{r}) \\ 0\end{bmatrix}. \tag{39}$$

The total solution to the zeroth and first-order problem can now be assembled. The total deflection and stress function can be expressed in Cartesian coordinates as

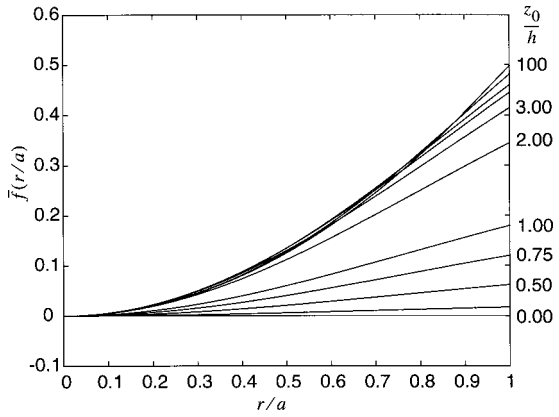


Fig. 2 The dimensionless displacement correction $\bar{f}=2w^c/(a^2T\beta_{\alpha\alpha}^1)$ as a function of dimensionless radius r/a for $z_0/h=0$, 0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 10, 100 and $\nu=0.262$

$$\begin{aligned}w &= -\frac{a^2T}{2}\{\beta_{\alpha\beta}\overline{x}_\alpha\overline{x}_\beta - \beta_{\alpha\alpha}\overline{f}(\overline{r})\} + O(\delta^2) \\ \Phi &= \frac{Eh^2a^2T}{2}\beta_{\alpha\alpha}\overline{s}(\overline{r}) + O(\delta^2)\end{aligned} \tag{40}$$

where $\overline{r}^2=\overline{x}_1^2+\overline{x}_2^2$. In Cartesian coordinates the total curvatures become

$$\begin{bmatrix}\kappa_{11} \\ \kappa_{22} \\ 2\kappa_{12}\end{bmatrix} = \frac{1}{2}T\begin{bmatrix}2\beta_{11}+\beta_{\alpha\alpha}\{\overline{x}_1^2\overline{h}_2(\overline{r})+\overline{x}_2^2\overline{h}_1(\overline{r})\}/\overline{r}^2 \\ 2\beta_{22}+\beta_{\alpha\alpha}\{\overline{x}_2^2\overline{h}_2(\overline{r})+\overline{x}_1^2\overline{h}_1(\overline{r})\}/\overline{r}^2 \\ 4\beta_{12}+2\beta_{\alpha\alpha}\overline{x}_1\overline{x}_2\{\overline{h}_2(\overline{r})-\overline{h}_1(\overline{r})\}/\overline{r}^2\end{bmatrix} + O(\delta^2) \tag{41}$$

whereas the total normal forces per unit length read

$$\begin{bmatrix}N_{11} \\ N_{22} \\ N_{12}\end{bmatrix} = \frac{1}{2}Eh^2T\beta_{\alpha\alpha}\frac{1}{\overline{r}^2}\begin{bmatrix}\overline{x}_1^2\overline{h}_3(\overline{r})+\overline{x}_2^2\overline{h}_4(\overline{r}) \\ \overline{x}_2^2\overline{h}_3(\overline{r})+\overline{x}_1^2\overline{h}_4(\overline{r}) \\ \overline{x}_1\overline{x}_2\{\overline{h}_3(\overline{r})-\overline{h}_4(\overline{r})\}\end{bmatrix} + O(\delta^2). \tag{42}$$

The total moment per unit length can be computed with the isotropic form of Eqs. (10), (12), and (41) whereas the total in-plane strain up to order one can be computed by using Eqs. (10), (12), and (42). For the in-plane strains, however, the term of order zero does not vanish as it does for the other quantities and it can be considered much larger than the terms of order one. The total middle surface strain can therefore approximately be written as $\varepsilon_{\alpha\beta}=T\alpha\delta_{\alpha\beta}+O(\delta)$ where α is the coefficient of thermal expansion of the substrate.

The effect of initial curvature and position on the dimensionless displacement correction is presented in Fig. 2. For large z_0/h, the total solution given by Eq. (40) thus implies that, for an isotropic film, the initial out-of-plane shape of the shell will remain unchanged. Figures 3 and 4 show the effect of initial curvature on the radial ($\overline{h}_2(\overline{r})$) and circumferential ($\overline{h}_1(\overline{r})$) dimensionless flat-plate curvature correction. It is seen that for very small initial curvatures, all corrections vanish as anticipated. Equation (41) and Figs. 3–4 show further that the corrections $\overline{h}_2(\overline{r})$ and $\overline{h}_1(\overline{r})$ are of significance when the initial out-of-plane distance z_0/h is of the order of one half or larger. For large z_0/h a boundary layer with dimensionless length of order $O(\sqrt{h/z_0})$ develops, see the Appendix.

2.4 Extraction of Average Layer Stresses From Curvature Measurements The theory presented in Sections 2.1–2.3 defines implicit relationships between curvature changes ($\kappa_{\alpha\beta}$) and thermal curvature coefficients $\beta_{\alpha\beta}$ which through Eq. (9) are directly connected to average layer stresses $\langle\sigma_{\alpha\beta}\rangle$. Here, explicit

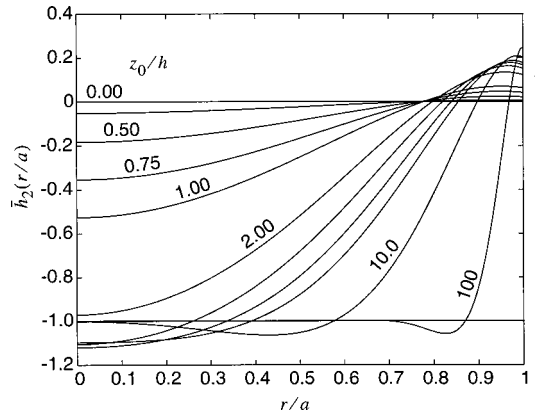


Fig. 3 The dimensionless radial curvature correction $\bar{h}_2 = 2\kappa_r^c/(T\beta_{\alpha\alpha}^1)$ as a function of dimensionless radius r/a for $z_0/h = 0$, 0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 10, 100 and $\nu = 0.262$

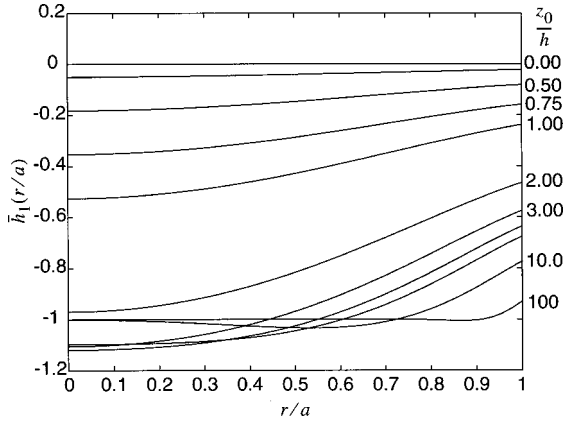


Fig. 4 The dimensionless tangential curvature correction $\bar{h}_1 = 2\kappa_\varphi^c/(T\beta_{\alpha\alpha}^1)$ as a function of dimensionless radius r/a for $z_0/h = 0$, 0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 10, 100 and $\nu = 0.262$

results will be presented for an isotropic substrate with spherical initial curvature (cf. Eq. (41)). It is assumed that the principal directions of $\beta_{\alpha\beta}$ coincide with the 1,2-directions, thus $\beta_{12} = 0$. The curvature measurement technique relies on evenly spaced measurements of the slope of the substrate along some radius before and after a temperature change has been applied. The slopes are then subtracted pointwise (therefore the effect of any initial curvature vanishes) and a straight line is fitted by the least squares method. The slope of the fitted line with respect to the radius is then proportional to the curvature change of the substrate. As long as the substrate is initially flat, the curvature will theoretically be constant over the radius. However, if the substrate is initially curved, the curvature obtained by this method is averaged in a sense and it will not coincide with the flat plate curvature $T\beta_{\alpha\beta}$. It is therefore of interest to find a way to extract accurate volume-averaged stresses from curvature measurements on initially curved substrates. Suppose that the slopes $(\partial w/\partial x_\alpha)$ along the two radial lines x_α are measured on evenly spaced points between $-s \leq x_\alpha \leq s$ where $s \leq a$. The pointwise error along the principal directions can then be expressed as

$$e_\alpha(\bar{x}_\alpha) = \frac{1}{a}\frac{dw(\bar{x}_\alpha)}{d\bar{x}_\alpha} - B_\alpha + a\kappa_\alpha^{LS}\bar{x}_\alpha \quad \text{(no sum over } \alpha\text{)} \tag{43}$$

where κ_α^{LS} represents the unknown averaged principal curvature. The continuous form of the least squares problem is to minimize the following mean square errors with respect to B_α and κ_α^{LS}:

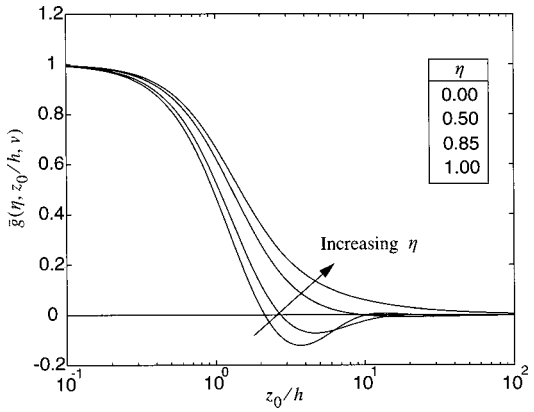


Fig. 5 The function $\bar{g}(\eta, z_0/h, \nu)$ (cf. Eqs. (46), (47)). The parameter z_0/h denotes the initial curvature and $\eta = 0.00$, 0.50, 0.85, 1.00 is the ratio between the scanning length and the substrate radius. An isotropic Si substrate with a Poisson's ratio of 0.262 and a spherical initial curvature is considered.

$$\|e_\alpha(\bar{x}_\alpha)\| = \frac{1}{2\eta}\int_{-\eta}^{\eta} e_\alpha(\bar{x}_\alpha)^2 d\bar{x}_\alpha \quad \text{(no sum over } \alpha\text{)} \tag{44}$$

where $\eta = s/a$. The fitted curvature can then be expressed as

$$\kappa_\alpha^{LS} = -\frac{3}{2}\frac{1}{a^2\eta^3}\int_{-\eta}^{\eta} \bar{x}_\alpha \frac{dw(\bar{x}_\alpha)}{d\bar{x}_\alpha} d\bar{x}_\alpha \quad \text{(no sum over } \alpha\text{)}. \tag{45}$$

Replacing the measured slopes with the theoretically obtained ones and solving the linear system gives

$$\beta_{11}T = \frac{1}{\bar{g}(\eta, z_0/h, \nu)}\frac{(\kappa_1^{LS} + \kappa_2^{LS})}{2} + \frac{(\kappa_1^{LS} - \kappa_2^{LS})}{2} \tag{46}$$

where β_{22} is obtained by interchanging the subscripts 1 and 2. The function $\bar{g}(\eta, z_0/h, \nu)$ becomes

$$\bar{g}(\eta, z_0/h, \nu) = \frac{3}{\eta^3}\int_0^{\eta} \xi^2(1 + \bar{h}_1(\xi))d\xi \tag{47}$$

which can be evaluated numerically. Equation (46) can now be used in conjunction with Eq. (47) to determine the coefficients of thermal curvature and thus to obtain average stresses in the thin layer. Very often, the curvature is only measured in one direction, it is then commonly assumed that the layer is macroscopically isotropic. The second term in Eq. (46) will then vanish and $\bar{g}(\eta, z_0/h, \nu)$ may be interpreted as the ratio between a measured curvature and the corresponding flat plate curvature. It is emphasized that the analysis is valid for elastic as well as inelastic layers. Figure 5 shows the function $\bar{g}(\eta, z_0/h, \nu)$ versus initial out-of-plane deflection z_0/h for different normalized scanning lengths η. It is observed that when $z_0/h \approx 0.5$, the measured curvatures deviate markedly from what is expected for an initially flat substrate. It is also seen that the curvature near the center of the substrate changes sign when $z_0/h \approx 2$. For situations when the initial curvature is large, the isotropic out-of-plane shape of the shell remain unchanged, therefore curvature measurements can in this case not be used to extract volume averaged stresses.

3 Conclusions

The present analysis of initially curved substrates has revealed a very useful first-order approximation. It has been shown that a homogeneous shallow substrate with a thin, generally inelastic anisotropic layer, may be modeled as a bare shallow substrate without layer and the effect of the layer is adequately described by the substrate/layer system's coefficients of thermal curvature. It

has further been shown that for certain combinations of initial curvatures and coefficients of thermal curvature, the solution to the general anisotropic shallow shell problem degenerates into the well-known flat-plate solution. For situations when the initial curvature is large, a boundary layer develops near the free edge of the substrate. In this case, the solution that is valid in the interior of the shell as well as the boundary layer solution is presented in closed form for a general anisotropic substrate of arbitrary in-plane shape.

For the particular case of an isotropic substrate with spherical initial curvature and a general thin layer deposit, a correction to the flat substrate solution that is valid for any constant initial curvature is presented in closed form. This correction exhibits radial symmetry. The analytic correction enables investigations of the effects of various parameters, such as the initial curvature and Poisson's ratio of the substrate. The initial out-of-plane deflection z_0 is observed to change the flat-plate solution significantly if $z_0 \sim h/2$ where h is the thickness of the substrate. The resulting curvatures (which now vary with position) cannot be used directly to interpret curvature measurements on shallow substrates in terms of volume-averaged stresses in the layer. However, this can be achieved by simulating curvature measurements on initially curved substrates. These simulations make it possible to solve the inverse problem and hence to accurately use measured curvatures for the determination of average layer stresses.

Acknowledgments

The authors acknowledge Prof. Subra Suresh (MIT) and Dr. Mårten Olsson (KTH) for valuable discussions during the course of this work.

Appendix

For the case of large initial curvatures, the solution in Section 2.2 has been uniquely divided into a compatible part and a noncompatible part. The general anisotropic asymptotic boundary layer solution (noncompatible part) can now be derived. Consider a coordinate system ξ_i at some point on the boundary of the arbitrarily shaped shell. The ξ_3-direction is defined to be perpendicular to the middle surface and the ξ_1-direction is defined to be perpendicular to the boundary and positive inside the shell. By using $\beta^P_{\alpha\beta}$ according to Eq. (23), the noncompatible parts of the coefficients of thermal curvatures in the general first-order problem are transferred to the boundary conditions. Asymptotically, a stress-function defined according to Eq. (27) as well as the displacements will vary much faster in the ξ_1-direction as compared to the tangential direction (ξ_2), therefore the stress function and displacements will asymptotically only be functions of ξ_1 and the following problem results:

$$h^3 C_{1111} w^{\mathrm{IV}} + 12\tilde{\kappa}_{22}\Phi'' = 0 \qquad S_{2222}\Phi^{\mathrm{IV}} - h\tilde{\kappa}_{22} w'' = 0$$
$$-\frac{h^3}{12} C_{1111} w''(0) = \hat{M}_{11} \equiv T\frac{h^3}{12} C_{11\alpha\beta}(\beta_{\alpha\beta} T - \beta^P_{\alpha\beta}) \qquad (A1)$$
$$w'''(0) = w(\infty) = w'(\infty) = N_{22}(\infty) = 0$$

where all tensorial quantities have been transformed to the local ξ_i frame and $N_{11}(0) = N_{12}(0) = 0$ are automatically fulfilled. The compliance tensor (inverse of $C_{\alpha\beta\gamma\delta}$) is denoted $S_{\alpha\beta\gamma\delta}$. The solution to this boundary layer problem may be written as

$$w = -\frac{6\hat{M}_{11}}{h^3 C_{1111}\lambda^2} e^{-\lambda\xi_1}(\cos(\lambda\xi_1) - \sin(\lambda\xi_1))$$
$$\Phi = \frac{\hat{M}_{11}}{\tilde{\kappa}_{22}} e^{-\lambda\xi_1}(\cos(\lambda\xi_1) - \sin(\lambda\xi_1)) \qquad (A2)$$

where $\lambda^4 = 3\tilde{\kappa}_{22}^2/(h^2 C_{1111} S_{2222})$. It can be shown that for situations when the local curvature $\tilde{\kappa}_{22}$ tend to zero, $\hat{M}_{11}$ also tends to zero.

References

[1] Flinn, P. A., Gardner, S., and Nix, W. D., 1987, ''Measurement and Interpretation of Stress in Aluminum-Based Metallization as a Function of Thermal History,'' IEEE Trans. Electron Devices, **ED-34**, No. 3, pp. 689–699.
[2] Yeo, I.-S., Ho, P. S., and Anderson, S. G. H., 1995, ''Characteristics of Thermal Stresses in Al(Cu) Fine Lines, I. Unpassivated Line Structures,'' J. Appl. Phys., **78**, pp. 945–952.
[3] Flügge, W., 1973, *Stresses in Shells*, 2nd ed., Springer-Verlag, New York.
[4] Leissa, A. W., and Qatu, M. S., 1991, ''Equations of Elastic Deformation of Laminated Composite Shallow Shells,'' J. Appl. Mech., **58**, pp. 181–188.
[5] Wikström, A., Gudmundson, P., and Suresh, S., 1999, ''Thermoelastic Analysis of Periodic Thin Lines Deposited on a Substrate,'' J. Mech. Phys. Solids, **47**, pp. 1113–1130.
[6] Reissner, E., 1946, ''Stresses and Small Displacements of Shallow Spherical Shells, I.,'' J. Math. Phys., **25**, pp. 80–85.
[7] Reissner, E., 1947, ''Stresses and Small Displacements of Shallow Spherical Shells, II.,'' J. Math. Phys., **25**, pp. 279–300.
[8] Reissner, E., 1948, ''Correction to Stresses and Small Displacements of Shallow Spherical Shells, II.'' J. Math. Phys., **27**, p. 240.

A Strain-Based Formulation for the Coupled Viscoelastic/Damage Behavior

K. Abdel-Tawab
Senior Engineer,
GE Power Systems,
Greenville, SC 29602

Y. J. Weitsman
Department of Mechanical and Aerospace Engineering and Engineering Science,
The University of Tennessee,
Knoxville, TN 37966
Fellow ASME

A strain-based thermodynamics framework is proposed for modeling the continuum damage behavior of viscoelastic materials. Damage is represented by an internal state variable in the form of a symmetric second rank tensor. The effect of damage on the constitutive behavior is introduced through direct coupling between the damage variable and the viscoelastic internal state variables. This approach accounts for time-dependent damage as well as damage-induced changes in material symmetry. Also, damage evolution is modeled by employing the concept of damage surfaces. This work is motivated by experimental observations of the response of swirl-mat and random chopped fiber mat polymeric composites where viscoelastic creep was accompanied by a multitude of fiber/matrix interfacial cracks. [DOI: 10.1115/1.1348013]

1 Introduction

The growing interest in the use of polymeric materials (e.g., plastics and polymeric composites) for structural applications mandates appropriate knowledge of the mechanical behavior as well as the durability of these materials. It is well known that polymeric materials creep viscoelastically. In addition, experimental investigations (e.g., [1–4]) indicate that polymeric composites may undergo distributed damage in the form of a multitude of microcracks. The objective of this article is to establish a framework for the constitutive modeling of the foregoing features of material behavior. Such a framework is essential for a reliable engineering design.

Up to the present time, most of the efforts for modeling distributed damage have been directed toward brittle materials exhibiting elastic behavior and metals exhibiting plastic or creep response (e.g., [5]). Less attention has been paid to the modeling of damage in viscoelastic materials. Notably, Schapery [6–8] established a basic formulation for viscoelastic response that is accompanied by microstructural changes, such as profuse microcracking. The microstructural changes are represented in Schapery's work by means of a set of internal state variables whose evolutionary laws are motivated by considerations of viscoelastic fracture mechanics. It may also be mentioned that Weitsman [9] attempted to model the coupling between viscoelasticity and damage for a special class of linear viscoelastic materials.

In a recent article by the present authors ([10]) a stress-based formulation for modeling the coupling between viscoelasticity and distributed damage was developed and applied to a swirl-mat polymeric composite. The effect of damage on material behavior was introduced through the concept of effective stress. Also, damage evolution was related by the empirical Kachanov-Rabotnov forms ([5]), which are best suited for monotonic creep loadings. For more complex loading histories the concept of damage surfaces ([11]) offers a more versatile approach to damage evolution.

It is well known that damage surfaces are better expressed in strain space than in stress space ([5,12]). This seems particularly appropriate for viscoelastic response where creep occurs at all stress levels. More specifically, while under an applied stress both strain and damage evolve with time, under an applied strain the stress relaxes with time and it may then be possible to identify states of stationary damage that are necessary to specify the damage surface. This supposition provided a motivation for the present strain-based formulation of viscoelasticity coupled with damage. An additional advantage is due to the fact that strain-based viscoelastic constitutive models are more convenient for implementation into finite element codes than stress-based ones (e.g., [13]). The present formulation employs concepts of continuum damage mechanics as well as several existing concepts of the thermodynamic theory of viscoelastic materials ([14,15]). This format accounts for time-dependent damage as well as damage-induced changes in material symmetry.

In Section 2 of this article, we present a general thermodynamics framework that accounts for both viscoelastic and damage processes. We proceed by modeling the coupling between these two processes in Section 3. In Section 4, damage evolution is modeled through the concept of damage surfaces and is illustrated by a simple example in Section 5. Section 6 concludes with a summary and some remarks pertinent to the present work.

2 Thermodynamics Framework

Consider a polymeric material and let γ_r $(r=1,2,\ldots,R)$ denote R scalar-valued internal state variables representing the internal degrees-of-freedom of molecular motion in the polymeric chains. The internal state variable representing damage can be related in terms of tensorial quantities of even ranks, which can be associated with the spatial distributions of microcracks ([5]). For simplicity, the damage variable is chosen as a symmetric second rank tensor ω_{ij} with dimensionless components. This damage variable is capable of simulating changes in material symmetry such that an initially damage-free isotropic material may become, at most, orthotropic upon damage formation ([16,17])[1]. Despite the shortcomings of the aforementioned damage variable ([18]), it was adopted by several workers in the field of damage mechanics (e.g., [16,19,20]) due to its relative simplicity and applicability to practical circumstances. It should be mentioned that the present formulation can be readily modified to accommodate damage variables of other tensorial ranks. Throughout this article the subscripts r and q are reserved for scalar quantities; and a, b, c, d, i, j, k, l, m, and n are associated with tensorial quantities and cover the range 1,2,3. Also, the summation convention is implied over the range of repeated indices unless stated otherwise.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 27, 1999; final revision, Mar. 24, 2000. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

[1]Also, an initially orthotropic symmetry may evolve into another orthotropy.

Viscoelasticity and damage are irreversible thermodynamic processes. For a closed system and small strains, the entropy production inequality can be written in the form ([21])

$$-\dot{\psi}+\sigma_{ij}\dot{\epsilon}_{ij}-\mathcal{S}\dot{T}-\frac{h_i T_{,i}}{T}\geqslant 0, \tag{1}$$

where ψ is the Helmholtz free energy (per unit volume), σ_{ij}–components of a suitably defined volume-averaged stress tensor, ϵ_{ij}–components of the infinitesimal strain tensor, $\mathcal{S}$–entropy (per unit volume), T–temperature, h_i–components of the heat flux vector, $T_{,i}=\partial T/\partial x_i$–components of the temperature gradient, and x_i–space coordinates. Also, in (1) an overdot signifies differentiation with respect to time.

Consider a Helmholtz free energy of the form

$$\psi=\psi(\epsilon_{ij},\gamma_r,\omega_{ab},T). \tag{2}$$

The function ψ is assumed to be continuous and sufficiently differentiable with respect to its arguments. Considerations of the entropy production inequality in (1) together with the functional dependence in (2) give the following familiar relations:

$$\sigma_{ij}=\frac{\partial\psi}{\partial\epsilon_{ij}}, \tag{3}$$

$$\mathcal{S}=-\frac{\partial\psi}{\partial T}, \tag{4}$$

and

$$\Gamma_r\dot{\gamma}_r+\Omega_{ab}\dot{\omega}_{ab}-\frac{h_i T_{,i}}{T}\geqslant 0, \tag{5}$$

where Γ_r and Ω_{ab} are the thermodynamic forces conjugate to the internal state variables γ_r and ω_{ab}, respectively, and are given by

$$\Gamma_r=-\frac{\partial\psi}{\partial\gamma_r}, \tag{6}$$

and

$$\Omega_{ab}=-\frac{\partial\psi}{\partial\omega_{ab}}. \tag{7}$$

Finally, from the dissipation inequality (5) we have the following requirements:

$$\Gamma_r\dot{\gamma}_r\geqslant 0, \tag{8}$$

$$\Gamma_r\dot{\gamma}_r+\Omega_{ab}\dot{\omega}_{ab}\geqslant 0. \tag{9}$$

Inequality (8) must be satisfied whenever viscoelastic deformation occurs, while when deformation is accompanied by damage inequality (9) should be satisfied as well.

3 Constitutive Modeling

3.1 General Formulation. In this section a constitutive model is formulated for the case of linear viscoelastic behavior coupled with damage. For simplicity, attention is restricted to the case of isothermal behavior. The extension to the general case of nonisothermal conditions can be made following the same approach adopted here. The underlying hypothesis in the present formulation is that γ_r and ω_{ab} are associated with disparate length scales; molecular for γ_r and, say, fiber/matrix interfacial cracks in a fibrous polymeric composite for ω_{ab}. The formulation will be first established for fixed strain ϵ_{ij} and damage ω_{ab} and subsequently extended to fluctuating ϵ_{ij} and ω_{ab}.

For fixed ϵ_{ij} and ω_{ab} an irreversible thermodynamic process is triggered in the material, which prompts the viscoelastic internal state variables γ_r to drift spontaneously toward their equilibrium values γ_r^e. Under isothermal conditions, all γ_r^e are independent of temperature, hence

$$\gamma_r^e=\gamma_r^e(\epsilon_{ij},\omega_{ab}). \tag{10}$$

These equilibrium values are assumed to be continuous and sufficiently differentiable functions of their arguments. Assuming that all γ_r and γ_r^e are sufficiently small, a Taylor series expansion for ψ about γ_r^e takes the form

$$\psi=\psi_e+\frac{1}{2}\psi_{rq}(\gamma_r-\gamma_r^e)(\gamma_q-\gamma_q^e)+\text{H.O.T.}, \tag{11}$$

where

$$\psi_e=\psi_e(\epsilon_{ij},\omega_{ab})$$

is the value of ψ at equilibrium,

$$\psi_{rq}=\left(\frac{\partial^2\psi}{\partial\gamma_r\partial\gamma_q}\right)_e$$

is a symmetric matrix considered to be constant, and H.O.T. refers to higher order terms neglected due to smallness of γ_r and γ_r^e. In the above relations, and in the sequel, the subscript ''e'' implies that a quantity is calculated at $\gamma_r=\gamma_r^e\ \forall r$. Note that at equilibrium ψ is minimum ([22,23]) and hence

$$\left(\frac{\partial\psi}{\partial\gamma_r}\right)_e=0,$$

and

$$\psi_{rq}\delta\gamma_r\delta\gamma_q>0.$$

Consequently, there is no linear term in (11) and ψ_{rq} is a positive definite matrix. It should be mentioned that an expansion similar to that in (11) was previously used by Lubliner [24].

Employing the usual assumption of viscous-like resistance ([14,15]), let

$$\Gamma_r=a_{rq}\dot{\gamma}_q, \tag{12}$$

where, according to Onsager's principle ([22,25]) a_{rq} is a symmetric matrix. Substitution of (12) into inequality (8) gives

$$a_{rq}\dot{\gamma}_r\dot{\gamma}_q\geqslant 0.$$

Hence, the matrix a_{rq} is positive semi-definite. Note that in the general case a_{rq} is a function of temperature, but since we are considering only isothermal conditions then a_{rq} is constant.

Equations (6), (11), and (12) yield

$$a_{rq}\dot{\gamma}_q+\psi_{rq}\gamma_q=\psi_{rq}\gamma_q^e. \tag{13}$$

Since a_{rq} is a constant symmetric positive semi-definite matrix and ψ_{rq} is a constant symmetric positive definite matrix, it is possible to rewrite (13) in a diagonalized form ([26]) as

$$A_r\dot{\hat{\gamma}}_r+\Psi_r\hat{\gamma}_r=\Psi_r\hat{\gamma}_r^e \quad (\text{no sum over } r), \tag{14}$$

where $\hat{\gamma}_r$ are transformed internal state variables, each being a linear combination of the original internal state variables γ_q. The parameters $\hat{\gamma}_r^e$ are the equilibrium values corresponding to $\hat{\gamma}_r$ and are obtained from γ_q^e by the same linear transformation as that for $\hat{\gamma}_r$. Also, A_r and Ψ_r are constants such that $A_r\geqslant 0$ and $\Psi_r>0$.

For fixed strain and damage, the solution of Eq. (14) is

$$\hat{\gamma}_r=\hat{\gamma}_r^e(1-e^{-t/\tau_r}) \quad (\text{no sum over } r), \tag{15}$$

where τ_r are relaxation times given by

$$\tau_r=\frac{A_r}{\Psi_r} \quad (\text{no sum over } r). \tag{16}$$

In terms of the transformed internal state variables, expansion (11) can be rewritten as

$$\psi=\psi_e+\frac{1}{2}\sum_r \Psi_r(\hat{\gamma}_r-\hat{\gamma}_r^e)^2+\text{H.O.T.} \tag{17}$$

The viscoelastic strain can now be obtained by substituting (17) into (3) bearing in mind that γ_r, and hence $\hat{\gamma}_r$, are to be kept fixed during the partial differentiation indicated in (3). Employing (15) we then obtain

$$\sigma_{ij}=\frac{\partial \psi_e}{\partial \epsilon_{ij}}+\sum_r \frac{\partial \Lambda_r}{\partial \epsilon_{ij}} e^{-t/\tau_r}, \tag{18}$$

where

$$\Lambda_r=\Lambda_r(\epsilon_{ij},\omega_{ab})=\frac{1}{2}\Psi_r(\hat{\gamma}_r^e)^2 \quad \text{(no sum over } r\text{)}. \tag{19}$$

The first term on the right-hand side of (18) represents the long-term (rubbery) part of the behavior, and the second term represents the transient (time-dependent) part.

Motivated by previous works on linear elasticity with damage (e.g., [17,27]), we now recast the formulation in a format that retains a linear viscoelastic relaxation modulus and introduces the effects of damage by mapping the stress and strain into ''damage effective'' stress and strain, respectively. To this end, consider first the transient part in (18). Expanding Λ_r in terms of strain around the reference state, $\epsilon_{ij}^{\text{ref}}=0$, up to quadratic terms to retain linearity one obtains

$$\Lambda_r=\frac{1}{2}\left(\frac{\partial^2 \Lambda_r}{\partial \epsilon_{ij}\partial \epsilon_{kl}}\right)_0 \epsilon_{ij}\epsilon_{kl}, \tag{20}$$

where the subscript 0 implies that a quantity is calculated at the reference state. Note that the constant term in (20) vanishes since $(\gamma_r^e)_0=0$, and hence $\hat{\gamma}_r^e=\Lambda_r=0$ at the reference state. In addition, the linear term in (20) is discarded since it corresponds to a residual stress at the reference state, which is disregarded in the present formulation.

A more specific functional form for Λ_r can be obtained by realizing that the internal molecular motions represented by γ_r occur on a much smaller dimensional scale than that of damage represented by ω_{ab}. This suggests that all γ_r^e, and hence all $\hat{\gamma}_r^e$ and Λ_r, are likely to be affected by damage in a common manner; i.e., they have common dependence on ω_{ab}. Consequently, we can rewrite (20) in the form

$$\Lambda_r=\frac{1}{2}P_{ijab}\Delta C_{abcd}^r P_{cdkl}\epsilon_{ij}\epsilon_{kl} \quad \forall r. \tag{21}$$

In (21), ΔC_{ijkl}^r is a double symmetric fourth rank tensor (i.e., $\Delta C_{ijkl}^r=\Delta C_{jikl}^r=\Delta C_{ijlk}^r=\Delta C_{klij}^r$), and $P_{ijkl}=P_{ijkl}(\omega_{ab})$ is a double symmetric, fourth rank tensor-valued function of the damage variable ω_{ab} such that

$$\text{at } \omega_{ab}=0 \rightarrow P_{ijkl}=I_{ijkl}, \tag{22}$$

where

$$I_{ijkl}=\frac{1}{2}(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}),$$

is the unit fourth-rank tensor and δ_{ij} is Kronecker delta.

Using (21), the transient part of (18) takes the form

$$\sum_r \frac{\partial \Lambda_r}{\partial \epsilon_{ij}} e^{-t/\tau_r}=P_{ijab}\Delta C_{abcd}(t)P_{cdkl}\epsilon_{kl}, \tag{23}$$

where

$$\Delta C_{ijkl}(t)=\sum_r \Delta C_{ijkl}^r e^{-t/\tau_r}. \tag{24}$$

From (18), (22), and (23) it is clear that ΔC_{ijkl} is the undamaged transient (time-dependent) stiffness tensor.

Consider now the long-term part of (18). The equilibrium Helmholtz free energy ψ_e can be expanded around the reference state in the form

$$\psi_e=\frac{1}{2}\left(\frac{\partial^2 \psi_e}{\partial \epsilon_{ij}\partial \epsilon_{kl}}\right)_0 \epsilon_{ij}\epsilon_{kl}. \tag{25}$$

Following common practice in continuum damage mechanics (e.g., [17,27]), we further assume that ψ_e depends on damage in the separable form

$$\psi_e=\frac{1}{2}Q_{ijab}C_{abcd}^e Q_{cdkl}\epsilon_{ij}\epsilon_{kl}, \tag{26}$$

where $Q_{ijkl}=Q_{ijkl}(\omega_{ab})$ is a double symmetric, fourth rank tensor-valued function of ω_{ab} such that $Q_{ijkl}=I_{ijkl}$ at $\omega_{ab}=0$, and consequently C_{ijkl}^e is the undamaged long-term (rubbery) stiffness tensor which is typically positive definite ([25]). Using (26), the long-term part of (18) takes the form

$$\frac{\partial \psi_e}{\partial \epsilon_{ij}}=Q_{ijab}C_{abcd}^e Q_{cdkl}\epsilon_{kl}. \tag{27}$$

Relations (23) and (27) indicate that the long-term part of the behavior can in general depend on damage in a manner that differs from that of the transient part. However, for simplicity, we assume here that both parts have the same dependence on damage so that

$$Q_{ijkl}=P_{ijkl}. \tag{28}$$

Assuming that the inverse P_{ijkl}^{-1} exists, define the following ''damage effective'' stress and strain tensors

$$\tilde{\sigma}_{ij}=P_{klij}^{-1}\sigma_{kl} \tag{29}$$

$$\tilde{\epsilon}_{ij}=P_{ijkl}\epsilon_{kl}, \tag{30}$$

then relation (18) can be rewritten in the compact form

$$\tilde{\sigma}_{ij}=C_{ijkl}(t)\tilde{\epsilon}_{kl}, \tag{31}$$

where

$$C_{ijkl}(t)=C_{ijkl}^e+\Delta C_{ijkl}(t), \tag{32}$$

is the overall (long-term and transient) stiffness tensor of the undamaged material. Relations (29) and (30) are consistent with the formulations of the concepts of effective stress and effective strain (e.g., [16,28]), where the mapping tensor for the effective stress is taken to be the inverse of that of the effective strain. Also, note that both $\tilde{\sigma}_{ij}$ and $\tilde{\epsilon}_{ij}$ are symmetric due to the hypothesized double-symmetry of P_{ijkl}.

Relation (31) suggests that for a given damage level ω_{ab}, instantaneous mapping of the actual stress σ_{kl} and strain ϵ_{kl} according to (29) and (30), respectively, lead to new stress $\tilde{\sigma}_{ij}$ and strain $\tilde{\epsilon}_{ij}$ quantities that are related by the usual linear viscoelastic constitutive relation for fixed strain (e.g., [25]). Upon hypothesizing time-translation invariance, and since $\tilde{\sigma}_{ij}$ is linear in $\tilde{\epsilon}_{kl}$, a straightforward application of the superposition principle ([29]) to expression (31) yields

$$\tilde{\sigma}_{ij}=\int_{0^-}^t C_{ijkl}(t-\tau)\frac{d\tilde{\epsilon}_{kl}}{d\tau}d\tau. \tag{33}$$

Allowing for spatial variations of stress and damage the total derivative $d/d\tau$ inside the integral is replaced by a partial derivative $\partial/\partial\tau$, holding the spatial coordinates x_i fixed. Thus

$$\tilde{\sigma}_{ij}=\int_{0^-}^t C_{ijkl}(t-\tau)\frac{\partial\tilde{\epsilon}_{kl}}{\partial\tau}d\tau. \tag{34}$$

Equation (34) is the stress-strain constitutive relation for the coupled linear viscoelastic/damage behavior, and can be expressed in terms of the actual stress and strain as

$$\sigma_{ij}=P_{ijab}\int_{0^-}^t C_{abcd}(t-\tau)\frac{\partial(P_{cdkl}\epsilon_{kl})}{\partial\tau}d\tau. \tag{35}$$

It should be noted that at $t=0$

$$C_{ijkl}(0)=C^o_{ijkl}=C^e_{ijkl}+\sum_r \Delta C^r_{ijkl}, \tag{36}$$

where C^o_{ijkl} is the initial (elastic) stiffness tensor which is positive definite ([25]).

3.2 The Isotropic Case. Consider now the special case of isotropic virgin material response, which applies to the class of randomly reinforced materials of current interest. The overall stiffness tensor C_{ijkl} takes the form ([25])

$$C_{ijkl}(t)=2G(t)I_{ijkl}+\left[K(t)-\frac{2}{3}G(t)\right]\delta_{ij}\delta_{kl}, \tag{37}$$

where $G(t)$ is the overall shear modulus and $K(t)$ is the overall bulk modulus given, respectively, by

$$G(t)=G_o+\Delta G(t), \tag{38}$$

and

$$K(t)=K_o+\Delta K(t). \tag{39}$$

In the above expressions G_o and K_o are the instantaneous shear and bulk moduli, respectively; and $\Delta G(t)$ and $\Delta K(t)$ are the transient shear and bulk moduli, respectively, obtained from (24) as

$$\Delta G(t)=\sum_r \Delta G_r e^{-t/\tau_r}, \tag{40}$$

and

$$\Delta K(t)=\sum_r \Delta K_r e^{-t/\tau_r}, \tag{41}$$

where ΔG_r and ΔK_r are positive constants.

3.3 The Dissipation Inequality. The thermodynamic force Ω_{ab} conjugate to ω_{ab} can be obtained by substituting (17) into (7) using (21), (26), and (28)

$$\Omega_{ab}=-P_{ijcd}\hat{C}_{cdmn}\frac{\partial P_{mnkl}}{\partial \omega_{ab}}\epsilon_{ij}\epsilon_{kl}, \tag{42}$$

where $\hat{C}_{ijkl}$ is given by

$$\hat{C}_{ijkl}=C^o_{ijkl}-\sum_r \left(\frac{\hat{\gamma}_r}{\hat{\gamma}^e_r}\right)\Delta C^r_{ijkl}. \tag{43}$$

Since

$$0\leqslant\left(\frac{\hat{\gamma}_r}{\hat{\gamma}^e_r}\right)\leqslant 1; \quad \forall r,$$

then $\hat{C}_{ijkl}$ is bounded by C^e_{ijkl} and C^o_{ijkl} corresponding, respectively, to the upper and lower limits of $\hat{\gamma}_r/\hat{\gamma}^e_r$. Since C^e_{ijkl} and C^o_{ijkl} are positive definite then it follows that $\hat{C}_{ijkl}$ is also positive definite.

Employing (6) and (7), the dissipation inequality (9) can be expressed as

$$\sum_r A_r\dot{\hat{\gamma}}^2_r-P_{ijab}\hat{C}_{abcd}\dot{P}_{cdkl}\epsilon_{ij}\epsilon_{kl}\geqslant 0, \tag{44}$$

where

$$\dot{P}_{cdkl}=\frac{\partial P_{cdkl}}{\partial \omega_{mn}}\dot{\omega}_{mn}. \tag{45}$$

Noting that the first term on the left-hand side of (44) is always non-negative, then a sufficient but not necessary condition to satisfy (44) is

$$-P_{ijab}\hat{C}_{abcd}\dot{P}_{cdkl}\rightarrow \text{positive semi-definite.} \tag{46}$$

3.4 The Mapping Tensor. The functional form of the mapping tensor P_{ijkl} is restricted by the requirement that P_{ijkl} is double symmetric in addition to the requirements in (22) and (44). In general, P_{ijkl} is an anisotropic fourth-rank tensor function of ω_{ab}. However, due to the complexity of anisotropic functional forms ([30]) and the fact that damage-induced anisotropy (or, more precisely, orthotropy) can be deduced from a symmetric second-rank damage tensor, then the more complex anisotropic functional form may be avoided.

Following Murakami and Imaizumi [19], a simpler representation of P_{ijkl} can be obtained by taking it as an isotropic fourth-rank tensor function of ω_{ab}. A further simplification of the representation of P_{ijkl} is obtained by considering a case of dilute concentration of microcracks in which P_{ijkl} is linear in ω_{ab}. In this case, P_{ijkl} can be written as ([19])

$$P_{ijkl}=c_1\delta_{ij}\delta_{kl}+c_2(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})+c_3\delta_{ij}\omega_{kl}+c_4\delta_{kl}\omega_{ij}+c_5(\delta_{ik}\omega_{jl}+\delta_{il}\omega_{jk}+\delta_{jk}\omega_{il}+\delta_{jl}\omega_{ik})+\text{H.O.T.}, \tag{47}$$

where c_α ($\alpha=1,2,\ldots,5$) are constants.

From (47), it is clear that double symmetry of P_{ijkl} dictates that $c_3=c_4$. Also, the requirement in (22) renders $c_1=0$ and $c_2=1/2$. To determine $c_3=c_4$ and c_5, we consider the special case of isotropic damage in which the ensuing damage pattern does not affect the symmetry of the virgin material. In this case damage is represented by a single scalar ω so that

$$\omega_{ab}=\omega\delta_{ab}; \quad 0\leqslant\omega<1, \tag{48}$$

and the mapping tensor P_{ijkl} takes the form ([28])

$$P_{ijkl}=(1-\omega)I_{ijkl}, \tag{49}$$

which is the inverse of the corresponding mapping tensor that maps the applied stress into the Kachanov effective stress in the case of scalar damage (e.g., [11]). The functional form in (49) can be recovered from (47) by setting $c_3=c_4=0$ and taking $c_5=-1/4$. Thus, the simplest possible form of P_{ijkl} becomes

$$P_{ijkl}=I_{ijkl}-\frac{1}{4}(\delta_{ik}\omega_{jl}+\delta_{il}\omega_{jk}+\delta_{jk}\omega_{il}+\delta_{jl}\omega_{ik}). \tag{50}$$

It is interesting to note that P^{-1}_{ijkl} obtained from (50) indeed coincides with one of the forms proposed by Chen and Chow [20] for the tensor that maps the applied stress into an effective stress.

The complete formulation of the constitutive model requires an expression for the evolution of the damage tensor $\dot{\omega}_{ab}$ such that (44) is satisfied. Such an expression can be formally derived from thermodynamic considerations (e.g., [17,27]), but the usefulness of such approach seems to be restricted to elastic or elastoplastic response with damage ([31]). In practice, the form of the damage evolution equation depends on the material considered and the applied loading. This dependence is better correlated within the concept of damage surfaces ([5]) as discussed in the following section.

4 Damage Evolution

The approach adopted here for describing damage evolution follows closely that presented by Simo and Ju [28] and Lubarda and Krajcinovic [32]. This approach has two main ingredients. First, a damage surface is introduced in strain space to distinguish between the material states associated with evolving damage and those with stationary damage. Second, a damage potential is assumed to exist, from which the constitutive law of damage growth (i.e., the damage rate $\dot{\omega}_{ab}$) can be derived.

To characterize damage evolution, i.e., *damage loading* conditions, a damage function $f(\epsilon_{ij},\kappa)$ is introduced so that

$$f(\epsilon_{ij},\kappa)\leqslant 0, \tag{51}$$

where κ is a positive scalar damage threshold history parameter and at the initial onset of damage $\kappa=\kappa_o$. The equality in (51), i.e., $f=0$, corresponds to strain states that lie on the damage sur-

face and for which damage can evolve. For simplicity, the function f is chosen in the simple isotropic hardening form

$$f(\epsilon_{ij},\kappa)=\mathcal{F}(\epsilon_{ij})-\kappa, \tag{52}$$

where $\mathcal{F}$ is a scalar function of the strain.

Introduce a monotonic scalar function $\mathcal{G}(\Omega_{ab})$ such that the damage rate can be expressed as

$$\dot{\omega}_{ab}=\dot{\lambda}\frac{\partial \mathcal{G}}{\partial \Omega_{ab}} \tag{53}$$

where λ is a monotonically increasing positive scalar, i.e.,

$$\dot{\lambda}\geqslant 0. \tag{54}$$

Physically, λ represents a measure of the cumulative damage at the considered instant of the deformation process. The function $\mathcal{G}$ is referred to as the ''damage potential.''

Following Simo and Ju [28] let

$$\dot{\lambda}=\dot{\kappa}, \tag{55}$$

and define *damage loading/unloading* conditions according to relations (51) and (54) together with

$$\dot{\lambda}f=0. \tag{56}$$

Thus, if $f<0$ then $\dot{\lambda}=0$ and from (53) no damage evolution takes place, i.e., the so-called *damage unloading* from the current state of strain on the damage surface takes place. If $f=0$ and $\dot{\lambda}=0$ then *damage neutral loading* occurs. Finally, if $\dot{\lambda}\neq 0$ then $f=0$ and *damage loading* takes place.

During *damage loading*, the consistency condition

$$\dot{f}=0, \tag{57}$$

must always be satisfied. From (52) and (57) we have

$$\dot{\kappa}=\frac{\partial \mathcal{F}}{\partial \epsilon_{ij}}\dot{\epsilon}_{ij}. \tag{58}$$

Assuming that no damage healing occurs, i.e., the damage surface can only expand, then κ is obtained from (52) and (55) as

$$\kappa=\max\{\kappa_o,\mathcal{F}_{\max}\}, \tag{59}$$

where $\mathcal{F}_{\max}$ is the maximum value of $\mathcal{F}$ over the entire loading history. Substitution of (58) and (55) into (53) yields

$$\dot{\omega}_{ab}=\frac{\partial \mathcal{G}}{\partial \Omega_{ab}}\frac{\partial \mathcal{F}}{\partial \epsilon_{ij}}\dot{\epsilon}_{ij}. \tag{60}$$

Thus, specification of the functional forms of $\mathcal{F}$ and $\mathcal{G}$ completes the formulation for the damage evolution. In practice, these functional forms depend on the material considered and the ensuing damage pattern. Example functional forms will be presented in the following section.

In the damage evolution Eq. (60) the thermodynamic force Ω_{ab} is given by expression (42) in which $\hat{C}_{ijkl}$ is given by (43). Thus an explicit expression for Ω_{ab} requires evaluation of the ratio $\hat{\gamma}_r/\hat{\gamma}_r^e$. This ratio can be determined from the differential equation

$$\frac{d}{dt}\left(\frac{\hat{\gamma}_r}{\hat{\gamma}_r^e}\right)+\left(\frac{1}{\tau_r}+\frac{1}{2\Lambda_r}\frac{d\Lambda_r}{dt}\right)\left(\frac{\hat{\gamma}_r}{\hat{\gamma}_r^e}\right)=\frac{1}{\tau_r}\quad (\text{no sum over } r), \tag{61}$$

where Λ_r is given by (21). Equation (61) is obtained after simple algebraic manipulations of Eq. (14) and making use of (16) and (19).

It should be noted that Ω_{cd} depends on ω_{ab} explicitly through P_{mnkl} and also implicitly through the ratio $\hat{\gamma}_r/\hat{\gamma}_r^e$. Thus, in practice, Eqs. (34), (60), and (61) need be implemented incrementally where for given strain and time increments an iterative procedure is required for determining the corresponding damage increment.

5 Illustrative Example for the Damage Functions: Isotropic Symmetry

Two scalar functions $\mathcal{F}(\epsilon_{ij})$ and $\mathcal{G}(\Omega_{ab})$ are needed in (60) to obtain an explicit damage evolution relation. The simplest possible representation of these functions is to take each as an isotropic function of its tensor argument. Thus

$$\mathcal{F}=\mathcal{F}(I_1^{\epsilon},I_2^{\epsilon},I_3^{\epsilon}), \tag{62}$$

and

$$\mathcal{G}=\mathcal{G}(I_1^{\Omega},I_2^{\Omega},I_3^{\Omega}), \tag{63}$$

where I_1^{β}, I_2^{β}, and I_3^{β} are the isotropic invariants of the corresponding second-rank symmetric tensor β_{ij}. These isotropic invariants can be written as ([30])

$$I_1^{\beta}=\beta_{kk},$$

$$I_2^{\beta}=\sqrt{\beta'_{ij}\beta'_{ij}},$$

and

$$I_3^{\beta}=\det[\beta_{ij}],$$

where β'_{ij} is the deviatoric part of β_{ij}

$$\beta'_{ij}=\beta_{ij}-\frac{1}{3}\beta_{kk}\delta_{ij}.$$

To simplify matters, discard the dependence of $\mathcal{F}$ on I_3^{ϵ}—this is a customary constitutive assumption in damage modeling ([5]). Further, $\mathcal{F}$ is expressed in the following simple form ([32]):

$$\mathcal{F}=F_1\epsilon'_{ij}\epsilon'_{ij}+F_2\epsilon_{kk}^2, \tag{64}$$

where F_1 and F_2 are constants.

Considering the case of isotropic damage, P_{ijkl} can be taken in the form (49) and (60) should reduce to the simple form

$$\dot{\omega}_{ab}=\dot{\omega}\delta_{ab}. \tag{65}$$

It follows that $\mathcal{G}$ can depend only on I_1^{Ω}, i.e., on the trace of Ω_{ab}, so that

$$\frac{\partial \mathcal{G}}{\partial \Omega_{ab}}\sim\delta_{ab}.$$

Consider the case of a dilute concentration of microcracks where the interaction between microcracks as well as the effect of accumulated microcracks on further microcrack formation can be neglected. In this case, the rate of damage evolution $\dot{\omega}$ may be taken to be independent of the accumulated damage ω. Dependence of $\dot{\omega}$ on ω, however, is implicit in Ω_{cc}. Thus, to eliminate dependence of $\dot{\omega}$ on ω the function $\mathcal{G}$ is taken to be linear in Ω_{cc}

$$\mathcal{G}=G\Omega_{cc}, \tag{66}$$

where G is a constant. Expression (66) can be thought of as the first term in a Taylor series expansion of $\mathcal{G}$ (the constant term in such expansion is immaterial to the present formulation). Thus addition of higher order terms in the expansion introduces dependence of $\dot{\omega}$ on ω.

Under *damage loading* conditions $\dot{\omega}>0$ and the left-hand side of the dissipation inequality (44) becomes

$$\sum_r A_r\hat{\gamma}_r^2+(1-\omega)\dot{\omega}\hat{C}_{ijkl}\epsilon_{ij}\epsilon_{kl},$$

which is always positive since $\hat{C}_{ijkl}$ is positive definite. Thus the requirement of positive dissipation is identically satisfied. Substitution of (64) and (66) into (60) and use of (65) yield damage evolution in the form

$$\dot{\omega}=\alpha\epsilon'_{ij}\dot{\epsilon}'_{ij}+\beta\epsilon_{kk}\dot{\epsilon}_{kk}, \tag{67}$$

where

$$\alpha = 2F_1G \quad \text{and} \quad \beta = 2F_2G$$

are free parameters that need to be determined from the damage evolution pattern in a considered problem. The first term on the right-hand side of (67) represents the effect of the deviatoric part of the behavior on damage evolution, whereas the second term represents the effect of the hydrostatic part.

6 Experimental Evidence and Its Interpretation

As noted earlier, the formulation presented in this paper is motivated by the experimentally recorded response of a class randomly reinforced composites. Though all data were collected under stress control, and was further limited to uniaxial loading, it was nevertheless possible to discern behavior that confirms several of the basic premises of the current strain-based model.

For instance, it was noticed ([4]) that viscoelastic creep can proceed at fixed levels of damage for an extended time duration. Namely, the imposition of a step stress, of magnitude that exceeds some threshold level, reduces the instantaneous stiffness of the aforementioned composites, but that stiffness does not degrade any further for substantially long time spans despite the ongoing creep process. This decoupling between damage growth and creep

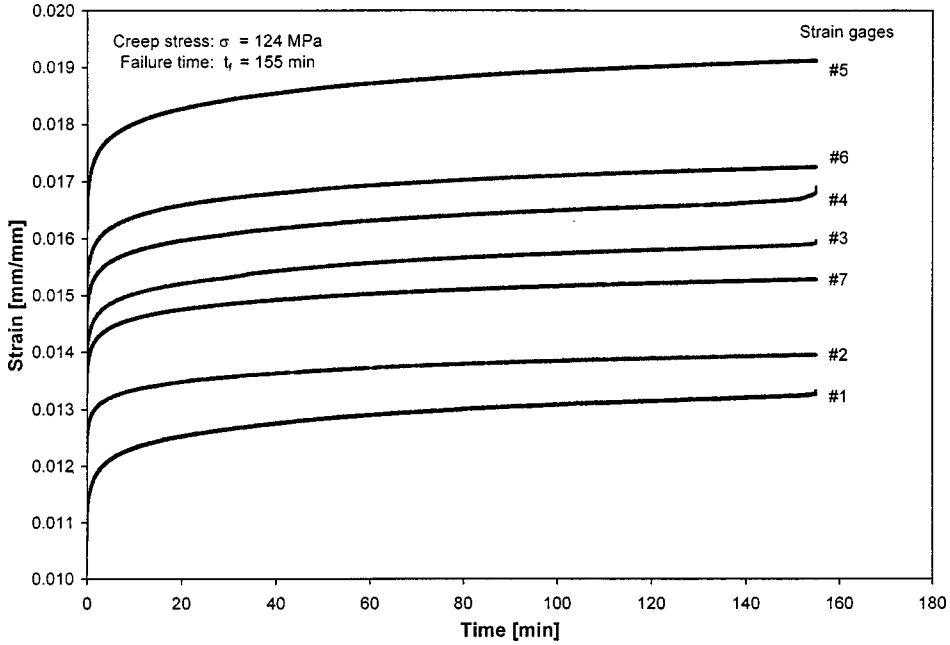


Fig. 1 Strain versus time for a multigaged random chopped-glass mat/urethane composite coupon. Failure occurred at time $t_f = 155$ min.

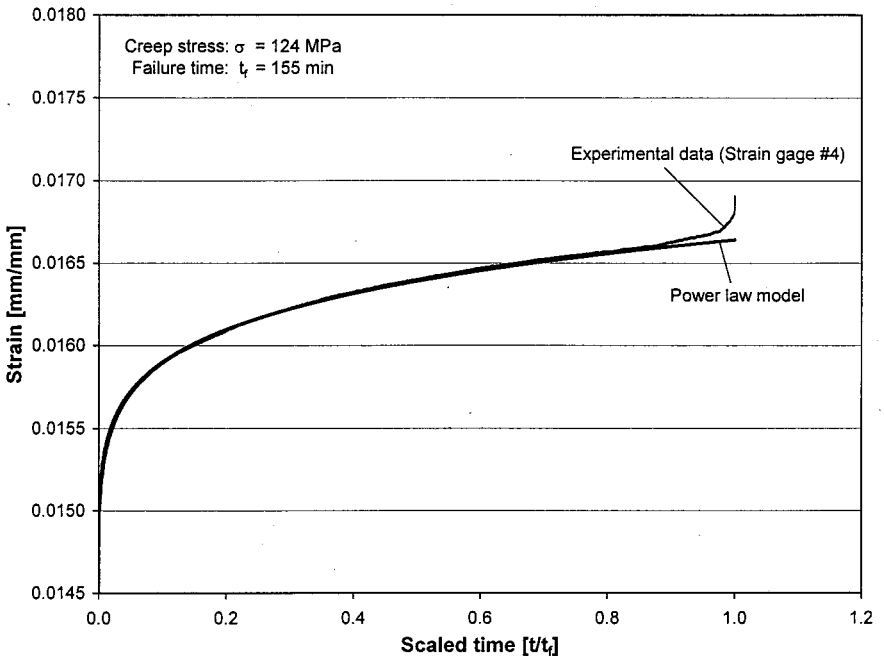


Fig. 2 Expanded exposition of the strain recorded by strain gage #4, detailing the experimental value and the prediction of power-law creep form

was observed to persist for relatively long times even at elevated temperatures ([33]). However, the contribution of damage growth to total deformation under sustained loads became evident for durations that exceeded 75 percent to 95 percent of failure time. In most case ultimate failure (''static fatigue'') was centered about the softest location within test samples ([34]).

Evidence for the above-mentioned creep, damage, and failure processes is provided by the typical data shown in Figs. 1 and 2[2]. Figure 1 exhibits typical creep data, recorded in a multigaged coupon made of randomly reinforced chopped glass mat/urethane matrix materials. The coupon was subjected to a step stress $\sigma=25$ MPa ($=17$ percent of σ_{ult}) and, upon unloading and full recovery, reloaded to $\sigma=124$ MPa ($=85$ percent of σ_{ult}). The lower stress level, which was well within the linear range, related the magnitudes of the undamaged initial stiffnesses. The high stress level was imposed intentionally in order to attain failure within a reasonably short time. The imposition of the higher stress resulted in reduced initial stiffnesses, which were attributed to the presence of instantaneous damage. As may be noted from Fig. 1, all the individual creep curves remain nearly parallel to each other up to time $t_r=142$ min. The creep response for $0<t<t_r$ can be fitted very accurately to a power-law form. This form is applicable for all stress levels, including the linear range ([4,33,34]).

Consider now Fig. 2, where attention is focused on the data that correspond to strain gage #4, which is the location of failure at time $t=t_f=155$ min. This figure exhibits an increasing discrepancy between the recorded values of total strain and the extrapolated prediction of power-law creep. This discrepancy was attributed to damage growth. In fact, this interpretation of the time-dependence of damage growth resulted in good predictions for time-to-failure in ''static fatigue'' at various stress levels[3] ([34]).

The foregoing observations suggest that damage can be quantified by the relative reduction in stiffness ($=$increase in compliance), while damage growth can be expressed by the difference between the total time-dependent strain and the amount associated with power-law creep. The latter portion was interpreted as representing the creep due to polymeric molecular motion. While that molecular motion, which occurs on a dimensional scale much smaller that that of micro-damage, seems to proceed independently of damage growth, the evolution of damage cannot proceed in the absence of molecular motion. It was this interpretation of the physical processes at hand, which was supported visually through microscopic observations ([4]), that led us to the selection of the inequalities (8) and (9)[4].

7 Concluding Remarks

In this article a thermodynamically consistent framework was proposed to model the coupling between linear viscoelastic deformation and microcrack damage. The effect of damage was incorporated into the constitutive equations in a form consistent with the well-known effective stress and effective strain concepts, and damage evolution was related by the concept of damage surfaces.

Several tensorial ranks for the damage variable can be employed in the context of the present formulation. However, for simplicity, in this article damage is represented by a symmetric second-rank tensor. This representation is capable of simulating some changes in material symmetry induced by microcrack damage. A more general representation is to take the damage variable in the form of a double symmetric fourth-rank tensor ω_{abcd}. Such a representation is capable of simulating more general damage-induced changes in material symmetry ([5]). This, however, complicates construction of the functional forms for the mapping tensor P_{ijkl} and the damage potential $\mathcal{G}$; since both have to be functions of double symmetric fourth-rank tensors (ω_{abcd} in the case of P_{ijkl} and the conjugate thermodynamic force Ω_{abcd} in the case of $\mathcal{G}$). For fourth-rank tensors, the definite forms of the integrity bases and invariants are not yet well established ([30]).

An important remark is that in the present formulation the thermodynamic force conjugate to damage depends on the viscoelastic internal state variables as can be seen from relations (42) and (43). This differs from previous formulations by Schapery [6,8,35], where the thermodynamic force conjugate to damage is taken to be independent of the viscoelastic internal state variables and to depend only on the elastic (instantaneous) part of the deformation. The argument put forth by Schapery is that based on a viscoelastic fracture mechanics analysis ([36]), it was found that the driving force for *existing* cracks is independent of the viscoelastic internal state variables. However, damage evolution occurs not only by the extension of existing microcracks, but also by the nucleation of new microcracks. In this general case, it is expected that the state of the viscoelastic deformation in the material should have a direct effect on the formation of new microcracks. Hence, the thermodynamic force conjugate to damage is expected to depend on the viscoelastic internal state variables. This is also in agreement with the internal state variable formulation for the coupled elastoplastic-damage behavior (e.g., [31]).

Acknowledgments

This work was performed, in part, under contract N00014-96-1-0821 from the Mechanics Division of the Engineering Science Directorate of the Office of Naval Research (Dr. Y. Rajapakse—Program Manager); and in part with support from the Office of Transportation Materials, U.S. Department of Energy, under contract DE-AC05-840R21400 with Lockheed-Martin Energy Systems, Inc., at Oak Ridge National Laboratory.

References

[1] Suvorova, J. V., 1985, ''The Influence of Time and Temperature on the Reinforced Plastic Strength,'' *Handbook of Composites*, Vol. 3—Failure Mechanics of Composites, G. C. Sih and A. M. Skudra, eds., Elsevier, New York, pp. 177–213.

[2] Schapery, R. A., 1989 ''Mechanical Characterization and Analysis of Inelastic Composite Laminates With Growing Damage,'' *Mechanics of Composite Materials and Structures*, J. N. Reddy and J. L. Tapley, eds., AMD-Vol. 100, ASME, New York, pp. 1–9.

[3] Tuttle, M. E., Pasricha, A., and Emery, A. F., 1995, ''The Nonlinear Viscoelastic-Viscoplastic Behavior of IM7/5260 Composites Subjected to Cyclic Loading,'' J. Compos. Mater., **29**, No. 15, pp. 2025–2046.

[4] Smith, L. V., and Weitsman, Y., 1999, ''The Visco-Damage Mechanical Response of Swirl-Mat Composites,'' Int. J. Fract., **97**, pp. 301–319.

[5] Krajcinovic, D., 1996, *Damage Mechanics*, Elsevier, New York.

[6] Schapery, R. A., 1981, ''On Viscoelastic Deformation and Failure Behavior of Composite Materials With Distributed Flaws,'' *Advances in Aerospace Structures and Materials*, S. S. Wang and W. J. Renton, eds., ASME, New York, pp. 5–20.

[7] Schapery, R. A., 1994, ''Nonlinear Viscoelastic Constitutive Equations for Composites Based on Work Potentials,'' Appl. Mech. Rev., **47**, No. 6, Part 2, pp. S269–S275.

[8] Schapery, R. A., 1996, ''Characterization of Nonlinear, Time-Dependent Polymers and Polymeric Composites for Durability Analysis,'' *Progress in Durability Analysis of Composite Systems*, A. H. Cardon, H. Fukuda, and K. Reifsnider, eds., Balkema, Rotterdam, pp. 21–38.

[9] Weitsman, Y., 1988, ''A Continuum Damage Model for Viscoelastic Materials,'' ASME J. Appl. Mech., **55**, pp. 773–780.

[10] Abdel-Tawab, K., and Weitsman, Y., 1998, ''A Coupled Viscoelasticity/Damage Model with Application to Swirl-Mat Composites,'' Int. J. Damage Mech., **7**, No. 4, pp. 351–380.

[11] Kachanov, L. M., 1986, *Introduction to Continuum Damage Mechanics*, Martinus Nijhoff, Dordrecht, The Netherlands.

[12] Ju, J. W., 1989, ''On Energy-Based Coupled Elastoplastic Damage Theories: Constitutive Modeling and Computational Aspects,'' Int. J. Solids Struct., **25**, No. 7, pp. 803–833.

[13] ABAQUS, 1998, *Theory Manual—Version 5.7*, HKS, Inc., Pawtucket, RI.

[14] Biot, M. A., 1954, ''Theory of Stress-Strain Relations in Anisotropic Viscoelasticity and Relaxation Phenomena,'' J. Appl. Phys., **25**, No. 6, pp. 1385–1391.

[15] Schapery, R. A., 1964, ''Application of Thermodynamics to Thermomechanical, Fracture and Birefringent Phenomena in Viscoelastic Media,'' J. Appl. Phys., **35**, No. 5, pp. 1451–1465.

[16] Cordebois, J. P., and Sidoroff, F., 1982, ''Damage Induced Elastic Anisotropy,'' *Mechanical Behavior of Anisotropic Solids, Proc. EUROMECH Col-*

[2]Experiments performed by Dr. S. Deng.

[3]Data collected independently by Dr. H. McCoy, Jr.

[4]While Eqs. (8) and (9) are judged to be necessary, the replacement of (9) by Eq. (46) is merely a sufficient condition, which is considered for mathematical expediency.

loque 115, June 1979, J. P. Boehler, ed., Martinus Nijhoff, The Netherlands, pp. 761–774.
[17] Lemaitre, J., 1992, *A Course on Damage Mechanics*, Springer-Verlag, New York.
[18] Lubarda, V. A., and Krajcinovic, D., 1993, ''Damage Tensors and the Crack Density Distributions,'' Int. J. Solids Struct., **30**, pp. 2859–2877.
[19] Murakami, S., and Imaizumi, T., 1982, ''Mechanical Description of Creep Damage State and Its Experimental Verification,'' J. Mec. Theor. Appl., **1**, No. 5, pp. 743–761.
[20] Chen, X. F., and Chow, C. L., 1995, ''On Damage Strain Energy Release Rate Y,'' Int. J. Damage Mech., **4**, pp. 251–263.
[21] Coleman, B. D., and Gurtin, M., 1967, ''Thermodynamics With Internal Variables,'' J. Chem. Phys., **47**, pp. 597–613.
[22] Callen, H. B., 1960, *Thermodynamics*, John Wiley and Sons, New York.
[23] Prigogine, I., 1967, *Thermodynamics of Irreversible Processes*, 3rd Ed., John Wiley and Sons, New York.
[24] Lubliner, J., 1972, ''On the Thermodynamic Foundation of Non-Linear Solid Mechanics,'' Int. J. Non-Linear Solid Mech., **7**, pp. 237–254.
[25] Fung, Y. C., 1965, *Foundations of Solid Mechanics*, Prentice-Hall, Englewood Cliffs, NJ.
[26] Meirovitch, L., 1967, *Analytical Methods in Vibrations*, Macmillan, New York.
[27] Lemaitre, J., and Chaboche, J.-L., 1985, *Mechanics of Solid Materials*, Cambridge University Press, New York (English translation).
[28] Simo, J. C., and Ju, J. W., 1987, ''Strain- and Stress-Based Continuum Damage Models—I. Formulation,'' Int. J. Solids Struct., **23**, pp. 821–840.
[29] Pipkin, A. C., 1986, *Lectures on Viscoelasticity Theory*, 2nd Ed. (Applied Mathematical Sciences, Vol. 7), Springer-Verlag, New York.
[30] Zheng, Q-S., 1994, ''Theory of Representation of Tensor Functions—A Unified Invariant Approach to Constitutive Equations,'' Appl. Mech. Rev., **47**, No. 11, pp. 545–586.
[31] Hansen, N. R., and Schreyer, H. L., 1994, ''A Thermodynamically Consistent Framework for Theories of Elastoplasticity Coupled With Damage,'' Int. J. Solids Struct., **31**, No. 3, pp. 359–389.
[32] Lubarda, V. A., and Krajcinovic, D., 1995, ''Constitutive Structure of Rate Theory of Damage in Brittle Elastic Solids,'' Appl. Math. Comput., **67**, pp. 81–101.
[33] Elahi, M., and Weitsman, Y. J., 1999, ''The Mechanical Response of Random Swirl-Mat Polymeric Composite,'' J. Eng. Mater. Technol., **121**, pp. 460–467.
[34] Leon, R., and Weitsman, Y. J., 2001, ''Time-to-Failure of Randomly Reinforced Glass Strand/Urethane Matrix Composites: Data, Statistical Analysis and Theoretical Prediction,'' Mech. Mater., **33**, No. 3, pp. 127–137.
[35] Schapery, R. A., 1996, ''Nonlinear Constitutive Relations for Polymers and Polymeric Composites With Distributed Damage,'' Univ. of Texas Report No. SSM-96-5, Sept.
[36] Schapery, R. A., 1984, ''Correspondence Principles and a Generalized *J*-Integral for Large Deformation and Fracture Analysis of Viscoelastic Media,'' Int. J. Fract., **25**, pp. 195–223.

Flow in Porous Media of Variable Permeability and Novel Effects

D. A. Siginer[1]
Dean, Fellow ASME,
College of Engineering,
Wichita State University,
105 Wallace Hall,
1845 N. Fairmount,
Wichita, KS 67260-0044
e-mail: siginer@engr.twsu.edu

S. I. Bakhtiyarov
Research Professor,
Mechanical Engineering Department,
Auburn University,
202 Ross Hall,
Auburn, AL 36849
Mem. ASME

The flow of polymeric liquids in a porous medium of variable permeability represented by a cylindrical tube randomly packed with glass spheres is studied. The cylinder represents two porous media of different permeabilities and same porosity arranged in series. We show that the energy loss is higher if the polymeric solution flows first through the porous medium with the smaller permeability rather than through the section of the cylinder with the larger permeability. The difference in energy requirements increases with increasing Reynolds number and may be as high as 25–35 percent for Reynolds numbers of O(1). This is a novel effect not observed for Newtonian and highly shear thinning inelastic fluids flowing through the same configuration. Energy requirements for the same volume flow rate are much higher than a Newtonian fluid of the same zero shear viscosity as the polymeric solution. Energy loss increases with increasing Reynolds number at a fixed concentration to level off at a Reynolds number of O(1). At a fixed Reynolds number, the loss is a strong function of the concentration and shows large increases with increasing concentration. For shear-thinning oil field spacer fluids $De\sim 0.1$ represents a good criterion for the onset of elasticity effects. For solutions of polyacrylamide $De\sim 0.1$ corresponds approximately to the flow rate at which pressure drop starts becoming dependent on the flow direction. Expressions for the friction factor and the resistance coefficient as a function of the Reynolds number have been developed using the inelastic KPK (Kutateladze-Popov-Kapakhpasheva) and viscoelastic eight constant Oldroyd models, respectively. The behavior of inelastic shear-thinning and viscoelastic fluids as represented by oil field spacer fluids and aqueous solutions of polyacrylamide is predicted qualitatively except the difference in energy requirements when the flow direction is reversed in the case of the latter. [DOI: 10.1115/1.1349120]

1 Introduction

The dynamics of the flow-through porous media is increasingly of pivotal importance to many petroleum engineering applications such as acidizing, fracturing, secondary recovery methods (water and gas flooding, steam injection, in situ combustion), gas cycling, etc. Polymer solutions of different concentrations and rheological properties are increasingly and widely used in these applications.

Polymeric flows exhibit significant elongational components in porous media due to the rapid changes in the cross-sectional area of the pore space in the flow direction. Unexpectedly high flow resistance has been observed in experiments performed with polyethylene oxide and polyacrylamide solutions ([1–3]). The influence of fluid properties such as molecular weight and polymer concentration, rheological properties of solvent, on flow resistance has been investigated theoretically and experimentally ([4–8]). Polymer degradation greatly affects flow resistance ([9]). Experiments also show a substantial increase in flow resistance with respect to that exhibited by Newtonian fluids in the case of nonuniform flows ([10]).

All the investigators mentioned above were concerned with flow through homogeneous porous media. However, in oil engineering applications, flow-through heterogeneous porous media is encountered frequently. In this paper, we report experimental results concerning the flow of Newtonian, viscoelastic, and viscoinelastic liquids in a porous medium with a step change in permeability, that is, two porous media of different permeabilities and same porosity in series. We also develop two theories to predict observed effects based on the KPK (Kutateladze-Popov-Kapakhpasheva) and 8-constant Oldroyd models, respectively.

2 Theoretical Analysis

2.1 A Theory Based on a Viscoinelastic Constitutive Structure. To describe the shear rate-dependent viscosity, we use the concept of fluidity developed by Kutateladze et al. [11]. Fluidity $\varphi(\tau)$ is defined as the reciprocal of the viscosity and is conceived of as depending on the shear stress τ. We may define the range of change of φ with τ as

$$\varphi_0 \leqslant \varphi(\tau) < \varphi_\infty, \quad \varphi \to \varphi_\infty \text{ as } \tau \to \infty, \quad \varphi = \varphi_0 \text{ when } \tau \leqslant \tau_1.$$

For most fluids, $\tau_1 \sim 0$. If we consider shear-thinning fluids, $d\varphi/d\tau > 0$ when $\tau > \tau_1$. Further, $d^2\varphi/d\tau^2 > 0, d\varphi = -\varphi^n d\tau$, which suggests that a phenomenological theory can be constructed based on the dimensionless fluidity φ^* and shear stress τ^*,

$$\varphi^* = \frac{\varphi_\infty - \varphi}{\varphi_\infty - \varphi_0}, \quad \tau^* = \theta \frac{\tau - \tau_1}{\varphi_\infty - \varphi_0},$$

by writing either

$$\varphi^* = e^{-\tau^*}, \quad n = 1, \quad \tau > \tau_1; \quad \varphi^* = [1 - \tau^*(1-n)]^{1/(n-1)}, \quad n \neq 1, \quad \tau > \tau_1.$$

Expanding the first of these in a series we obtain

$$\varphi(\tau) = \varphi_0 + \theta(\tau - \tau_1) - \frac{\theta^2}{2} \frac{(\tau - \tau_1)^2}{\varphi_\infty - \varphi_0} + O(\theta^3).$$

The coefficient θ is called the structural fluidity coefficient. Keeping only the first two terms in this expansion yields the linear fluidity relationship between the fluidity φ and the shear stress τ. Kutateladze et al. [11] show that the linear fluidity law

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 23, 2000; final revision, August 15, 2000. Editor: L. T. Wheeler. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$\varphi=\varphi_0+\theta\tau,\quad \tau_1\sim 0 \tag{1}$$

represents the function $\varphi(\tau)$ rather well for diverse fluids in the range of shear stresses of practical interest.

The structure defined by (1) together with the linear momentum balance yields the equation of motion for the axial flow of a linear fluidity fluid in a circular tube,

$$-\frac{\partial P}{\partial x}+\frac{1}{r}\frac{d}{dr}\left(r\varphi^{-1}\frac{du}{dr}\right)=0.$$

The longitudinal velocity profile is given by

$$u=\frac{\varphi_0\Delta P(R^2-r^2)}{4L}+\frac{\theta(\Delta P)^2(R^3-r^3)}{12L^2},$$

where L and R are the length and radius of the tube, respectively, with ΔP denoting the pressure drop. The average velocity in the tube is defined by

$$\langle u\rangle=\frac{Q}{\pi R^2}=\frac{\varphi_0 R^2\Delta P}{8L}+\frac{\theta R^3(\Delta P)^2}{20L^2}.$$

If the porous medium is regarded as a conduit with a complicated cross section with mean hydraulic radius R_h ([12]), the average velocity becomes

$$\langle u\rangle=\frac{\varphi_0 R_h^2\Delta P}{2L}+\frac{2\theta R_h^3(\Delta P)^2}{5L^2}.$$

The mean hydraulic radius R_h is the ratio of the cross sectional area available to the flow to the wetted perimeter, and is related to the particle diameter D_p and porosity ϵ ([13])

$$R_h=\frac{D_p\epsilon}{6(1-\epsilon)}. \tag{2}$$

The superficial velocity u_0 is defined in terms of the average velocity $\langle u\rangle$,

$$u_0=\frac{\langle u\rangle\epsilon}{C}=\frac{\varphi_0\epsilon^3 D_p^2\Delta P}{72(1-\epsilon)^2LC}+\frac{\theta D_p^3\epsilon^4(\Delta P)^2}{540(1-\epsilon)^3L^2C},$$

where C is the tortuosity factor which accounts for the extra length of the particle paths in the porous bed. Analysis of most of the experimental data suggests $C=25/12=2.0833$ ([13]). When the permeabilities of the pieces of a porous medium vary, the average permeability depends on the manner in which the pieces are arranged ([14]). When N porous media of different permeabilities are in series, the average permeability K of the composite medium is the harmonic mean,

$$K=N\left(\sum_i^N K_i^{-1}\right)^{-1}. \tag{3}$$

According to the Blake-Kozeny-Carman equation, the permeability K_i of the ith medium in series is related to the particle diameter D_{pi} by

$$K_i=\frac{D_{pi}^2\epsilon^3}{180(1-\epsilon)^2}. \tag{4}$$

The actual size of the spherical particles has no influence on the porosity ([15]). We assume that the porosities of the porous media in series is the same even though the permeabilities are different as is the case in our experiments (Section 3)

$$\epsilon=\epsilon_i,\quad i=1,\ldots,N. \tag{5}$$

Substitution of Eqs. (4) and (5) in Eq. (3) leads to the following expression for the average particle diameter:

$$D_p=\sqrt{\frac{N}{\sum_{i=1}^N D_{pi}^2}}\prod_{i=1}^N D_{pi}. \tag{6}$$

Hence, the superficial velocity in a porous medium with different permeabilities in series is given by

$$u_0=\frac{\varphi_0\epsilon^3 D_p^2\Delta P}{150(1-\epsilon)^2L}+\frac{\theta D_p^3\epsilon^4(\Delta P)^2}{1125(1-\epsilon)^3L^2}$$

where D_p is defined in Eq. (6). When $\theta=0$ we obtain the expression for the superficial velocity of a Newtonian fluid ([13]).

Pressure drop in porous media is usually represented in terms of the resistance coefficient Λ, a function of the friction factor f, and the Reynolds number Re,

$$\Lambda=\mathrm{Re}\cdot f,\quad f=\frac{\Delta P D_p\epsilon^3}{\rho u_0^2L(1-\epsilon)},\quad \mathrm{Re}=\frac{D_p u_0\rho\varphi}{1-\epsilon}, \tag{7}$$

where

$$\varphi=\varphi_0+\theta\tau,\quad \tau=\frac{R_h}{C}\frac{\Delta P}{L}=\frac{2D_p\epsilon\Delta P}{25(1-\epsilon)L}.$$

We obtain

$$\mathrm{Re}=\beta(1+0.13\psi)(1+0.08\psi),$$

$$f=\frac{150}{\beta(1+0.13\psi)^2},\quad \Lambda=\frac{150(1+0.08\psi)}{1+0.13\psi},$$

where

$$\beta=\frac{\varphi_0^2\epsilon^3D_p^3\rho\Delta P}{150(1-\epsilon)^3L},\quad \psi=\frac{\theta D_p\epsilon\Delta P}{\varphi_0(1-\epsilon)L}.$$

2.2 A Theory Based on a Viscoelastic Constitutive Structure. We start with the 8-constant Oldroyd model,

$$\tau_{ij}+\lambda_1\left\{\frac{D}{Dt}\tau_{ij}+u_{m,i}\tau_{mj}+u_{m,j}\tau_{mi}\right\}+\mu_o\tau_{kk}d_{ij}$$
$$-\mu_1(\tau_{im}d_{mj}+\tau_{jm}d_{mi})+a_1\delta_{ij}\tau_{km}d_{km}$$
$$=2\eta_0\left\{d_{ij}+\lambda_2\left[\frac{D}{Dt}d_{ij}+u_{m,i}d_{mj}+d_{im}u_{m,j}\right]\right.$$
$$\left.-2\mu_2 d_{im}d_{mj}+2a_2\delta_{ij}d_{mn}d_{mn}\right\},$$

$$\sigma_{ij}+P\delta_{ij}=\tau_{ij},$$

where τ_{ij}, σ_{ij}, and d_{ij} represent the extra stress, total stress, and rate of deformation tensors. It can be shown that this constitutive structure yields for the viscosity η at a given shear rate $\dot\gamma$ ([16])

$$\frac{\eta}{\eta_0}=\frac{1+\sigma_2\dot\gamma^2}{1+\sigma_1\dot\gamma^2}, \tag{8}$$

where η_0 is the zero shear viscosity, and σ_1 together with σ_2 represent combinations of the eight material parameters in the 8-constant Oldroyd model, and can be considered to be independent material parameters in their own right. If this model is to reflect the behavior of real fluids in simple shear, that is, when the shear stress curve always seems to rise monotonically with shear rate, σ_1 and σ_2 must be restricted, $\sigma_2<\sigma_1, \sigma_2/\sigma_1\geqslant 1/9$ ([16]). We meet these criteria in our experiments.

The steady flow problem of an 8-constant Oldroyd fluid in a cylindrical tube has been solved by Williams and Bird [17]. They determine that the pressure drop ΔP in a tube of length L and radius R is given by

$$\frac{\sqrt{\sigma_1}}{2\eta_0}\frac{\Delta P}{L}R = A\sqrt{X}, \quad X = \sigma_1 \dot{\gamma}^2|_{r=R}, \quad n = \frac{\sigma_2}{\sigma_1}, \quad A = \frac{1+nX}{1+X}. \tag{9}$$

An expression for the average velocity $\langle u \rangle$ in terms of the wall shear rate can be developed,

$$\frac{3\sqrt{\sigma_1}\langle u \rangle}{R} = \sqrt{X} - \frac{F}{2(A\sqrt{X})^3},$$

$$F = \frac{1}{2}n^3X^3 - 3n^2(n-1)X + 3n(n-1)(2n-1)\ln(1+X) - \frac{1}{2}X\left(\frac{n-1}{1+X}\right)^2[6n+(7n-1)X]. \tag{10}$$

We use the concept of hydraulic radius as defined in Eq. (2), introduce the average velocity $\langle u \rangle$ and the superficial velocity u_0,

$$u_0 = \langle u \rangle \frac{\epsilon}{C}, \quad \langle u \rangle = \frac{R_h}{4}\dot{\gamma}^2\Big|_{r=R}, \tag{11}$$

and finally obtain the pressure drop ΔP over a length L from Eqs. (9) through (11),

$$\Delta P = \frac{6\eta_0 LA}{R_h}\sqrt{\frac{X}{\sigma_1}}, \tag{12}$$

in terms of the average particle diameter D_p imbedded in R_h given by Eq. (6) and Eq. (2), respectively. Substitution of Eqs. (12), (11), and (8) into Eq. (7) yields the friction factor, the resistance coefficient and the Reynolds number,

$$f = \frac{486\eta_0 C^2B^2A}{\rho R_h^2}\sqrt{\frac{\sigma_1}{X}}, \quad \Lambda = \frac{108CX^2}{2X^2 - A^3F}, \tag{13}$$

$$\mathrm{Re} = \frac{\rho R_h^2 B}{9CA\eta_0}\sqrt{\frac{X}{\sigma_1}}, \quad B = 1 - \frac{F}{2X^2 - A^3}. \tag{14}$$

We note that in contrast to the case of the viscoinelastic constitutive structure used in the previous section the tortuosity factor used in Eq. (11) in the case of a viscoelastic constitutive structure is not constant. It is a strong function of the elastic properties of the fluid together with the properties of the porous medium. In fact, we determine in Section 4 of the present paper that if we choose $C=1.3$ and $C=3.0$ for two solutions of polyacrylamide one percent and two percent in concentration by weight, respectively, we get the best qualitative description of the flow behavior that this model can provide in the context of our experiments, which suggests that the tortuosity factor may assume increasingly larger values with increasingly elastic fluids, that is, higher concentrations of polyacrylamide in this case.

3 Experimental Apparatus and Procedure

The experimental apparatus used in this work is sketched in Fig. 1.

The liquid is supplied from a pressurized feed tank 1 to the test section 2 with the porous medium. The flow rate is controlled through the valve 3 located at the exit of the test tube. At the exit, the test liquid is collected in the graduated cylinder 4 and the mean volumetric flow rate is measured by the weight method. A stainless steel flow cell with an internal diameter of 4.5 cm and a length of 30 cm, packed with glass spheres with diameters of 1000 ± 50 and 3000 ± 50 microns in series has been used for the experiments. Spheres of different diameters fill each half of the tube. To prevent the mixing of glass spheres of different diameters and to contain the packed bed, fine mesh screens are attached to the inlet, exit, and middle sections of the flow cell. To achieve a stable packing arrangement, the spheres have been packed into the cell by mechanical vibration. The flow cell packed with glass spheres of a kind was weighed before and after saturation by distilled water to determine total pore volume. The porosities of both parts of the packed bed were found to be 0.40. As a check the porosities were also determined by weighing the spheres filling the tube. The same value was obtained for the porosities of both porous media in series.

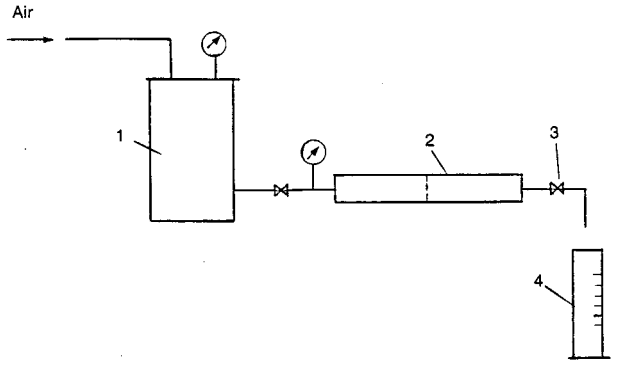


Fig. 1 Experimental set-up: (1) filling tank; (2) test tube; (3) valve, (4) measuring cylinder

To avoid the entrance and exit effects at the ends of the porous medium, the difference in required pressure for two porous media of different lengths (30 and 50 cm) was measured at the same flow rate ([18]). This pressure difference is the pressure drop across a porous medium whose length is the difference between the lengths of the two porous media used (e.g., 20 cm) which is free of all end effects. The permeability of each porous medium in the series arrangement was determined by pumping water at a constant rate into the cell packed with particles of the same diameter and by measuring the pressure drop. Using Darcy's equation,

$$u_0 = \frac{K}{\mu}\frac{\Delta P}{L},$$

the permeabilities of flow cells packed with glass spheres of 1000 μm and 3000 μm in diameter were calculated to be 1.085×10^{-5} cm^2 and 9.665×10^{-5} cm^2, respectively. The average permeability according to Eq. (3) is 1.951×10^{-5} cm^2.

3.1 Liquids Used. Newtonian liquids used in the experiments are distilled water and a glycerol-water solution with a viscosity of 0.0184 Pa·s at 20°C. The non-Newtonian liquids investigated are a polyacrylamide based oil field spacer fluid and aqueous polyacrylamide solutions of one percent and two percent concentration by weight prepared with distilled water and polyacrylamide of molecular weight 5×10^6. The oil field spacer fluid contains 0.6 percent polyacrylamide by weight with several different proprietary additives.

The polymer solutions were carefully prepared by slow shaking to ensure that no degradation of the polymer occurred. Viscosities were measured in a rotational viscometer Rheotest RV-2 at 20°C. The results show that the non-Newtonian viscosity of the spacer fluid and the polymeric solutions of both concentrations (one percent and two percent by weight) is adequately described by the linear fluidity model in the shear stress range relevant to our experiments, $0<\tau\leqslant35$ Pa (Fig. 2).

We present in Fig. 3 viscosity versus shear rate data for the liquids used. We determine the values of the parameters σ_1, σ_2, and η_0 in Eq. (8) for both one percent and two percent solutions of polyacrylamide by a best fit of the data points. We find that for one percent and two percent PAA, ($n=\sigma_2/\sigma_1=0.286$, $\sigma_1=3.10^{-5}$ s^{-2}, $\eta_0=0.133$ Ns/m^2) and ($n=\sigma_2/\sigma_1=0.250$, $\sigma_1=3.10^{-5}$ s^{-2}, $\eta_0=0.150$ Ns/m^2), respectively.

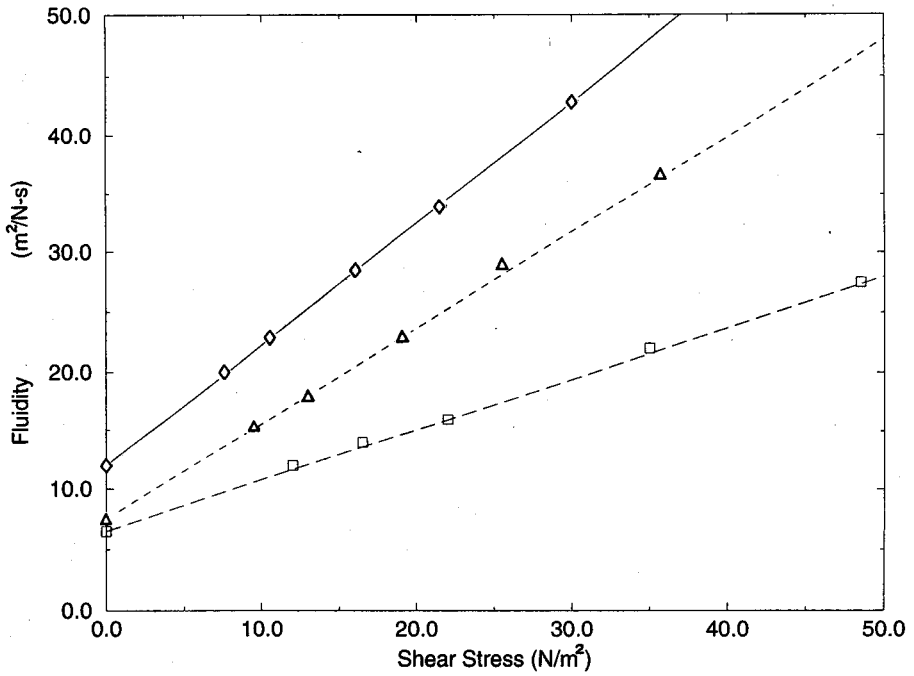


Fig. 2 Fluidity as a function of the shear stress for two aqueous solutions of polyacrylamide (△–one percent, □–two percent) and for the spacer fluid (◇)

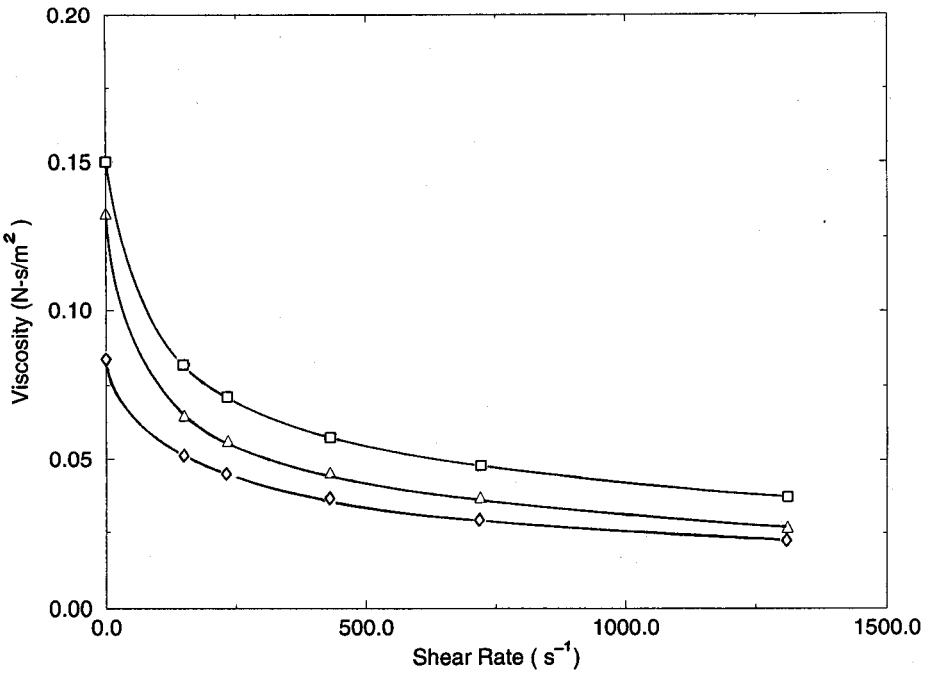


Fig. 3 Viscosity as a function of the shear rate for two aqueous solutions of polyacrylamide (△–1 percent, □–2 percent) and for the spacer fluid (◇)

4 Results and Discussion

Two sets of experiments have been conducted for each fluid:

Tests (I): Flow proceeds through the porous medium with the smaller permeability towards the porous medium with the larger permeability.

Tests (II):Flow proceeds through the porous medium with the larger permeability towards the porous medium with the smaller permeability.

Figures 4 and 5 show the results obtained in the experiments. For water and the glycerol-water solution experimental results are quite well represented by Eq. (7) with a constant resistance coefficient $\Lambda = 150$. We conclude that at Reynolds numbers less than 10, inertial effects are negligible and viscous forces are dominant which implies that the pressure gradient is directly proportional to the superficial velocity. As a consequence, the resistance coefficient assumes constant values. For Reynolds numbers 10 and

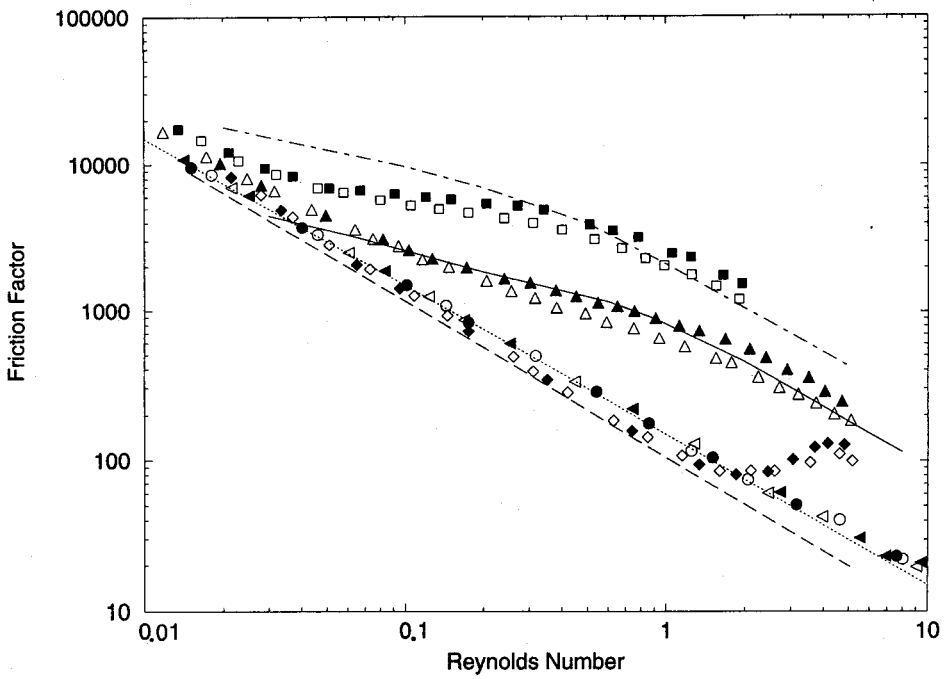


Fig. 4 Friction factor as a function of the Reynolds number. Tests I: (●)–distilled water, (◀)–glycerol/water, (▲)–one percent PAA, (■)–two percent PAA, (♦)–spacer fluid; Tests II: (○)–distilled water, (◁)–glycerol-water, (△)–one percent PAA, (□)–two percent PAA, (◇)–spacer fluid. Solid curves correspond to theoretical predictions obtained using Eqs. (13) and (14) or Eqs. (7): (—) Oldroyd model, one percent PAA, $\eta_0 = 0.133$ Ns/m^2, $n = 0.286$, $\sigma_1 = 3.10^{-5}$ s^{-2}, $C = 1.3$; (— · —) Oldroyd model, two percent PAA, $\eta_0 = 0.150$ Ns/m^2, $n = 0.250$, $\sigma_1 = 3.10^{-5}$ s^{-2}, $C = 3.0$; (– – –) spacer fluid, $\theta/\varphi_0 = 0.06$ m^2/N; (. . .) Newtonian fluids, distilled water, and glycerol/water solution.

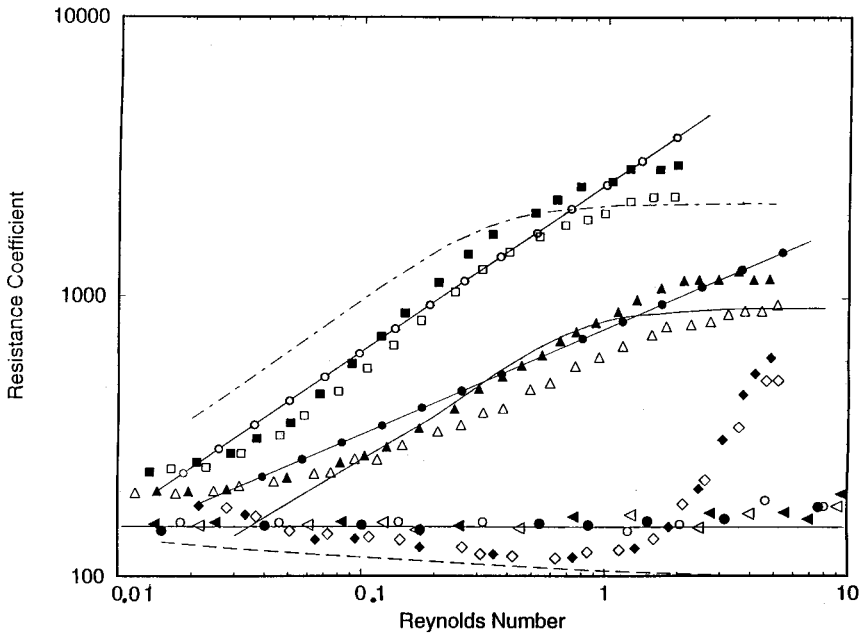


Fig. 5 Resistance coefficient as a function of the Reynolds number. Tests I: (●)–distilled water, (◀)–glycerol/water, (▲)–one percent PAA, (■)–two percent PAA, (♦)–spacer fluid; Tests II: (○)–distilled water, (◁)–glycerol-water; (△)–one percent PAA, (□)–two percent PAA, (◇)–spacer fluid. Solid curves correspond to theoretical predictions obtained using Eqs. (13) and (14) or Eqs. (7): (—) Oldroyd model, one percent PAA, $\eta_0 = 0.133$ Ns/m^2, $n = 0.286$, $\sigma_1 = 3.10^{-5}$ s^{-2}, $C = 1.3$; (— · —) Oldroyd model, two percent PAA, $\eta_0 = 0.150$ Ns/m^2, $n = 0.250$, $\sigma_1 = 3.10^{-5}$ s^{-2}, $C = 3.0$; (-●-) KPK model, one percent, PAA; (–○–) KPK model, 2 percent PAA; (– – –) spacer fluid, $\theta/\varphi_0 = 0.06$ m^2/N; (· · · ·) Newtonian fluids, distilled water, and glycerol/water solution.

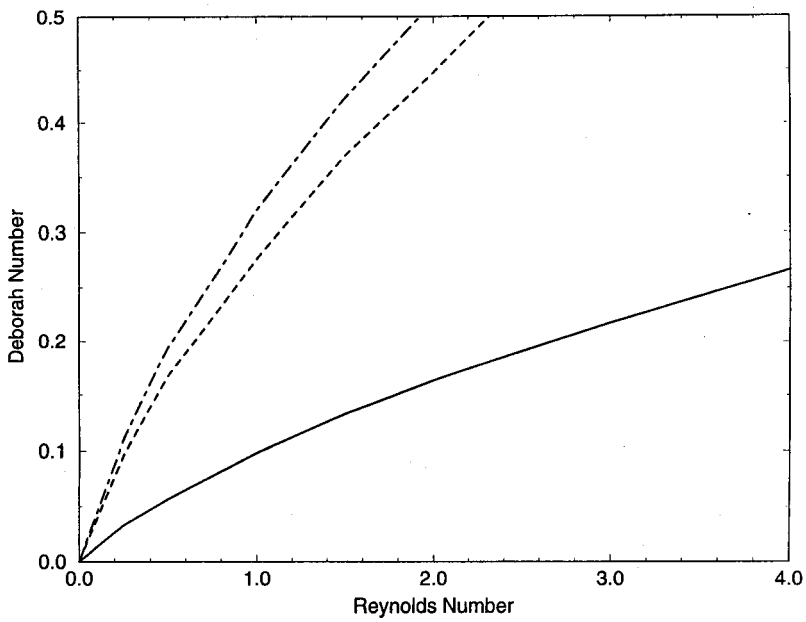


Fig. 6 The variation of the Deborah number (De) with the Reynolds number (Re): (—) spacer fluid; (— — —) one percent PAA; (— · —) two percent PAA

larger, the magnitude of the inertial forces becomes large enough to influence the flow and the resistance coefficient starts increasing. Similar results were obtained in previous investigations ([10]). The friction factor and the resistance coefficient for Newtonian fluids do not depend on flow direction as we obtain the same data when the flow direction is reversed, that is, data for tests I and II are the same.

When the Reynolds number is smaller than 0.04, the pressure drop required in the case of the spacer fluid for the same flow rate as the Newtonian fluid is higher than that of the Newtonian fluid, and increases with decreasing Reynolds numbers. For Reynolds numbers larger than the crossover value 0.04 and smaller than 0.7 the pressure drop required for the same flow rate as the Newtonian fluid is smaller than the Newtonian case, and decreases with increasing Reynolds numbers. At Re=0.7 frictional effects are 20 percent less than the Newtonian case.

A criterion for the importance of elastic effects can be established on the basis of a Deborah number defined as

$$\mathrm{De}=\frac{\eta_0 \tau_{1/2}}{D_p/u_0},$$

where, η_0, $\tau_{1/2}$, D_p, and u_0 are the zero-shear-rate viscosity, the value of the shear stress at which the viscosity η is half of η_0, the particle diameter of the porous media in series given by Eq. (6) and the superficial velocity, respectively ([12]). The variation of De versus Re is shown in Fig. 6. Sadowski [19] has established experimentally that elastic effects seem to set in at about De $\sim$0.1. This critical Deborah number corresponds to a Reynolds number equal to one in the case of the spacer fluid, Fig. 6. Re $\sim$1 is approximately the Reynolds number at which deviations from the predictions of the linear fluidity model start, Fig. 5. We are led to conclude that Sadowski's criterion works quite well for shear-thinning fluids of this type. When the Reynolds number exceeds 2, both the friction factor and the resistance coefficient increase rapidly and at Re=5 they are at least three times larger than the data for Newtonian fluids. The friction factor and the resistance coefficient obtained experimentally are reasonably well predicted theoretically by the linear fluidity model when $\theta/\varphi_0 = 0.066\ \mathrm{m}^3/\mathrm{N}$ for Re$<$1. Flow direction again has no effect on the flow parameters in the case of the spacer fluid, and the pressure drop required for the same flow rate in tests I and II is the same (Fig. 5).

In the case of the polymeric fluids, the flow resistance becomes higher and the pressure drop larger compared with Newtonian fluids as the Reynolds number is increased. The resistance coefficient increases dramatically with concentration. For instance, at Re=1.0, the resistance coefficient Λ and the pressure drop for the one percent solution in the case of tests (II) are four times larger than the Newtonian case. But when the concentration is doubled (two percent solution), the resistance coefficient becomes 12 times larger than that for the Newtonian liquid. Frictional effects increase at a much faster rate than the rate of increase of the concentration, and much larger pressure gradients are required at the same Reynolds number, that is, for the same flow rate. At any concentration, the resistance coefficient increases rapidly with increasing Reynolds numbers, and the rate of increase of the resistance coefficient becomes steeper as the concentration increases at the same Reynolds number. The resistance coefficient seems to be constant at very small and relatively large Reynolds numbers with a steep change in between. As the concentration increases, the Reynolds numbers denoting the upper and lower boundaries of the region where the change takes place become smaller. For instance, for the one percent and two percent PAA, Λ is constant when $\mathrm{Re}_1<0.0115$ and, $\mathrm{Re}_1<0.01$, respectively. But, it shows an order of magnitude change with increasing Reynolds numbers when Re$>$Re$_1$ and becomes constant again when Re$>$2 and, Re $>$1.2, respectively.

For the same polymeric solution, flow resistance depends on the flow direction. Experimental data shows that energy loss is higher if the polymeric solution flows first through the medium with the smaller permeability K_1 rather than through the section of the flow cell with the larger permeability K_2 first (Figs. 4 and 5). The dependence of the resistance coefficient on the flow direction is quite distinct, and the difference may be as large as 25–35 percent. Flow resistance increases with increasing Reynolds numbers at a fixed concentration when flow direction is changed from $K_2 \rightarrow K_1$ to $K_1 \rightarrow K_2 (K_1 < K_2)$. Flow resistance also shows strong dependence on concentration when flow direction is reversed. That is, at the same Reynolds number increasingly larger pressure

drops are required for the same flow rate with increasing concentration, or equivalently increasing elasticity. As the concentration becomes larger, the difference in energy requirements for the same flow rate in the $K_2 \rightarrow K_1$ and $K_1 \rightarrow K_2 (K_1 < K_2)$ directions starts becoming noticeable at smaller Reynolds numbers. For instance, whereas the critical Reynolds number for the one percent solution is around 0.11, it recedes to 0.065 for the two percent solution. Assuming that the criterion De$\sim$0.1 for the onset of elasticity effects applies as well to the polyacrylamide solutions we determine the critical Reynolds numbers of $\sim$0.2 and $\sim$0.15 for the one percent and two percent solutions of polyacrylamide, respectively, from Fig. 6. These critical Reynolds numbers correspond approximately to the flow rate at which a significantly higher pressure drop is required when the fluid flows first through the lower permeability medium, in particular for the one percent polyacrylamide solution, Fig. 5.

The predictions of the 8-constant Oldroyd model using the viscosity functions for the polyacrylamide solutions determined using the data in Fig. 3 are also shown in Fig. 5 together with the predictions of the inelastic KPK model using the fluidity data given in Fig. 2. There is better qualitative agreement with the data in the case of the 8-constant Oldroyd model for the lower concentration fluid except for very small Reynolds numbers where the plateau is not predicted at all in the case of either concentration. The inelastic KPK model also predicts the trend for both concentrations, that is the steep increase in the resistance coefficient with increasing Reynolds numbers, but fails to predict the plateaus for the resistance coefficient at both high (Re$>$1) and low (Re $<$0.05) Reynolds numbers. The fact that both constitutive models can qualitatively describe the steep increase in the resistance coefficient with increasing Reynolds numbers between the high and low plateaus may not be surprising after all and may be expected as the average velocity expression developed for the 8-constant Oldroyd model (Eq. (10)) does not reflect elastic effects. What is surprising is that the 8-constant Oldroyd model can predict the upper plateau regardless of the particular value of the tortuosity factor used whereas the KPK model completely fails to do so in either concentration. We remark that the tortuosity factor C when used in conjunction with viscoelastic models (Eq. $(11)_1$) becomes a function of the elastic properties of the fluid. The theoretical curves in Fig. 5 for the 8-constant Oldroyd model have been obtained by assigning values to C for the best qualitative description of the flow data. We find that C must be 1.3 and 3.0 in the case of one percent and two percent PAA, respectively, which suggests that the tortuosity factor may increase with increasing elasticity. It is clear that the theory presented here is incapable of describing the considerable increase in pressure drop when flow direction is switched from $K_2 \rightarrow K_1$ to $K_1 \rightarrow K_2 (K_1 < K_2)$. At this point in time, we cannot offer a theory to explain this novel phenomenon which is closely governed by the behavior of the test liquids in elongational flows. Due to rapid changes of the cross-sectional area of the pore space in the flow direction, flow through packed beds exhibits large elongational components of velocity. An increase in the extension rate may cause an increase in the apparent viscosity and in the flow resistance in elongational flows of high molecular weight polymer solutions.

Experimental results obtained concerning the energy loss with a step change in permeability can be heuristically extended to porous media arranged in series over a length L with increasing permeabilities $K_1 < K_2 < \cdots < K_{N-1} < K_N$. If N is taken large enough, we may obtain a close enough approximation to a nonhomogeneous porous media with a continuous permeability gradient in the flow direction. Experiments imply therefore that in an anisotropic, nonhomogeneous medium with constant porosity the energy loss is considerably higher for the same flow rate if flow proceeds in the direction of the positive permeability gradient.

Conclusions

Experiments were run in a porous medium with a step change in permeability represented by a flow cell with two adjoining porous media in series of equal length and of different permeabilities and same porosity. The results can be extended heuristically to an anisotropic porous medium of constant porosity and continuous permeability gradient.

We find that for two Newtonian liquids, water and a water/glycerol solution, the resistance coefficient is constant when inertial effects are negligible, Re$<$10. A highly shear-thinning oilfield spacer fluid requires less energy (smaller resistance coefficient) than a Newtonian fluid for the same volume flow rate for Reynolds numbers between 0.08 and 1.5. Elastic effects start becoming important at a critical Deborah number of 0.1, or equivalently at a critical Reynolds number of one. The pressure drop required for the same volume flow rate is higher than the Newtonian case for Re$>$2, and increases rapidly with increasing Reynolds numbers.

Experiments with two polymeric solutions (one percent and two percent PAA) show that at all Reynolds numbers the pressure drop required for the same volume flow rate is much higher than the Newtonian liquid of the same zero shear rate viscosity. Energy loss increases with increasing Reynolds numbers to level off at a Reynolds number of $O(1)$. The pressure drop required at that Reynolds number is an order of magnitude larger than the pressure drop for the Newtonian liquid, and increases with increasing concentration.

The pressure drop required for the same volume flow rate when the flow direction is switched from $K_1 \rightarrow K_2$ to $K_2 \rightarrow K_1 (K_1 > K_2)$, where K_i represents the permeability, is considerably higher. The difference in energy requirements increases with increasing Reynolds numbers at any concentration to reach an almost constant value at Re$\sim O(1)$. At Re$\sim O(1)$ it is as much as 25 percent and 35 percent for the one percent and two percent solutions of polyacrylamide, respectively.

The Newtonian behavior is well predicted theoretically when inertial effects are negligible. Two theories built on the inelastic KPK (Kutateladze-Popov-Kapakhpasheva) and viscoelastic Oldroyd models show only qualitative agreement with experimental data for the nonlinear fluids used. But, predicting the difference in energy requirements for viscoelastic fluids even qualitatively when the flow direction is reversed remains a challenge.

Nomenclature

d_{ij} = rate of deformation tensor
D_p = particle diameter
C = tortuosity coefficient
f = friction factor
L = length of the test tube
K = permeability
R = radius of the test tube
R_h = hydraulic radius
u = longitudinal (axial) velocity
$\langle u \rangle$ = average velocity
u_0 = superficial velocity
Q = volume flow rate
P = pressure
φ = fluidity or reciprocal of viscosity
φ_0 = lower bound of fluidity below a limiting value τ_1 of the shear stress
φ_∞ = upper bound of fluidity reached at large shear stresses
θ = structural fluidity coefficient
ϵ = porosity
Λ = resistance coefficient
ρ = density
σ_{ij} = total stress tensor
τ_{ij} = extra-stress tensor
τ = shear stress
$\tau_{1/2}$ = shear stress at which $\eta = \eta_0/2$
$\dot{\gamma}^2$ = shear rate
η = viscosity
η_0 = zero shear viscosity

σ_1 = material parameter—a function of the eight constants in the Oldroyd Model
σ_2 = material parameter—a different function of the eight constants in the Oldroyd Model
n = the ratio of σ_2 over σ_1
δ_{ij} = Kronecker tensor
Re = Reynolds number

References

[1] Dauben, D. L., and Menzie, D. E., 1967, ''Flow of Polymer Solutions Through Porous Media,'' J. Pet. Technol., pp. 1065–1072.
[2] Marshall, R. G., and Metzner, A. B., 1967, ''Flow of Viscoelastic Fluids Through Porous Media,'' Ind. Eng. Chem. Fundam., **6**, No. 3, pp. 393–400.
[3] James, D. F., and McLaren, D. R., 1975, ''The Laminar Flow of Dilute Polymer Solutions Through Porous Media,'' J. Fluid Mech., **70**, pp. 733–752.
[4] Mirzajanzade, A. Kh., 1959, *Hydrodynamics of Viscoplastic and Viscous Fluids Used in Oil Production*, Aznefteizdat, Baku (in Russian).
[5] Christopher, R. K., and Middleman, S., 1965, ''Power-Law Flow Through a Packed Tube,'' Ind. Eng. Chem. Fundam., **4**, pp. 422–426.
[6] Kulicke, W. M., and Haas, R., 1984, ''Flow Behavior of Dilute Polyacrylamide Solutions Through Porous Media. 1. Influence of Chain Length, Concentration and Thermodynamic Quality of the Solvent,'' Ind. Eng. Chem. Fundam., **23**, No. 3, pp. 308–315.
[7] Dharmadhikari, R. V., and Kale, D. D., 1985, ''Flow of Non-Newtonian Fluids Through Porous Media,'' J. Chem. Eng. Sci., **40**, No. 3, pp. 527–529.
[8] Bakhtiyarov, S. I., and Suleymanzade, N. S., 1991, ''Herschel-Bulkley Fluid Flow in Porous Medium,'' J. Oil Gas, Baku, No. 7, pp. 50–53 (in Russian).
[9] Farinato, R. S., and Yen, W. S., 1987, ''Polymer Degradation in Porous Media Flow,'' J. Appl. Polym. Sci., **33**, pp. 2353–2368.
[10] Rodriguez, S., Romero, C., Sargenti, M. L., Müller, A. J., Sáez, A. E., and Odell, J. A., 1993, ''Flow of Polymer Solutions Through Porous Media,'' J. Non-Newtonian Fluid Mech., **49**, pp. 63–85.
[11] Kutateladze, S. S., Popov, V. I., and Khabakhpasheva, E. M., 1966, ''The Hydrodynamics of Fluids of Variable Viscosity,'' J. Appl. Mech. Tech. Phys., **1**, pp. 45–49 (in Russian).
[12] Bird, R. B., Armstrong, R. C., and Hassager, O., 1987, *Dynamics of Polymeric Liquids*, Vol. 1, Fluid Mechanics, 2nd Ed. John Wiley and Sons, New York.
[13] Bird, R. B., Stewart, W. E., and Lightfoot, E. N., 1960, *Transport Phenomena*, John Wiley and Sons, New York.
[14] Scheidegger, A. E., 1974, *The Physics of Flow Through Porous Media*, University of Toronto Press, New York.
[15] Muskat, M., and Wyckoff, K. D., 1946, *The Flow of Homogeneous Fluids Through Porous Media*, Edwards, Ann Arbor, MI.
[16] Tanner, R. I., 1985, *Engineering Rheology*, Clarendon Press, Oxford, U.K.
[17] Williams, M. C., and Bird, R. B., 1962, ''Steady Flow of an Oldroyd Viscoelastic Fluid in Tubes, Slits, and Narrow Annuli,'' AIChE J., **8**, pp. 378–382.
[18] Fredrickson, A. G., 1964, *Principles and Applications of Rheology*, Prentice-Hall, Englewood Cliffs, N.J.
[19] Sadowski, T. J., 1965, ''Non-Newtonian Flow Through Porous Media. II. Experimental,'' Trans. Soc. Rheol., **9**, No. 2, pp. 251–271.

Rotary Inertia in the Classical Nonlinear Theory of Shells and the Constitutive (Non-Kinematic) Kirchhoff Hypothesis

J. G. Simmonds
Department of Civil Engineering,
University of Virginia,
Charlottesville, VA 22903
Fellow ASME

A general nonlinear theory of isothermal shells is presented in which the only approximations occur in the conservation of energy and in the consequent constitutive relations, which include expressions for the shell velocity and spin. No thickness expansions or kinematic hypotheses are made. The introduction of a dynamic mixed-energy density avoids ill-conditioning associated with near inextensional bending or negligible rotational momentum. It is shown that a variable scalar rotary inertia coefficient exists that minimizes the difference between the exact kinetic-energy density and that delivered by shell theory. Finally, it is shown how specialization of the dynamic mixed-energy density provides a simple and logical way to introduce a constitutive form of the Kirchhoff hypothesis, thus avoiding certain unnecessary constraints (such as no thickness changes) imposed by the classical kinematic Kirchhoff hypothesis. [DOI: 10.1115/1.1357870]

1 Introduction

In a *classical* theory of shells the basic kinetic ingredients are a stress resultant tensor **N** and a stress couple tensor **M**. That is, no moments higher than the zeroth and first appear.

Chapter VIII of Libai and Simmonds [1] shows that exact integral-impulse equations for a shell follow naturally and simply, *without any kinematic assumptions*, from analogous equations for a three-dimensional continuum. The descent from three to two dimensions leads to definitions of **N** and **M** in terms of integrals through the (possibly variable) shell thickness of the three-dimensional first Piola-Kirchhoff stress tensor and to definitions of translational and rotational momenta, **L** and **R**, in terms of thickness integrals of the velocity and the deformed position of the three-dimensional shell-like body.

From the local (differential) form of the equations of motion of a shell, which are *exact* and contain the translational and rotatory inertia terms, $\dot{\mathbf{L}}$ and $\dot{\mathbf{R}}$, where the superior dot denotes differentiation with respect to time, one obtains a power identity by introducing a two-dimensional velocity and spin, **v** and $\boldsymbol{\omega}$,—both initially undefined—and applying Green's theorem. This produces two-dimensional extensional-shear and bending strain tensors, **E** and **K**, whose local rates, $\mathbf{E}^*$ and $\mathbf{K}^*$, are conjugate to **N** and **M**, respectively. However, in dealing with energy and constitutive relations, it is convenient to introduce ''back-rotated'' stress resultants and couples, **n** and **m**, and extensional-shear and bending strains, **e** and **k**, as explained later.

The inclusion of rotary inertia in shell theories derived from the three-dimensional equations of motion by integrating through the thickness has been considered by Habip and Ebcioglu [2], Habip [3], Naghdi [4], Antman [5], and Libai and Simmonds [1], among others. All the definitions in [2–5] resulting in a *finite* set of shell equations use a priori kinematic hypotheses on the three-dimensional motion of the shell; see Eq. (27) of [2], Eq. (19) of [3], Eq. (7.5) and (12.25) of [4], and equation (XIV.9.11) of [5]. No such kinematic hypotheses are introduced in [1]. The present note modifies and extends the approaches in [1–5] in several ways:

1 In place of a constant rotary inertia coefficient (or, more generally, a constant rotary inertia tensor), I introduce a *variable* rotary inertia coefficient, depending on position and time and chosen so that *the kinetic energy of the shell is as close as possible to the exact three-dimensional kinetic energy*. (Note that if one attempts to describe the gross motion of any deformable body by the motion of its center of mass and some mean rotation about the center of mass only, then, in general, it is impossible to account for all the kinetic energy of the body. Think of a ball that undergoes purely radial motion.)

2 I define an isothermal *elastodynamic* shell as one for which there exists an energy density ε (kinetic plus elastic), depending on **L**, **R**, **e**, **k**, and possibly **y**, the position on **M**, a reference surface of mass to be specified presently. Then, using an idea suggested to me independently by Ladevèze and Makowski, I define $\mathbf{v} \equiv \varepsilon_{,\mathbf{L}}$ and $\boldsymbol{\omega} \equiv \varepsilon_{,\mathbf{R}}$. That is, I consider inertia terms as part of the constitutive relations.

3 By means of a partial Fenchel-Lagrange transformation, I introduce a *mixed* elastodynamic-energy density $\psi(\mathbf{L},\boldsymbol{\omega},\mathbf{n},\mathbf{k};\mathbf{y})$. One motivation is that, in the static theory of shells, the associated Euler equations are well-conditioned in the sense that they require no special treatment in the extreme cases of membrane or bending-dominated behavior ([1,6]).

4 Another motivation for introducing the dynamic mixed-energy density ψ is that a rather obvious specialization yields what Libai and [1] have called the *constitutive Kirchhoff hypothesis*—a hypothesis based, not on a priori kinematic assumptions, but on certain approximations in the constitutive relations. The latter, by their very nature, can only approximate actual material behavior. Thus, *all* approximations in isothermal shell theory are thrown into those parts of the theory that are unavoidably approximate.

2 Geometry and Exact Equations of Motion of a Shell

A shell may be defined as a material body such that the *initial position* **x** of each material particle belongs to a family of nonintersecting surfaces $\mathcal{S}$ in three-dimensional Euclidean space. That is,

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Aug. 7, 2000; final revision, Oct. 19, 2000. Associate Editor: R. C. Benson. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$\mathbf{x}\in\mathcal{S}(\mathbf{y},\zeta),\quad \mathbf{y}\in\mathcal{R},\quad -H\leqslant\zeta\leqslant H, \tag{1}$$

where $\mathcal{R}$ is some reference surface and H is some given positive constant with the dimensions of LENGTH. (In general, the shell may have variable thickness and ζ is *not* distance along the normal to $\mathcal{R}$.) The *motion* of the shell is then a vector-valued function of the form $\overline{\mathbf{x}}(\mathbf{y},\zeta,t)$.

As in [1], it simplifies the equations that follow and entails no loss of generality to take $\mathcal{R}=\mathcal{M}(\mathbf{y},0)$, where $\mathcal{M}$ is the time-dependent *surface of mass*, whose position is defined by

$$m(\mathbf{y})\overline{\mathbf{y}}(\mathbf{y},t)\equiv\int_{-}^{+}\rho\overline{\mathbf{x}}\mu d\zeta. \tag{2}$$

Here,

$$m(\mathbf{y})\equiv\int_{-}^{+}\rho\mu d\zeta \tag{3}$$

is the mass per unit area of $\mathcal{R}$, $\int_{-}^{+}\equiv\int_{-H}^{H}$, ρ is the initial mass/volume of the shell material, μ is a geometric factor defined (implicitly) by the relation $dV=\mu(\mathbf{y},\zeta)dRd\zeta$, where dV is a differential element of volume in the shell at $(\mathbf{y},\zeta)$, and dR is a differential element of area on $\mathcal{R}$. (See the footnote on p. 456 of [1] for an explicit formula for μ.) For convenience, assume that $\mathcal{R}$ is smooth and orientable with an associated unit normal $\mathbf{b}$ at each point $\mathbf{y}$. As shown in [1], the local (differential) equations of motion of a shell, under suitable smoothness conditions, can be written in coordinate-free form as

$$\boldsymbol{\nabla}\cdot\mathbf{N}^{T}+\mathbf{p}=\dot{\mathbf{L}}\quad\text{and}\quad\boldsymbol{\nabla}\cdot\mathbf{M}^{T}-(\mathbf{N}\cdot\boldsymbol{\nabla})\times\overline{\mathbf{y}}+\mathbf{l}=\dot{\mathbf{R}}. \tag{4}$$

Here, ''T'' denotes ''transpose'' and $\boldsymbol{\nabla}$ is the surface del operator, defined by

$$d\mathcal{F}\equiv d\mathbf{y}\cdot\boldsymbol{\nabla}\mathcal{F}, \tag{5}$$

where $d\mathcal{F}$ is the differential of any suitably smooth scalar, vector, or tensor defined on $\mathcal{R}$.

Expressions for $\mathbf{N}$ and $\mathbf{M}$ in terms of thickness integrals of the three-dimensional first Piola-Kirchhoff stress tensor are given in Chapter VIII of [1]. These tensors have the following physical meaning: let $\boldsymbol{\nu}ds$ denote an oriented element of arc on $\mathcal{R}$ at $\mathbf{y}$, where $\boldsymbol{\nu}$ is a unit vector perpendicular to $\mathbf{b}$ at $\mathbf{y}$ and ds is arc length. Then $\mathbf{N}\cdot\boldsymbol{\nu}$ and $\mathbf{M}\cdot\boldsymbol{\nu}$ are, respectively, the net force and moment exerted across the strip generated by the differential element of area μds as ζ goes from $-H$ to H by the material lying on the same side of the strip as $\boldsymbol{\nu}$.

The present study is mainly concerned with the inertia terms in (4) which are defined as

$$\mathbf{L}(\mathbf{y},t)\equiv\int_{-}^{+}\rho\dot{\overline{\mathbf{x}}}\mu d\zeta\quad\text{and}\quad\mathbf{R}(\mathbf{y},t)\equiv\int_{-}^{+}\rho\overline{\mathbf{z}}\times\dot{\overline{\mathbf{z}}}\mu d\zeta. \tag{6}$$

In $(6)_2$,

$$\overline{\mathbf{z}}(\mathbf{y},\zeta,t)\equiv\overline{\mathbf{x}}-\overline{\mathbf{y}} \tag{7}$$

is the *deviation* from the surface of mass $\mathcal{M}$.

3 The Mechanical Power Identity and Strains

Let $\mathbf{v}$ and $\boldsymbol{\omega}$ be a shell velocity and spin, as yet undefined. Then take the dot product of $(4)_1$ with $\mathbf{v}$, the dot product of $(4)_2$ with $\boldsymbol{\omega}$, add the resulting equations, integrate over $\mathcal{R}$, and use Green's Theorem to remove derivatives on $\mathbf{N}$ and $\mathbf{M}$. This leads to the *mechanical power identity*

$$\mathcal{W}=I+\mathcal{D}, \tag{8}$$

where

$$\mathcal{W}=\int_{\partial\mathcal{R}}(\mathbf{v}\cdot\mathbf{N}+\boldsymbol{\omega}\cdot\mathbf{M})\cdot\mathbf{v}ds+\int_{\mathcal{R}}(\mathbf{v}\cdot\mathbf{f}+\boldsymbol{\omega}\cdot\mathbf{l})dR \tag{9}$$

is the *external mechanical power*,

$$I=\int_{\mathcal{R}}(\mathbf{v}\cdot\dot{\mathbf{L}}+\boldsymbol{\omega}\cdot\dot{\mathbf{R}})dR \tag{10}$$

is the *inertial power* and

$$\mathcal{D}=\int_{\mathcal{R}}[\mathbf{N}:(\boldsymbol{\nabla}\mathbf{v}+\boldsymbol{\nabla}\overline{\mathbf{y}}\times\boldsymbol{\omega})+\mathbf{M}:\boldsymbol{\nabla}\boldsymbol{\omega}]dR \tag{11}$$

is the *deformation power*.

To define extensional-shear and bending strains, imagine the motion of the shell at each point $\mathbf{y}$ carrying $\boldsymbol{\nabla}\mathbf{y}^{T}=\boldsymbol{\nabla}\mathbf{y}$ into $\boldsymbol{\nabla}\overline{\mathbf{y}}^{T}$ by a rigid-body rotation, represented by the *rotator* $\mathbf{Q}(\mathbf{y},t)$, followed by a *stretch*, $\mathbf{1}+\mathbf{E}(\mathbf{y},t)$, where $\mathbf{1}=\boldsymbol{\nabla}\mathbf{y}+\mathbf{bb}$ is the three-dimensional identity tensor. That is,

$$\boldsymbol{\nabla}\overline{\mathbf{y}}^{T}=(\mathbf{1}+\mathbf{E})\cdot\mathbf{Q}\cdot\boldsymbol{\nabla}\mathbf{y},\quad\mathbf{E}\cdot\mathbf{Q}\cdot\mathbf{b}=\mathbf{0}. \tag{12}$$

(This represents a *modified* version of the polar decomposition theorem because, in general, $\mathbf{E}\neq\mathbf{E}^{T}$.) The rotator is defined in terms of $\boldsymbol{\omega}$—which itself is still undefined—as the unique solution of the following differential equation with initial condition:

$$\dot{\mathbf{Q}}=\boldsymbol{\omega}\times\mathbf{Q},\quad\mathbf{Q}(\mathbf{y},0)=\mathbf{1}. \tag{13}$$

Note that $\mathbf{Q}\cdot\mathbf{Q}^{T}\equiv\mathbf{1}$ implies that

$$\boldsymbol{\omega}\times\mathbf{1}=\dot{\mathbf{Q}}\cdot\mathbf{Q}^{T}. \tag{14}$$

From (12),

$$(\boldsymbol{\nabla}\overline{\mathbf{y}}\cdot\mathbf{Q})^{T}=\mathbf{Q}^{T}\cdot\boldsymbol{\nabla}\overline{\mathbf{y}}^{T}=(\mathbf{1}+\mathbf{e})\cdot\boldsymbol{\nabla}\mathbf{y},\quad\text{where}\quad\mathbf{e}\equiv\mathbf{Q}^{T}\cdot\mathbf{E}\cdot\mathbf{Q}. \tag{15}$$

Thus, by $(13)_1$ and the identity $\boldsymbol{\nabla}\overline{\mathbf{y}}\cdot(\boldsymbol{\omega}\times\mathbf{Q})\equiv(\boldsymbol{\nabla}\overline{\mathbf{y}}\times\boldsymbol{\omega})\cdot\mathbf{Q}$, the time derivative of (15) yields

$$(\boldsymbol{\nabla}\dot{\overline{\mathbf{y}}}+\boldsymbol{\nabla}\overline{\mathbf{y}}\times\boldsymbol{\omega})\cdot\mathbf{Q}=\boldsymbol{\nabla}\mathbf{y}\cdot\dot{\mathbf{e}}^{T}=(\mathbf{1}-\mathbf{bb})\cdot\dot{\mathbf{e}}^{T}=\dot{\mathbf{e}}^{T}. \tag{16}$$

The last step in this equation follows from $(12)_2$ and $(15)_2$, which imply that $\mathbf{b}\cdot\mathbf{e}^{T}=0$ and hence that $\mathbf{b}\cdot\dot{\mathbf{e}}^{T}=0$. With the definitions

$$\mathbf{v}\equiv\dot{\overline{\mathbf{y}}}\quad\text{and}\quad\mathbf{n}\equiv\mathbf{Q}^{T}\cdot\mathbf{N}, \tag{17}$$

it follows that

$$\mathbf{N}:(\boldsymbol{\nabla}\mathbf{v}+\boldsymbol{\nabla}\overline{\mathbf{y}}\times\boldsymbol{\omega})=\mathbf{n}:\dot{\mathbf{e}}. \tag{18}$$

Because of the factor $\mathbf{Q}^{T}$ in $(15)_2$ and $(17)_2$, $\mathbf{e}$ *and* $\mathbf{n}$ *may be called a back-rotated extensional-shear strain and stress resultant, respectively*.

As in [1], the bending strain tensor $\mathbf{K}$ is defined by the spatial analogue of (14):

$$\mathbf{K}\times\mathbf{1}\equiv\boldsymbol{\nabla}\mathbf{Q}\cdot\mathbf{Q}^{T}. \tag{19}$$

As shown in [1], (14) and (19) imply the relation

$$\boldsymbol{\nabla}\boldsymbol{\omega}=\dot{\mathbf{K}}+\mathbf{K}\times\boldsymbol{\omega}. \tag{20}$$

Note by $(13)_1$ and the identity $\mathbf{K}\cdot(\boldsymbol{\omega}\times\mathbf{Q})\equiv(\mathbf{K}\times\boldsymbol{\omega})\cdot\mathbf{Q}$ that

$$\boldsymbol{\nabla}\boldsymbol{\omega}\cdot\mathbf{Q}=(\dot{\mathbf{K}}\cdot\mathbf{Q})^{\cdot}. \tag{21}$$

Thus, with the definitions

$$\mathbf{m}\equiv\mathbf{Q}^{T}\cdot\mathbf{M}\quad\text{and}\quad\mathbf{k}\equiv\mathbf{K}\cdot\mathbf{Q}, \tag{22}$$

it follows from (21) and (22) that

$$\mathbf{M}:\boldsymbol{\nabla}\boldsymbol{\omega}=\mathbf{m}:\dot{\mathbf{k}}. \tag{23}$$

The *motion* of a shell is the pair of vector-tensor functions $\{\overline{\mathbf{y}}(\mathbf{y},t),\mathbf{Q}(\mathbf{y},t)\}$. Because the nine parameters that define the orthogonal tensor $\mathbf{Q}$ *are not all independent, it may be convenient to express* $\mathbf{Q}$ *in terms of a finite rotation vector* $\boldsymbol{\psi}$ ([1]), *in which case the motion of the shell may be defined as the pair of vector-valued functions* $\{\overline{\mathbf{y}}(\mathbf{y},t),\boldsymbol{\psi}(\mathbf{y},t)\}$. *Given a suitably smooth motion, a compatible set of spins and strains may be computed from* (12), (14), *and* (19). *On the other hand, if one wishes to work with*

the extensional-shear and bending strains $\mathbf{E}$ *and* $\mathbf{K}$ (*or, alternatively, with* $\mathbf{e}$ *and* $\mathbf{k}$), *then compatibility conditions must be satisfied, as explained in* [1].

4 An Elastodynamic Shell

For conciseness, let

$$\Lambda \equiv (\mathbf{L}, \mathbf{R}, \mathbf{e}, \mathbf{k}) \quad \text{and} \quad \Xi \equiv (\mathbf{v}, \boldsymbol{\omega}, \mathbf{n}, \mathbf{m}) \tag{24}$$

denote, respectively, an *argument list* and its *conjugate*. Then an isothermal shell is *elastodynamic* if there exists an energy density ε depending on Λ and possibly position $\mathbf{y}$ such that

$$\Xi \cdot \dot{\Lambda} = \dot{\varepsilon}(\Lambda, \mathbf{y}) \tag{25}$$

and such that the equations of motion (4) are satisfied. These latter conditions may always be met by a suitable choice of the surface load $\mathbf{p}$ and the surface couple $\mathbf{l}$.

Now suppose that the unknowns composing Ξ also depend on Λ and $\mathbf{y}$ only. Then, because $\dot{\Lambda}$ may always be assigned any value at any fixed point on $\mathcal{R}$ and at any fixed time, it follows from (25) that $\Xi = \varepsilon_{,\Lambda}$, providing ε is a differentiable function of Λ. That is, (25) implies the *constitutive relations*

$$\mathbf{v} = \varepsilon_{,\mathbf{L}}, \quad \boldsymbol{\omega} = \varepsilon_{,\mathbf{R}}, \quad \mathbf{n} = \varepsilon_{,\mathbf{e}}, \quad \mathbf{m} = \varepsilon_{,\mathbf{k}}. \tag{26}$$

This relation *defines* $\boldsymbol{\omega}$, whereas the expression for $\mathbf{v}$ must be consistent with $(17)_1$.

5 A Dynamic Mixed-Energy Density

Despite the popularity of minimum energy formulations both in the finite element literature and from a mathematical standpoint (uniqueness and convergence proofs), there are compelling reasons to work with a *dynamic mixed-energy density* $\psi(\mathbf{L}, \boldsymbol{\omega}, \mathbf{n}, \mathbf{k}; \mathbf{y})$ instead of with the positive definite energy density $\varepsilon(\mathbf{L}, \mathbf{R}, \mathbf{e}, \mathbf{k}; \mathbf{y})$. First, depending on the external loading and kinematic boundary conditions, a shell may experience (nearly) inextensional bending. In this circumstance, the constitutive relation $\mathbf{n} = \varepsilon_{,\mathbf{e}}$ becomes ill-conditioned. The components of the stress resultant tensor are then, essentially, reactive quantities, best computed from the equations of motion. (Think of the classical theory of planar curved beams: the axial strain vanishes, by assumption, but not, in general, the axial force.) And second, by working with a dynamic mixed-energy density, the a priori *kinematic Kirchhoff hypothesis* (''normals to the undeformed reference surface deform without stretching into normals to the deformed reference surface'') may be replaced by the (dynamic) *constitutive Kirchhoff hypothesis*: ψ does not depend on the spin $\boldsymbol{\omega}$ or the transverse shear stress resultant $\mathbf{Q}$. The effect, as will be seen, is to make the transverse shearing strain and the rotary inertia vanish, *without implying that* $\mathbf{Q}$ *or* $\boldsymbol{\omega}$ *do*.

A dynamic mixed-energy density may be defined via the partial *Legendre-Fenchel transformation*

$$\psi(\mathbf{L}, \boldsymbol{\omega}, \mathbf{n}, \mathbf{k}; \mathbf{y}) \equiv \inf_{\mathbf{R}, \mathbf{e}} \{ \varepsilon(\mathbf{L}, \mathbf{R}, \mathbf{e}, \mathbf{k}; \mathbf{y}) - \boldsymbol{\omega} \cdot \mathbf{R} - \mathbf{n} : \mathbf{e} \}. \tag{27}$$

If ε is differentiable and grows faster than $\mathbf{R}$ and $\mathbf{e}$ as these unknowns approach infinity (so that the infimum occurs at finite values of $\mathbf{R}$ and $\mathbf{e}$), then we have the *partially inverted constitutive relations*

$$\mathbf{v} = \psi_{,\mathbf{L}}, \quad \mathbf{R} = -\psi_{,\omega}, \quad \mathbf{e} = -\psi_{,\mathbf{n}}, \quad \mathbf{m} = \psi_{,\mathbf{k}}. \tag{28}$$

6 Choosing the Kinetic-Energy Density

From (2), (3), $(6)_1$, (7), and $(17)_1$

$$\mathbf{L} = m\mathbf{v}. \tag{29}$$

Hence, $(28)_1$ implies that the dynamic mixed-energy density has the form

$$\psi = \tfrac{1}{2} m^{-1} \mathbf{L} \cdot \mathbf{L} + \hat{\theta}(\boldsymbol{\omega}, \mathbf{n}, \mathbf{k}; \mathbf{y}). \tag{30}$$

What can be said of the $\boldsymbol{\omega}$-dependence of ψ? If we mimic rigid-body dynamics but at the same time recognize that a shell is deformable, then we might assume that

$$\mathbf{R} = \mathbf{l}(\mathbf{n}, \mathbf{k}; \mathbf{y}) \cdot \boldsymbol{\omega}, \tag{31}$$

where $\mathbf{l}$ is a rotary inertia (or moment of inertia) tensor. This would lead to a dynamic mixed-energy density of the form

$$\psi = \tfrac{1}{2} [m^{-1} \mathbf{L} \cdot \mathbf{L} + \boldsymbol{\omega} \cdot \mathbf{l}(\mathbf{n}, \mathbf{k}; \mathbf{y}) \cdot \boldsymbol{\omega}] + \tilde{\theta}(\mathbf{n}, \mathbf{k}; \mathbf{y})]. \tag{32}$$

The determination of $\mathbf{l}$, like the determination of $\tilde{\theta}$ must come from experiments or from some sort of descent from a three-dimensional energy density.

A simpler approach is to assume that

$$\mathbf{R} = I(\mathbf{n}, \mathbf{k}; \mathbf{y}) \boldsymbol{\omega}, \tag{33}$$

where I is a scalar rotatory inertia coefficient. Then,

$$\psi = \tfrac{1}{2} [m^{-1} \mathbf{L} \cdot \mathbf{L} + I(\mathbf{n}, \mathbf{k}; \mathbf{y}) \boldsymbol{\omega} \cdot \boldsymbol{\omega} + \theta(\mathbf{n}, \mathbf{k}; \mathbf{y})]. \tag{34}$$

This choice may be supported on two grounds: First, in practical problems involving not-too-thick shells, the rotational momentum $\mathbf{R}$ plays a minor role, as reflected by the magnitude of $\mathbf{l}$ or I which is $O(H^2)$, whereas $\boldsymbol{\omega}$ is an important, non-negligible kinematical ingredient. Thus, ignoring the contribution of $\mathbf{R}$ to the energy density ε would imply that $\boldsymbol{\omega} = \mathbf{0}$, i.e., the constitutive relation $(26)_2$ would be ill-conditioned, whereas ignoring the contribution of $\boldsymbol{\omega}$ in (34) has little consequence, *and does not imply that* $\boldsymbol{\omega} = \mathbf{0}$. And second, as I now show, there always exists some scalar I that minimizes the difference between K_{ext}, the exact kinetic energy per unit area of $\mathcal{R}$ of the shell, and $\frac{1}{2}(m^{-1}\mathbf{L} \cdot \mathbf{L} + I \boldsymbol{\omega} \cdot \boldsymbol{\omega})$.

By (2), (3), (7), and $(17)_1$,

$$K_{\text{ext}} \equiv \frac{1}{2} \int_{\mathcal{R}} \int_{-}^{+} \rho \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} \mu \, d\zeta \, d\mathbf{R} = \frac{1}{2} \left(m \mathbf{v} \cdot \mathbf{v} + \int_{\mathcal{R}} \int_{-}^{+} \rho \dot{\mathbf{z}} \cdot \dot{\mathbf{z}} \mu \, d\zeta \right). \tag{35}$$

Let

$$\mathbf{z} = \lambda(\mathbf{y}, \zeta, t) \mathbf{u}(\mathbf{y}, \zeta, t), \quad \text{where } |\mathbf{u}| = 1, \tag{36}$$

so that

$$\dot{\mathbf{z}} = \dot{\lambda} \mathbf{u} + \lambda \dot{\mathbf{u}}, \quad \mathbf{u} \cdot \dot{\mathbf{u}} = 0. \tag{37}$$

Then,

$$\int_{-}^{+} \rho \dot{\mathbf{z}} \cdot \dot{\mathbf{z}} \mu \, d\zeta = \int_{-}^{+} \rho (\dot{\lambda}^2 + \lambda^2 |\dot{\mathbf{u}}|^2) \mu \, d\zeta, \tag{38}$$

whereas, by $(6)_2$, (36), and (37),

$$\mathbf{R} = \int_{-}^{+} \rho \lambda^2 \mathbf{u} \times \dot{\mathbf{u}} \mu \, d\zeta. \tag{39}$$

The scalar and vector functions $\dot{\lambda}$ and $\lambda^2 \mathbf{u} \times \dot{\mathbf{u}}$ that appear in the integrands of (38) and (39) are independent in the sense that assigning of just one of these is not sufficient to determine $\mathbf{z}$ in (36). Thus, if $\dot{\lambda} = g(\mathbf{y}, \zeta, t)$, some given function, then

$$\lambda = \lambda(\mathbf{y}, \zeta, 0) + \int_0^t g(\mathbf{y}, \zeta, \tau) \, d\tau, \tag{40}$$

where the first function on the right may be assigned at will. With λ known, the relation $\lambda^2 \mathbf{u} \times \dot{\mathbf{u}} = \mathbf{g}(\mathbf{y}, \zeta, t)$, some given function, can be rewritten as

$$\dot{\mathbf{u}} = \boldsymbol{\Omega} \times \mathbf{u}, \quad \boldsymbol{\Omega} \equiv \lambda^{-2} \mathbf{g}. \tag{41}$$

At any $\mathbf{y}$ and ζ, this is just the equation for a unit vector rotating at a variable angular velocity $\boldsymbol{\Omega}$, so (41) obviously has a solution, starting at any arbitrary initial condition $\mathbf{u}(\mathbf{y}, \zeta, 0)$.

Thus, since $m^{-1}\mathbf{L} \cdot \mathbf{L} = m\mathbf{v} \cdot \mathbf{v}$, it follows from (35), (38), (39) that

$$\min_I\left\{K_{\text{ext}}-\frac{1}{2}(m\mathbf{v}\cdot\mathbf{v}+I\boldsymbol{\omega}\cdot\boldsymbol{\omega})\right\}=\frac{1}{2}\int_-^+\rho\dot{\lambda}^2\mu\,d\zeta>0, \tag{42}$$

providing the inertia coefficient is defined as

$$I\equiv\frac{\left|\int_-^+\rho\lambda^2\mathbf{u}\times\dot{\mathbf{u}}\mu\,d\zeta\right|^2}{\int_-^+\rho\lambda^2|\dot{\mathbf{u}}|^2\mu\,d\zeta}. \tag{43}$$

Of course, I cannot be computed from its definition (43) any more than, say, $\mathbf{L}$ can be computed from its definition $(6)_1$: a formula for $I(\mathbf{n},\mathbf{k};\mathbf{y})$ is part of the unavoidably approximate description of the dynamic mixed-energy density ψ. The simplest approximation is to set $I=\int_-^+\rho|\mathbf{z}|^2\mu\,d\zeta$, which is the initial value of the right side of (43) in a *Kirchhoff motion*, $\overline{\mathbf{z}}=\mathbf{Q}(\mathbf{y},t)\cdot\mathbf{z}(\mathbf{y},\zeta)$. Other, more elaborate, approximations may be constructed, especially for shells that undergo significant thickness changes.

7 The (Dynamic) Constitutive Kirchhoff Hypothesis

Perhaps the simplest way to describe this new form of the hypothesis is to introduce, in the tangent plane at each point of the reference surface $\mathcal{R}$, a set of independent vectors $\mathbf{y}_\alpha$, $\alpha=1,2$. These vectors need not be associated with a set of surface coordinates. The back-rotated stress resultant and bending strain tensors may then be given the component forms

$$\mathbf{n}=N^{\alpha\beta}\mathbf{y}_\beta\mathbf{y}_\alpha+Q^\alpha\mathbf{b}\mathbf{y}_\alpha \tag{44}$$

and

$$\mathbf{k}=K_{\alpha\beta}\mathbf{y}^\alpha\mathbf{y}^\beta\times\mathbf{b}+K_\alpha\mathbf{y}^\alpha\mathbf{b}, \tag{45}$$

where a repeated index is to be summed from 1 to 2 and $\mathbf{y}^\alpha\cdot\mathbf{y}_\beta=\delta^\alpha_\beta$, the Kronecker delta. The *(dynamic) constitutive Kirchhoff hypothesis* is that ψ in (32) depends only on $\mathbf{L}$,

$$N^{(\alpha\beta)}\equiv\tfrac{1}{2}(N^{\alpha\beta}+N^{\beta\alpha}),\quad K_{(\alpha\beta)}\equiv\tfrac{1}{2}(K_{\alpha\beta}+K_{\beta\alpha}), \tag{46}$$

and possibly $\mathbf{y}$.

If the back-rotated extensional-shear strain and stress couple tensors are put into the analogous component forms,

$$\mathbf{e}=E_{\alpha\beta}\mathbf{y}^\alpha\mathbf{y}^\beta+E_\alpha\mathbf{b}\mathbf{y}^\alpha \tag{47}$$

and

$$\mathbf{m}=M^{\alpha\beta}\mathbf{y}_\beta\times\mathbf{y}_\alpha+M^\alpha\mathbf{b}\mathbf{y}_\alpha, \tag{48}$$

then the constitutive relations $(28)_{2,3,4}$ imply that $\mathbf{R}=\mathbf{0}$ and

$$E_{(\alpha\beta)}=-\psi,_N{}^{(\alpha\beta)},\quad E_\alpha=0,\quad M^{(\alpha\beta)}=\psi,_{K_{(\alpha\beta)}},\quad M^\alpha=0. \tag{49}$$

In coordinate-free form, the constitutive dynamic Kirchhoff hypothesis is that ψ depends only on $\mathbf{L}$, $\overline{\mathbf{n}}$, $\overline{\mathbf{k}}$, and possibly $\mathbf{y}$, where

$$\overline{\mathbf{n}}\equiv\tfrac{1}{2}(\boldsymbol{\nabla}\mathbf{y}\cdot\mathbf{n}+\mathbf{n}^T\cdot\boldsymbol{\nabla}\mathbf{y}) \tag{50}$$

and

$$\overline{\mathbf{k}}\equiv\tfrac{1}{2}(\mathbf{k}\cdot\boldsymbol{\nabla}\mathbf{y}+\boldsymbol{\nabla}\mathbf{y}\cdot\mathbf{k}^T). \tag{51}$$

(Recall that $\boldsymbol{\nabla}\mathbf{y}=\boldsymbol{\nabla}\mathbf{y}^T$.)

8 Conclusions

I have presented a general nonlinear theory of isothermal, elastodynamic shells which makes assumptions only in the constitutive relations, which are intrinsically approximate. (They rest ultimately on experiments.) In particular, in the equations of motion of a shell, the inertia terms—which, of course, involve material properties—are regarded as given by constitutive relations, following a suggestion made to me independently by Ladevèze and Makowski. Problems of ill-conditioning that can arise in near inextensional bending are avoided by the introduction of a dynamic mixed-energy density. This last maneuver allows an alternative form of the Kirchhoff hypothesis to be introduced which avoids any a priori kinematic constraints imposed by the classic form of the hypothesis.

References

[1] Libai, A., and Simmonds, J. G., 1998, *The Nonlinear Theory of Elastic Shells*, 2nd ed., Cambridge University Press, Cambridge, UK.

[2] Habip, L. M., and Ebcioglu, I. K., 1965, ''On the Equations of Motion of Shells in the Reference State,'' Ing.-Arch., **34**, pp. 28–32.

[3] Habip, L. M., 1965, ''Theory of Elastic Shells in the Reference State,'' Ing-Arch., **34**, pp. 228–237.

[4] Naghdi, P. M., 1984, ''The Theory of Shells and Plates,'' *Mechanics of Solids*, Vol. II, C. Truesdell, ed., Springer-Verlag, Berlin, pp. 425–640.

[5] Antman, S. S., 1995, *Nonlinear Problems of Elasticity*, Springer-Verlag, New York.

[6] Koiter, W. T., and Simmonds, J. G., 1973, ''Foundations of Shell Theory,'' *Proc. 13th International Congress of Theoretical and Applied Mechanics*,'' E. Becker and G. K. Mikhailov, eds., Springer-Verlag, Berlin, pp. 150–176.

Stability of the Shanley Column Under Cyclic Loading

E. Corona
Department of Aerospace and Mechanical Engineering,
University of Notre Dame,
Notre Dame, IN 46556

This paper presents a numerical study of the Shanley column under cyclic loading. The model includes intermediate deformation kinematics. The constitutive model is based on the Dafalias-Popov model. Results for fully reversed, symmetric load or displacement-controlled loading are presented. Under displacement-controlled loading, the model can exhibit a transient response which takes it away from its initial configuration, but the response eventually reaches a stable cycle. Under load-controlled loading, the model can either reach a limit cycle or develop a limit load instability that causes collapse. The responses that result in collapse are imperfection sensitive. In addition, the response of the model is also sensitive to the amplitude of the applied loading cycles.
[DOI: 10.1115/1.1349118]

Introduction

Many metal structures currently in service are subjected to cyclic loading. Examples include structures located in seismically active areas, offshore structures excited by wind and waves, structures subjected to cyclic thermal loading, etc. In some cases, mostly under off-design conditions, the amplitude of the cyclic loading may cause repeated yielding of the material. It is well known that, in those cases, structural degradation is possible and can lead to buckling and collapse. Bertero and Popov ([1]) presented one of the earliest demonstrations of such behavior in cyclically loaded cantilever beams. Later studies of the response and stability of structural members under cyclic loading ([2–7]) have shown that structural degradation and collapse under cyclic loads affect a wide range of structural members. In many of these cases the cause of degradation is the progressive growth of the buckling mode, which eventually localizes and leads to collapse. Most of the studies conducted have been experimental. Attempts to conduct analytical studies to capture experimentally measured behavior ([8,9]) have been successful in some cases, but only after overcoming significant challenges in the area of constitutive modeling. The main difficulty has been the accurate prediction of ratcheting under general, multiaxial loading conditions that are present in many of the structural members studied, as shown in ([10–12]). In view of this state of affairs, one avenue to analytically explore the stability of structural members subjected to cyclic loading is to consider the response of simplified models.

The objective of the present work is to study the response of a relatively simple structural model subjected to cyclic load or displacement controlled loading in order to study the degradation and collapse mechanism resulting from the growth in amplitude of the buckling mode. The model considered was first proposed by Shanley ([13]) to study column plastic buckling and is shown in Fig. 1(*a*). All members are rigid, except for the elastic-plastic springs, or links, that support the model at points B and C. The state of stress in the links is uniaxial. Due to the importance of ratcheting in the calculated response, the cyclic plasticity model employed has to have shown success in predicting ratcheting under uniaxial stress. Such a model will be discussed in more detail in the next section. The main concern of the investigation is the stability of the model under fully reversed loading, and how it is related to the geometric and loading parameters imposed. This type of loading has been employed in several experimental studies of members under bending and hence is employed here. The investigation consists of numerical experiments on the model for various geometries and loading conditions and analysis of the results.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 20, 2000; final revision, June 28, 2000. Associate Editor: S. Kyriakides. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Formulation

The Shanley column, shown in Fig. 1(*a*), consists of two rigid members that deflect in the plane. The vertical member OA has length L and connects to the horizontal member BC of length b at point O. OA and BC remain perpendicular at all times. Elastic-plastic links of length h and unit cross-sectional area support the structure at points B and C.

Figure 1(*b*) shows the model in its deflected configuration. The point of application of the external load N is at A. The line of action of N remains vertical at all times. The model has two degrees-of-freedom: the vertical displacement of point O (u) and the rotation (θ). The initial configuration can be imperfect. The imperfection consists of initial values of u and θ given by u_o and θ_o, respectively. For the range of parameters considered in this investigation, u and θ remain relatively small, so an intermediate class of kinematics is sufficient to capture the behavior of the model.

The equilibrium equations of the model can be easily obtained from the free-body diagram in Fig. 1(*b*). They are

$$F_1 + F_2 = N$$

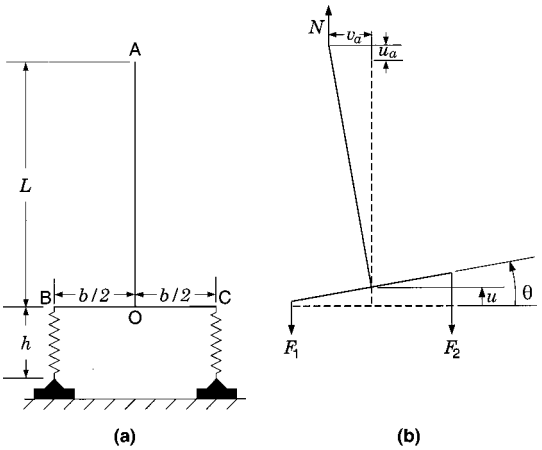


Fig. 1 (*a*) Shanley column, (*b*) deflected configuration and free-body diagram

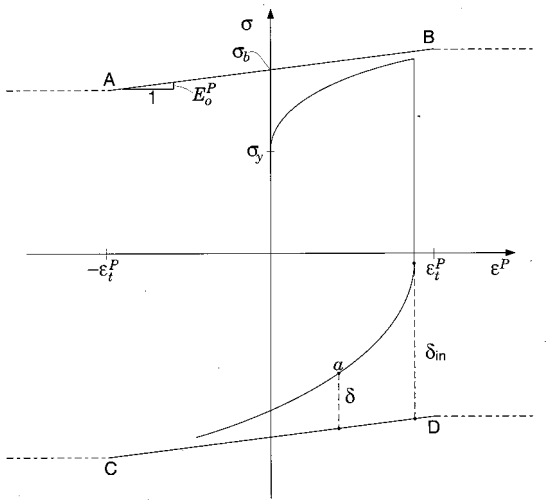


Fig. 2 Plasticity model parameters

$$F_1 - F_2 = 2N\frac{L}{b}\theta. \tag{1}$$

The strain in each link is given by

$$\epsilon_1 = \frac{\overline{h}_1 - \overline{h}_{1o}}{\overline{h}}, \quad \epsilon_2 = \frac{\overline{h}_2 - \overline{h}_{2o}}{\overline{h}}, \tag{2}$$

where the bar indicates normalization with respect to L ($\overline{h} = h/L$, etc.). Here,

$$\overline{h}_1 = \overline{h} + \overline{u} - \frac{b}{2L}\theta, \quad \overline{h}_2 = \overline{h} + \overline{u} + \frac{b}{2L}\theta, \tag{3}$$

and $\overline{h}_{1o}$ and $\overline{h}_{2o}$ are the values of $\overline{h}_1$ and $\overline{h}_2$ due to the initial imperfection. The horizontal and vertical deflections of point A are given by

$$\overline{u}_a = \overline{u} - \theta^2/2, \quad \overline{v}_a = \theta, \tag{4}$$

respectively. It is clear that the model contains two geometric parameters, $\overline{h}$ and L/b.

A second nonlinearity in the model is due to the constitutive relation of the links. It desirable to employ a plasticity model that has shown success in predicting uniaxial ratcheting in actual materials. The Dafalias-Popov cyclic plasticity model ([14,15]) governs the constitutive behavior of the links. This model is based on classical incremental plasticity with kinematic hardening. In the uniaxial setting, as applicable in this work, the basic characteristics of the model can be explained with the help of the stress-plastic strain diagram in Fig. 2. As is customary, the strain increment is decomposed into an elastic and a plastic part

$$d\epsilon = d\epsilon^e + d\epsilon^P \tag{5}$$

where

$$d\epsilon^e = \frac{d\sigma}{E} \tag{6}$$

and E is Young's Modulus.

The stress-plastic strain response of the material is bounded between the two lines AB and CD, which are called the ''bounds.'' In the simplest case, the bounds are linear with slope E_o^P. The relation between the plastic strain increment and the stress increment is given by

$$d\epsilon^P = \frac{d\sigma}{E_t^P} \tag{7}$$

where E_t^P is the plastic tangent modulus. At a given point, say point a in the figure, E_t^P is calculated based on the distances δ between a and the bound and δ_{in} between the previous yield point and the same bound. The plastic tangent modulus is given by

$$E_t^P = E_o^P + \eta\left(\frac{\delta}{\delta_{in} - \delta}\right) \tag{8}$$

where η controls the shape of the curve and is given by

$$\eta = \frac{\alpha}{1 + \beta\left(\frac{\delta_{in}}{2\sigma_b}\right)^m}. \tag{9}$$

If the stress-strain curve touches a bound, then $E_t^P = E_o^P$. All parameters of the model can be evaluated from two uniaxial stress-strain curves.

Hassan and Kyriakides [10] introduced a modification to the model by cutting the bounds at the points A, B, C, and D instead of letting them extend to infinity as in the original model. If the absolute value of the plastic strain exceeds the value ϵ_t^P, the points A, B, C, and D translate along the strain axis in the direction of the strain increment at the same rate as the plastic strain. Using this modification, they showed that the model could accurately predict experimentally measured uniaxial ratcheting.

A simple plastic bifurcation analysis of the perfect system yields the bifurcation buckling load

$$N_c = -\frac{E_t}{2}\frac{(b/L)^2}{(h/L)} \tag{10}$$

where E_t is the instantaneous tangent modulus given by

$$\frac{1}{E_t} = \frac{1}{E} + \frac{1}{E_t^P}. \tag{11}$$

The negative sign indicates that the buckling load is compressive.

If E_t is not constant, (10) must be solved by trial and error. The corresponding buckling mode is $\hat{u} = 0$ and $\hat{\theta}$ arbitrary.

The response of the column under load control can be obtained by prescribing the external load N and then solving Eqs. (1)–(3) using Newton's method. Alternatively, the response under displacement control can be obtained by prescribing the displacement u_a. In this case, the first of Eq. (4) must also be included in the solution procedure.

It is understood that this simple, two-degree-of-freedom model cannot account for all aspects of the measured response of actual structural members, such as localization-driven collapse, collapse under displacement-controlled loading, and the degradation due to prebuckling deformations observed in circular tubes under bending ([2,5,8]). It provides, however, a simple way to simulate the growth of the amplitude of buckling modes under cyclic loading as has been observed, for example, in tubes of square cross section ([6]) and T-beams ([7]) under cyclic bending.

Results[1]

The investigation of the behavior of the model was carried out as a series of numerical experiments by varying the geometric and loading parameters. The material properties are those of carbon steel 1026 used in ([10]) and are listed in Table 1. These properties remained fixed in all cases presented here. Note that the yield load of the perfect model is $N_y = 2\sigma_y$. The corresponding displacement is $\overline{u}_y = \sigma_y \overline{h}/E$.

The results can be divided into three cases depending on the loading conditions: monotonic response, cyclic response under displacement control, and cyclic response under load control. In all cases considered $\overline{h} = 0.01$, so the only geometric parameter

[1]Preliminary results of the response of the model were presented in [16].

Table 1 Material properties

E, GPa (ksi)	E_o^P, GPa (ksi)	σ_y, MPa (ksi)	σ_b, MPa (ksi)	α, GPa (ksi)	β	m	ϵ_t^P, %
181 (26,320)	1.38 (200)	131 (19)	269 (39)	490 (71,100)	27	2	2.35

varied was L/b. Three values of L/b (50, 100, and 150) were chosen to illustrate the various aspects of the response of the model.

Monotonic Response

Shanley developed his model to analytically demonstrate that the critical load calculated based on the tangent modulus represents the lowest load at which lateral deflection of a perfect, elastic-plastic column becomes possible. Furthermore, he showed that the load supported by the column increases through the critical point as the response follows the stable branch.

Table 2 lists the bifurcation loads and the corresponding axial displacements of the three models considered, as calculated using Eq. (10). Both quantities are normalized by the corresponding yield quantities. It is clear that as L/b decreases, bifurcation occurs further into the plastic range of the material.

The post-buckling behavior of the model can be studied by introducing a small initial imperfection in the form of an initial rotation θ_o. Figure 3(*a*) shows the load-axial deflection ($N-u_a$) responses calculated for the three values of L/b considered with an initial deflection of $\theta_o=10^{-4}$ rad. The loads are normalized by the absolute value of the bifurcation loads $|N_{cr}|$ while the deflections u_a are normalized by the absolute value of the bifurcation deflections $|u_{cr}|$. In all cases the response is characterized by a limit load instability. The limit loads occur at values of $N/|N_{cr}|=0.978$, 0.957, and 0.928 for $L/b=50$, 100, and 150, respectively. Note that as L/b decreases the N-u_a curves exhibit more softening prior to reaching the limit load. This is due to the increase in N_{cr}/N_y as shown in Table 2. In other words, columns with lower L/b have a more severe plastic state at the limit load. Figure 3(*b*) shows the corresponding load-rotation curves. Note that the values of θ at the limit load are all below 0.001 rad so that the assumption of intermediate rotations is justified. Further analysis indicated that, as expected, the limit loads and the corresponding displacements are moderately imperfection sensitive.

Table 2 Brifurcation loads and displacements for three values of L/b considered

L/b	N_{cr}/N_y	u_a/u_y
50	−1.93	−3.16
100	−1.58	−1.84
150	−1.22	−1.25

Cyclic Loading: Displacement Control

Under displacement control, the prescribed variable is the vertical deflection of the tip of the column, u_a in Fig. 1. The first case to be discussed is that of a model with $L/b=50$ under symmetric loading about $u_a=0$ and with amplitude $\tilde{u}_a/|u_{cr}|=0.95$ ($\tilde{u}_a$ indicates the amplitude of the cycle in u_a). Figure 4(*a*) shows the $N-u_a$ response. The initial imperfection is $\theta_o=10^{-4}$ rad. This imperfection will be the default value for the rest of the paper, unless stated otherwise. The loading cycles start by compressing the column. Note that, initially, the response goes through a transient stage where the loops are shifting but eventually approaches a repeatable, or limit, cycle. The reason for the shifting can be seen in Fig. 4(*b*), which shows the N-θ response. It is clear that the model progressively leans so that the peak value of θ in each cycle increases. This causes softening in compression. Hardening in tension develops from the increase in the width of the loops.

This type of displacement controlled loading imposes a kinematic constraint on the system which keeps the lateral deflection from growing too large. As a result, it can be expected that the system will be stable in the long term even though the transient response takes the system away from its unloaded configuration. The mechanics of how stability is achieved at the material level can be explained by looking at the stress-strain (σ-ε) response of the two elastic-plastic links in Figs. 5(*a*) and (*b*). As the lateral deflection increases, the model leans on link 1, hence it ratchets in the negative strain direction. Link 2 ratchets in the positive strain direction. Note that the rate of ratcheting decreases with cycling in both links. Ratcheting can be slowed down in two manners: by a decrease of the mean stress or by a decrease in the amplitude of the stress cycles. It is clear that in link 1 the mean stress decreases (in absolute value) during the transient period, while in spring 2 the mean stress increases significantly, but the amplitude of the cycles decreases very quickly simultaneously. The net effect is that ratcheting stops after sufficient cycling. As a final detail, careful consideration of the stress-strain loops reveals that the mean stress in link 1 does not decrease all the way to zero when ratcheting stops. This arrest of ratcheting with nonzero mean stress is a characteristic of the Dafalias-Popov model. Should a different plasticity model which requires zero mean stress to stop ratcheting

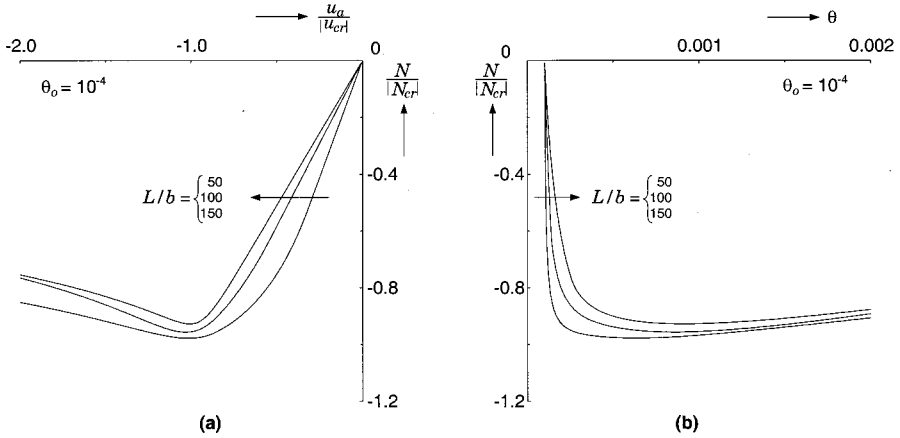


Fig. 3 Monotonic response, (*a*) load-axial deflection response of three models, (*b*) corresponding load-rotation response

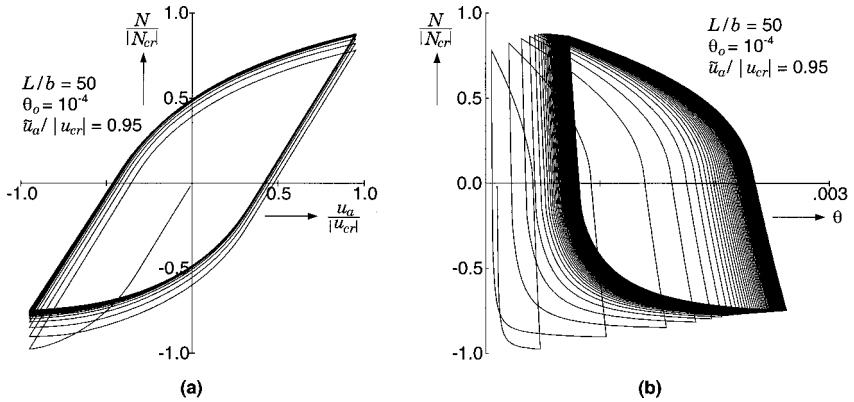


Fig. 4 Cyclic response under symmetric displacement control for a model with $L/b=50$, (*a*) load-axial deflection response, (*b*) corresponding load-rotation response

be used, the mean stress in link 1 would have to decrease to zero. Similarly, link 2 stops ratcheting even though some plastic deformation occurs in each cycle and the mean stress is relatively high.

The effect of varying the amplitude of the loading cycle can be seen in the plot in Fig. 6. It shows the peak value of θ (θ_p) in each cycle as a function of the number of cycles. As expected, the response eventually stabilizes in all cases. The stable value of θ_p and the rate at which it is approached depend on the amplitude of the cycle. The higher the amplitude, the larger the stable value and the faster it is approached. The imperfection sensitivity of the response of the model was also studied by considering $\theta_o = 10^{-5}$, 10^{-4}, and 10^{-3}. The results indicate that, in this case, the stable value of θ_p is relatively insensitive to the amplitude of the initial imperfection.

Figures 7(*a*) and (*b*) show the load-axial deflection and load-rotation plots for a model with $L/b=100$ loaded at $\tilde{u}_a/|u_{cr}| = 0.95$. As in the previous case, the response approaches a limit cycle. The load-axial deflection plot looks very similar to the one in the previous case. Note that limit loads occur in the first few several cycles, but they are inconsequential under displacement control. The load-rotation plot shows significant differences when compared to that in Fig. 4, the most obvious being that the sign of θ in all loops is opposite to that of θ_o. This reversal in θ occurs frequently in models with $L/b=100$ and 150. The reason for the reversal can be seen in Fig. 8, which shows the stress-strain histories of both links for the first cycle. At the first reversal, indicated by ①, link 1 has reached a higher value of stress and strain than link 2, as expected. Upon unloading and reverse loading, link 1 yields before link 2 due to the Bauschinger effect and softens significantly more during the rest of this cycle segment. This causes link 1 to develop a larger strain at point ②. The strain difference between the links is sufficient to overcome the initial imperfection and accommodate the change of sign of θ. In the case shown in Fig. 4 for $L/b=50$ the stress-strain history is similar to this point, but the difference between the strains in the links is not sufficient to overcome the initial imperfection and the sign of θ does not change. During reloading from ② to ③, link 2 develops much larger strain than link 1 and hence pushes the model further in the negative θ-direction. From this point on, θ remains negative as shown in Fig. 7(*b*).

Further investigation into cases with $L/b=50$, 100, and 150, imperfections of 10^{-5}, 10^{-4} and 10^{-3}, and various cycle amplitudes indicated that the transient responses of the models are varied. The long term behavior is stable, but the long term values of θ_p depend on the parameters listed above. In general, the behavior of the model becomes more predictable for lower values of L/b, larger initial imperfections and larger loading amplitudes. In these cases, the plastic state in the links is more severe, which more effectively ''locks'' the deflection of the model in a more predictable direction.

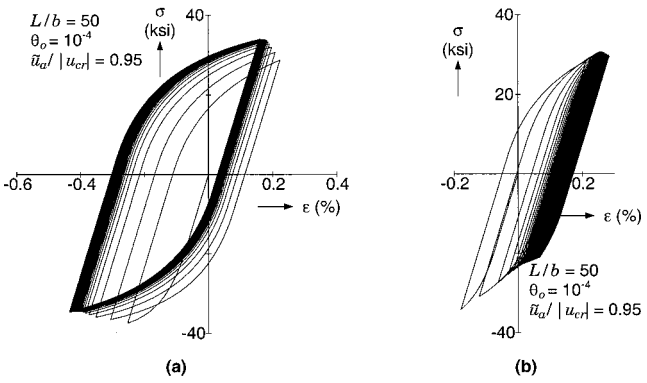


Fig. 5 Stress-strain response for the case in Fig. 4, (*a*) link 1, (*b*) link 2

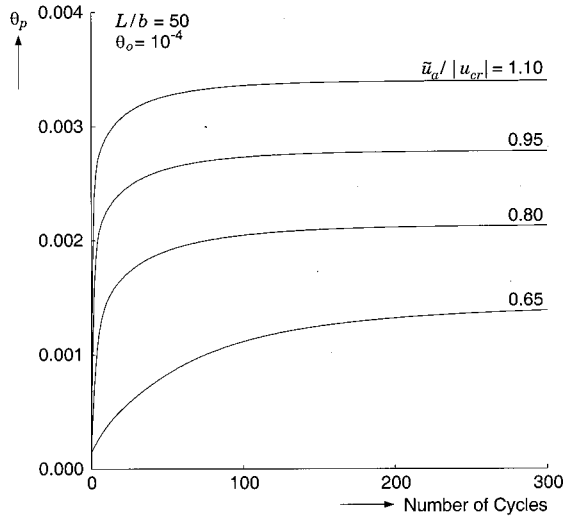


Fig. 6 Peak values of θ in each cycle as a function of cycle number for various cycle amplitudes under displacement control

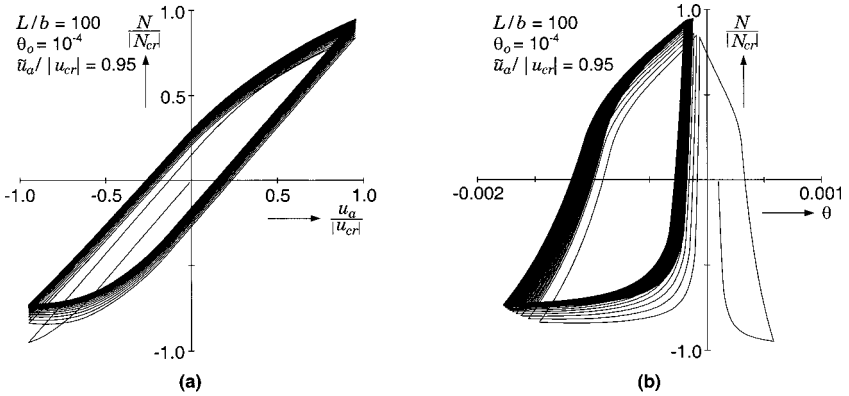


Fig. 7 Cyclic response under symmetric displacement control for a model with $L/b=100$, (*a*) load-axial deflection response, (*b*) corresponding load-rotation response

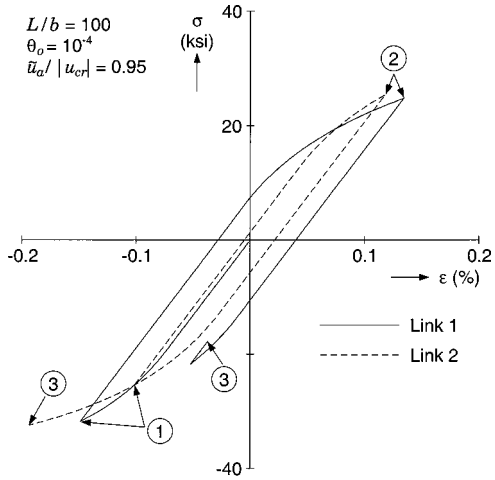


Fig. 8 Stress-strain histories for the first cycle of the response shown in Fig. 7

Cyclic Loading: Load Control

The second mode of cyclic loading considered is load control, where N is the prescribed variable. As in the case of displacement control, the cycles considered are symmetric about $N=0$. The amplitude of the cycles, denoted by $\tilde{N}$, necessarily has to be smaller than the limit loads presented in Fig. 3, in contrast with displacement control where the amplitude of the cycle could be larger than the displacement corresponding to the limit load under monotonic loading. In parallel with the discussion of displacement controlled loading, the first case considered is one where significant plastic deformation takes place.

Figure 9 shows the response of a model with $L/b=50$ with a load amplitude $\tilde{N}/|N_{cr}|=0.8$. The N-u response, shown in Fig. 9(*a*) is significantly different from the one obtained under displacement control. In this case the loops change slowly at the start but progressively ratchet in the compressive direction. The ratcheting accelerates as cycling progresses and becomes relatively fast for the last few cycles. The $N-\theta$ response, shown in Fig. 9(*b*), shows that ratcheting in θ is also present. In fact, a small amount of the deflection u_a is due to the leaning of the model. Ratcheting in θ also starts at a slow rate and accelerates throughout the loading history, moving away from θ_o. Also, note that the instantaneous stiffness of the model at the compressive load reversal point decreases with cycling. The calculations stop when the solution procedure diverges as the model approaches a limit load instability during the last cycle. At this point the model can no longer support the load required by the loading history, and it collapses.

The model exhibits two possible instabilities. The first, which is also present in the displacement controlled case in Fig. 4 is the departure of the configuration of the model away from the initial one. The speed of departure, however, is significantly different. Under displacement control, the departure is initially fast and progressively slows down until it stops. Under load control, the de-

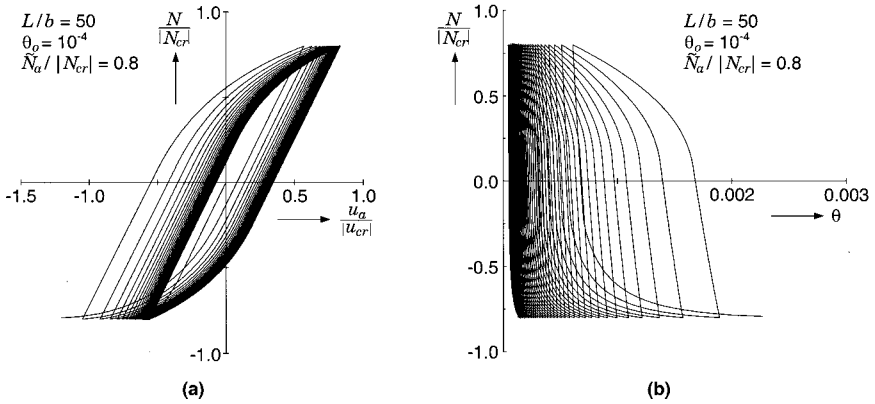


Fig. 9 Cyclic response under symmetric load control for a model with $L/b=50$, (*a*) load-axial deflection response, (*b*) corresponding load-rotation response

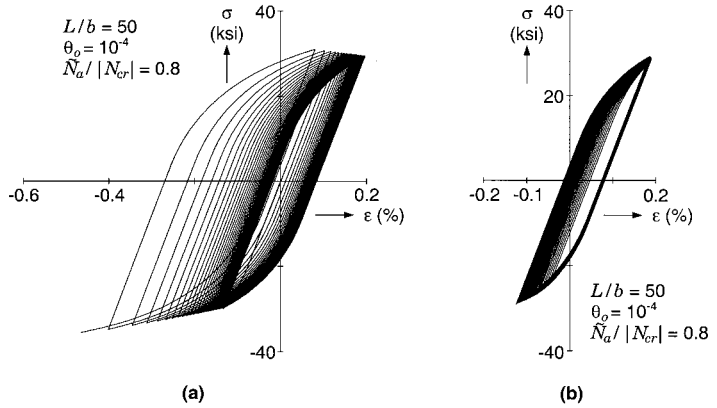


Fig. 10 Stress-strain response under symmetric load control for the case in Fig. 9, (*a*) link 1, (b) link 2

parture is initially slow and progressively accelerates. The second instability is due to the fact that, under load control, the kinematic constraint present in the displacement control case does not exist. Therefore, both u and θ may attain relatively large values leading to the collapse of the model.

Figure 10 shows the σ-ε response of the two links for the case shown in Fig. 9. As cycling progresses, the model leans on link 1 (Fig. 10(*a*)) and hence it ratchets in the compressive direction. Note that both the stress amplitude and the absolute value of the mean stress increase with cycling. These increases have the effect

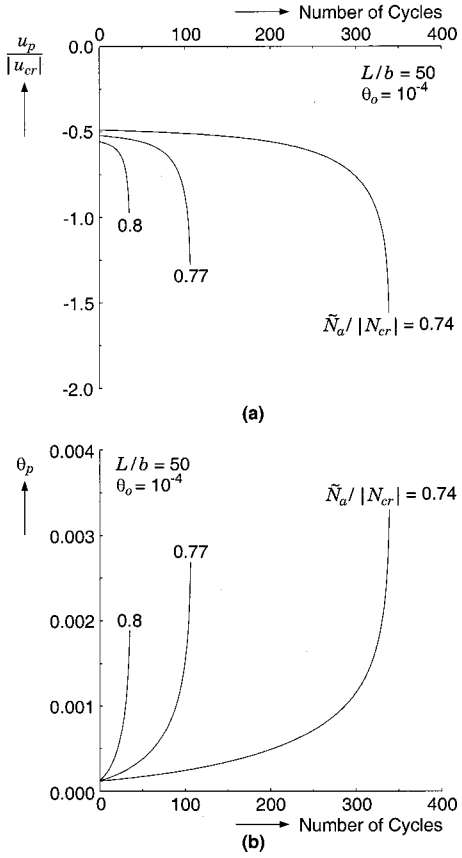


Fig. 11 Response history under load control for a model with $L/b=50$ for different load amplitudes, (*a*) peak values of vertical displacement in each cycle as function of cycle number, (*b*) peak values of θ in each cycle as function of cycle number

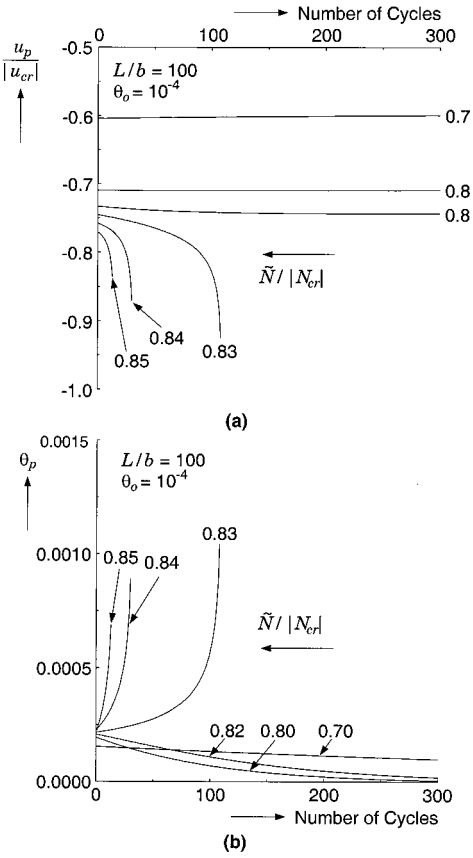


Fig. 12 Response history under load control for a model with $L/b=100$ for different load amplitudes, (*a*) peak values of vertical displacement in each cycle as function of cycle number, (*b*) peak values of θ in each cycle as function of cycle number

of accelerating ratcheting in this link. The response of link 2, shown in Fig. 10(*b*), exhibits a more moderate response within a smaller strain range.

The peak values of u_a and θ in each cycle, shown in Fig. 11 as a function of number of cycles, indicate that as $\tilde{N}$ decreases the initial rates of accumulation of u_a and θ decrease and that the number of cycles required to induce the limit load instability increases. Note that the range of $\tilde{N}$ which could lead to instability is bounded by the limit load above and approximately by the yield load of the perfect model below. In this case, $0.52 \leqslant \tilde{N} \leqslant 0.978$. Therefore, the amplitude range presented in the figure is rather small, yet the number of cycles required to induce the limit load span one order of magnitude. This indicates that cycle count at the limit load instability is very sensitive to the amplitude of the load cycle. Indeed, for $\tilde{N}/|N_{cr}|=0.65$ the limit load occurs at 2.7×10^4 cycles.

Figure 12 shows peak values of u_a and θ in each cycle for a model with $L/b=100$. The cases considered here show a remarkable difference when compared with the cases in Fig. 11. The three cases with higher $\tilde{N}/|N_{cr}|$ develop a limit load instability as discussed previously. The three cases with lower $\tilde{N}/|N_{cr}|$ do not. The difference can be attributed to the material nonlinearity associated with the Bauschinger effect along the lines of the discussion of the case in Fig. 8. For the cases with lower $\tilde{N}/|N_{cr}|$, the responses tend to limit cycles. This was verified by carrying the calculations to a much higher number of cycles than shown in the figure. For example, Fig. 13 shows plots of the peak values of u_a and θ in each half-cycle for a range of 4000 cycles when $\tilde{N}/|N_{cr}|=0.7$. It is clear that while the range of u_a in each cycle

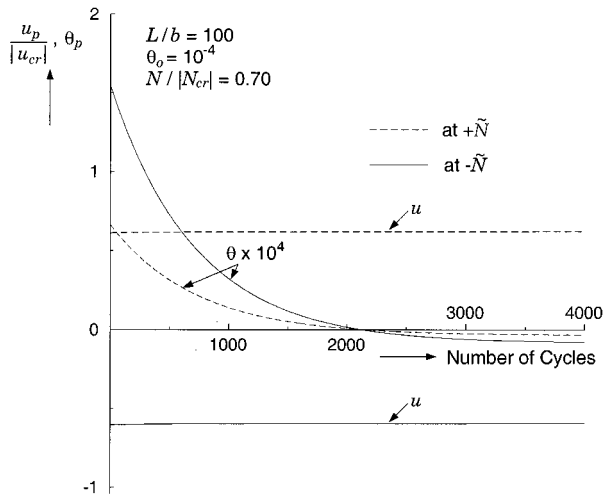


Fig. 13 Response in terms of peak values of u and θ in each half-cycle for a model with $L/b=100$ and $\tilde{N}/|N_{cr}|=0.70$ under load control

remains nearly constant, the range of θ decreases continuously until the 2110th cycle. At this point the loop flips and, by the 2150th cycle, all values of θ in the cycle are negative. The range of θ subsequently increases but now the cycles approach a stationary condition.

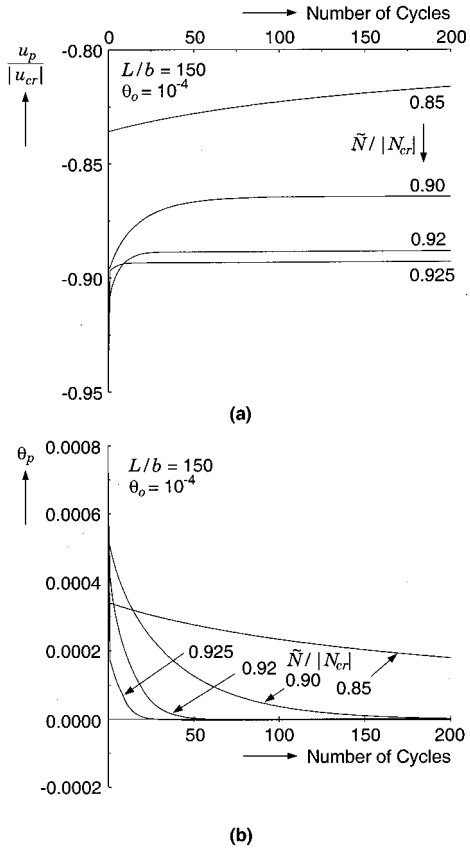


Fig. 14 Response history under load control for a model with $L/b=150$ for different load amplitudes, (*a*) peak values of vertical displacement in each cycle as function of cycle number, (*b*) peak values of θ in each cycle as function of cycle number

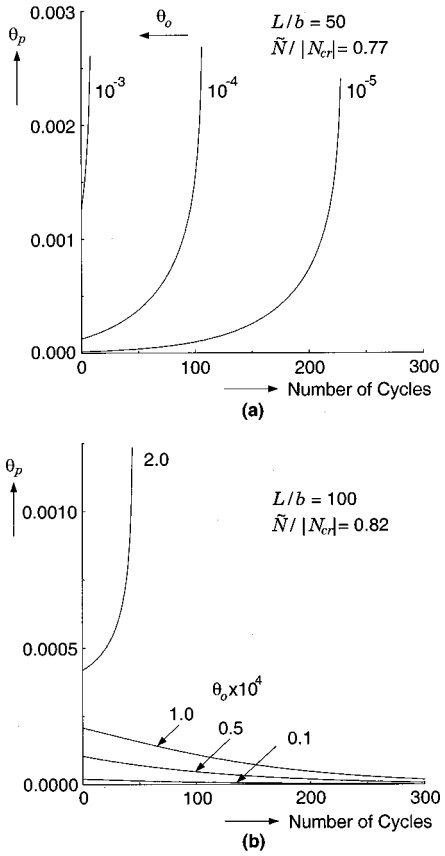


Fig. 15 Response history in terms of peak θ values as function of cycle number under load control for different values of initial imperfection, (*a*) $L/b=50$, (*b*) $L/b=100$

Figure 14 shows peak values of u_a and θ in each cycle for a model with $L/b=150$. The range of possible values of $\tilde{N}$ is most restricted in this case because, for the perfect model, yielding occurs only for $\tilde{N}/|N_{cr}|>0.82$. The plots in the figures show that the responses in all cases considered did not march toward limit load instabilities. Therefore, the responses were stable in the long term.

Imperfection sensitivity is another important aspect of the response of the model when under load control. Figure 15(*a*) shows plots of θ_p versus number of cycles for three values of θ_o one order of magnitude apart. In these cases $L/b=50$ and $\tilde{N}/|N_{cr}|=0.77$. It is clear that the rate of accumulation of θ_p and number of cycles to collapse are sensitive to the initial imperfection. Further analysis indicates that the number of cycles to collapse continues to increase as the imperfection amplitude decreases. This trend continued down to amplitudes of 10^{-18}. The number of cycles to collapse became a random variable for smaller imperfections but never exceeded 2000 cycles. This effect is most likely due to numerical effects that surface when the imperfection is very small. It, however, may indicate that for given material, geometric and loading conditions, a maximum possible number of cycles to collapse exists due to unavoidable perturbations.

Figure 15(*b*) shows similar plots for a model with $L/b=100$ and $\tilde{N}/|N_{cr}|=0.82$. The value of $\tilde{N}$ was chosen to be close to the boundary between long-term stable and unstable responses in Fig. 12. The results indicate that the initial imperfection can also have an effect deciding whether cycling will drive the model to a limit cycle or to collapse.

Conclusions

This paper discussed the elastic-plastic stability of the Shanley column model when subjected to fully reversed cyclic loading in either displacement or load control. The model can exhibit two instabilities: the first is a progressive departure from the initial configuration of the model. It can occur under load and displacement control. Under displacement control the model approaches a limit cycle, whereas under load control the departure from the initial state may accelerate in some cases. The second is a limit load instability that under load control results in the collapse of the model. The results indicate, however, that as L/b increases and $\tilde{N}$ decreases, the model may reach a stable cycle even under load control. Another aspect of the results worth emphasizing is that, in addition to imperfection sensitivity, the model exhibits load amplitude sensitivity. For example, under load control, the number of cycles to collapse strongly depends on the amplitude of the cycle.

It is clear from the results that the response of even such a simple model is complex and rich in features. The nature of the response can be affected by the parameters varied in this study, namely the amplitude of the loading cycle, L/b and the amplitude of the initial imperfection. Other parameters kept constant in this investigation that can be of significant importance in actual structures and that can affect the response and stability of the model include the presence of mean values of displacement or load in the loading cycles and variations in the material properties.

The number of variables that can affect the stability of actual structures can obviously be much larger than in the present model. Furthermore, as the complexity of the structure increases, it is possible to have more than one load, multiaxial states of stress, different stress histories at different points, possible localization instabilities, collapse under displacement control, etc. These factors increase the complexity of the response well above what was discussed in this paper and, coupled with current difficulties in ratcheting prediction, make quantitative predictions of the development of structural instabilities under cyclic loading very difficult to obtain. Therefore, experimentation remains an essential tool that must accompany analytical studies of the response and stability of actual structures or structural elements under cyclic loading. The results from the simple model presented here demonstrate that the combination of geometric nonlinearities and ratcheting at the material level can lead to the growth of buckling modes and induce collapse under cyclic loading as has been observed experimentally.

Acknowledgments

This work was carried out with the financial support of the National Science Foundation under Grant No. CMS96-10510.

References

[1] Bertero, V. V., and Popov, E. P., 1965, ''Effect of Large Alternating Strains of Steel Beams,'' J. Struct. Div. ASCE, **91**, pp. 1–12.
[2] Kyriakides, S., and Shaw, P. K., 1987, ''Inelastic Buckling of Tubes Under Cyclic Bending,'' ASME J. Pressure Vessel Technol., **109**, pp. 169–178.
[3] Clément, G., Acker, D., and Lebey, J., 1989, ''A Practical Rule for Progressive Buckling,'' Nucl. Eng. Des., **111**, pp. 209–216.
[4] Taylor, N., and Hirst, P., 1989, ''Strut Behavior: Sub-Buckling Excursion Effects and Risk Assessment,'' Res. Mech., **28**, pp. 139–189.
[5] Corona, E., and Kyriakides, S., 1991, ''An Experimental Investigation of the Degradation and Buckling of Circular Tubes Under Cyclic Bending and External Pressure,'' Thin-Walled Struct., **12**, pp. 229–263.
[6] Vaze, S. P., and Corona, E., 1998, ''Degradation and Collapse of Square Tubes Under Cyclic Bending,'' Thin-Walled Struct., **31**, pp. 325–341.
[7] Ellison, M. S., and Corona, E., 1998, ''Buckling of T-Beams Under Cyclic Bending,'' Int. J. Mech. Sci., **40**, pp. 835–855.
[8] Shaw, P. K., and Kyriakides, S., 1985, ''Inelastic Analysis of Thin-Walled Tubes Under Cyclic Bending,'' Int. J. Solids Struct., **21**, pp. 1073–1100.
[9] Corona, E., 1997, ''Cyclic Plasticity Applications in Structural Degradation of Shells Under Cyclic Loading,'' *Proceedings of Plasticity '97*, A. S. Khan, ed., Neat Press, Fulton, MD, pp. 345–346.
[10] Hassan, T., and Kyriakides, S., 1992, ''Ratcheting in Cyclic Plasticity, Part I: Uniaxial Behavior,'' Int. J. Plast. **8**, pp. 91–116.
[11] Hassan, T., Corona, E., and Kyriakides, S., 1992, ''Ratcheting in Cyclic Plasticity, Part II: Multiaxial Behavior,'' Int. J. Plast. **8**, pp. 117–146.
[12] Corona, E., Hassan, T., and Kyriakides, S., 1996, ''On the Performance of Kinematic Hardening Rules in Predicting a Class of Biaxial Ratcheting Histories,'' Int. J. Plast. **12**, pp. 117–145.
[13] Shanley, F. R., 1947, ''Inelastic Column Theory,'' J. Aeronaut. Sci., **14**, pp. 261–268.
[14] Dafalias, Y. F., and Popov, E. P., 1975, ''A Model of Nonlinearly Hardening Materials for Complex Loading,'' Acta Mech., **21**, pp. 173–192.
[15] Dafalias, Y. F., and Popov, E. P., 1976, ''Plastic Internal Variables Formalism of Cyclic Plasticity,'' ASME J. Appl. Mech., **43**, pp. 645–651.
[16] Corona, E., 1996, ''Buckling of Elastic-Plastic Structural Members Under Cyclic Loading,'' NSF Proposal ID CMS-9610510.

R. Fan
Graduate Student

S. K. Singh
Graduate Student

C. D. Rahn
Associate Professor,
Mem. ASME

Department of Mechanical and Nuclear Engineering,
The Pennsylvania State University,
University Park, PA 16802

Modal Analysis of Ballooning Strings With Small Curvature

During the manufacture and transport of textile products, yarns are rotated at high speed. The surface of revolution generated by the rotating yarn is called a balloon. The dynamic response of the balloon to varying rotation speed, boundary excitation, and aerodynamic disturbances affects the quality of the associated textile product. Resonance, in particular, can cause large tension variations that reduce product quality and may cause yarn breakage. In this paper, the natural frequencies and mode shapes of a single loop balloon are calculated to predict resonance. The three-dimensional nonlinear equations of motion are simplified under assumptions of small displacement and quasi-static axial motion. After linearization, Galerkin's method is used to calculate the mode shapes and natural frequencies. Experimental measurements of the steady-state balloon shape and the first two natural frequencies and mode shapes are compared with theoretical predictions.
[DOI: 10.1115/1.1355776]

Introduction

Textile processes such as spinning, twisting, and unwinding involve the rotation of yarns at high speed. To an inertial observer, the rotating string blurs to produce a balloon or surface of revolution formed by the rotating yarn. Relative to a frame that rotates with the string, a steady balloon is stationary. For the heavy yarns with negligible air drag studied in this research, the steady yarn displacement resembles a planar catenary relative to the rotating frame. The dynamic response of the balloon to varying rotation speed, boundary excitation, and disturbances governs the quality of the textile product. In unwinding ([1]), for example, the yarn rotates around a stationary package as it is axially withdrawn. The back and forth boundary motion of the yarn on the package causes varying balloon rotation speed ([2]), introducing both axial and transverse disturbances. Resonance can cause large tension variations that reduce product quality and may cause yarn breakage.

Recent studies on ballooning strings have shown a variety of interesting dynamic behaviors. The numerical studies by Batra et al. [3] and Fraser [4] and theoretical and experimental studies by Zhu et al. [5] show balloons can have multiple shapes for fixed parameters due to the highly nonlinear governing equations. The linear vibration analysis by Zhu et al. [5] captures the experimentally observed jump phenomena and flutter instabilities of lightweight yarns with significant air drag. The number of loops in the balloon shape increases with increasing string length or decreasing tension. Linear dynamic investigations by Zhu et al. [5] and Stump et al. [6] show that single loop balloons are stable, one and a half loop balloons are divergent unstable, and double loop balloons may be flutter unstable for sufficiently low air drag. Zhu et al. [7] also show string extensibility has limited effect on balloon stability for most textile yarns. The limit cycling of double loop balloons is studied analytically and numerically by Zhu et al. [8] and Clark et al. [9], respectively.

In this paper, the natural frequencies and mode shapes of a single loop balloon are calculated to predict resonance. The three-dimensional nonlinear equations of motion are simplified via small steady-state displacement and vibration assumptions ([10]), that have been experimentally shown to be accurate for cable analysis ([11]). Axial vibration is assumed to propagate instantaneously (i.e., quasistatically) ([12]). Finally, the equations are linearized and Galerkin's method is used to calculate mode shapes and natural frequencies. Experimental measurements of the steady-state balloon shape and the first two natural frequencies and mode shapes are compared with theoretical predictions.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 7, 1999; final revision, Aug. 29, 2000. Associate Editor: A. A. Ferri. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS

Equations of Motion

Nonlinear Equations. Figure 1 shows a schematic diagram of the ballooning string system. The string is modeled as a perfectly flexible one-dimensional continuum pinned at the top and attached to an eyelet of length D at the bottom. The bottom eyelet is separated by perpendicular distance H from the top eyelet and rotates with angular velocity Ω. The unstressed, steady-state, and final configurations are given by χ^0, χ^i, and χ^f, respectively. The rotation speed and yarn mass per unit length ρA are assumed to be sufficiently large so that gravitational and aerodynamic forces can be neglected, respectively.

The steady-state displacement $\mathbf{R}^i(S,T)=R_1\mathbf{e}_1+R_2\mathbf{e}_2$ locates χ^i where S is the arc length coordinate measured along the steady state configuration and $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are Cartesian coordinates rotating about $\mathbf{e}_1$ with angular velocity Ω. The relative displacement between the final configuration χ^f located by $\mathbf{R}^f(S,T)$ and the steady-state configuration is

$$\mathbf{U}(S,T)=\mathbf{R}^f-\mathbf{R}^i=U_1\mathbf{l}_1+U_2\mathbf{l}_2+U_3\mathbf{l}_3, \tag{1}$$

where $\mathbf{l}_1$, $\mathbf{l}_2$, and $\mathbf{l}_3$ are the unit tangential, normal, and binomial vectors, respectively.

Following Zhu et al. [5], we obtain the nonlinear strain

$$\varepsilon=\frac{1}{2}(U_{1,S}^2+U_{2,S}^2+U_{3,S}^2)+\frac{1}{2}K^2(U_1^2+U_2^2)+K(U_1U_{2,S}-U_2U_{1,S})+U_{1,S}-KU_2, \tag{2}$$

where $K=\sqrt{R_{1,SS}^2+R_{2,SS}^2}$ is the steady-state curvature and comma subscripts indicate partial differentiation. Note that Eq. (2) neglects the strain between the initial and steady-state configurations. This inextensibility assumption implies

$$\left|\frac{\partial\mathbf{R}^i}{\partial S}\right|=1 \tag{3}$$

and is appropriate for most textile yarns because of their high axial stiffness ([7]).

The field equations are

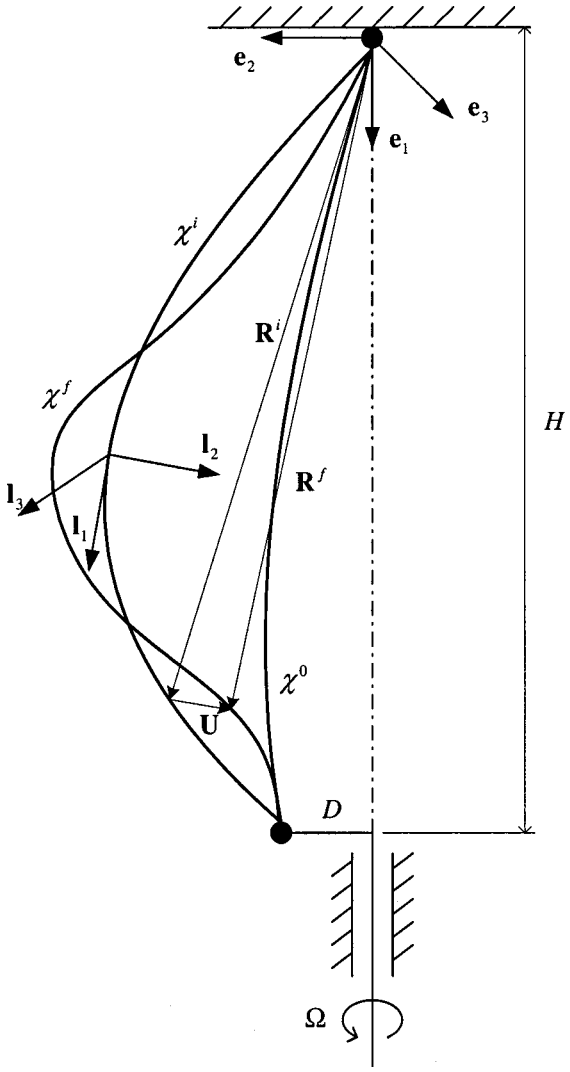


Fig. 1 Schematic diagram of a ballooning string system

$$\rho A U_{1,TT} + 2\rho A\beta\Omega U_{3,T} - \rho A\Omega^2(\beta^2 U_1 + \beta R_2 - \alpha\beta U_2) - [(P + EA\varepsilon)(1 + U_{1,S} - KU_2)]_{,S} + K(P + EA\varepsilon)(U_{2,S} + KU_1) = 0 \quad (4)$$

$$\rho A U_{2,TT} - 2\rho A\alpha\Omega U_{3,T} - \rho A\Omega^2(\alpha^2 U_2 - \alpha R_2 - \alpha\beta U_1) - [(P + EA\varepsilon)(U_{2,S} + KU_1)]_{,S} - K(P + EA\varepsilon)(1 + U_{1,S} - KU_2) = 0 \quad (5)$$

$$\rho A U_{3,TT} - 2\rho A\Omega(\beta U_{1,T} - \alpha U_{2,T}) - \rho A\Omega^2 U_3 - [(P + EA\varepsilon)(U_{3,S}]_{,S} = 0 \quad (6)$$

where $P(S)$ and EA are the string steady-state tension and longitudinal stiffness, respectively. The direction cosines $\alpha = \mathbf{e}_1 \cdot \mathbf{l}_1 = dR_1/dS$ and $\beta = \mathbf{e}_2 \cdot \mathbf{l}_1 = dR_2/dS$. The boundary conditions are

$$\mathbf{R}(0) = 0, \quad R_1(L) = H, \quad R_2(L) = D, \qquad \mathbf{U}(0,T) = 0, \quad \mathbf{U}(L,T) = 0, \quad (7)$$

where H is the balloon height.

To simplify the analysis, the following nondimensional variables are introduced:

$$\mathbf{U} = \mathbf{u}L, \quad S = sL, \quad T = \phi t, \quad R_2 = r_2 L, \quad H = hL, \quad p = \frac{P}{P_e}, \quad \phi = \sqrt{\frac{\rho AL^2}{P_e}}, \quad \omega = \Omega\phi, \quad \lambda = \frac{EA}{P_e}, \quad D = dL, \quad K = \frac{k}{L}, \quad (8)$$

where $P_e = P(0)$. Substitution of Eq. (8) into Eqs. (4)–(6) produces

$$u_{1,tt} + 2\beta\omega u_{3,t} - \omega^2(\beta^2 u_1 + \beta r_2 - \alpha\beta u_2) - [(p + \lambda\varepsilon)(1 + u_{1,s} - ku_2)]_{,s} + k(p + \lambda\varepsilon)(u_{2,s} + ku_1) = 0 \quad (9)$$

$$u_{2,tt} - 2\alpha\omega u_{3,t} - \omega^2(\alpha^2 u_2 - \alpha r_2 - \alpha\beta u_1) - [(p + \lambda\varepsilon)(u_{2,s} + ku_1)]_{,s} + k(p + \lambda\varepsilon)(1 + u_{1,s} - ku_2) = 0 \quad (10)$$

$$u_{3,tt} - 2\omega(\beta u_{1,t} - \alpha u_{2,t}) - \omega^2 u_3 - [(p + \lambda\varepsilon)u_{3,s}]_{,s} = 0 \quad (11)$$

The boundary conditions Eqs. (7) become

$$\mathbf{u}(0,t) = \mathbf{u}(1,t) = 0, \qquad r_2(0) = 0, \quad r_2(1) = d. \quad (12)$$

The strain ε has the same form as Eq. (2) with lowercase letters.

Ordering of the system variables as follows:

$$u_1 = \epsilon^2 u_1, \quad u_2 = \epsilon u_2, \quad u_3 = \epsilon u_3, \quad k = \epsilon k, \quad (13)$$

where ϵ is a small parameter, produces the simplified strain equation

$$\varepsilon = u_{1,s} + \frac{1}{2}(u_{2,s}^2 + u_{3,s}^2) - ku_2 + O(\epsilon^4). \quad (14)$$

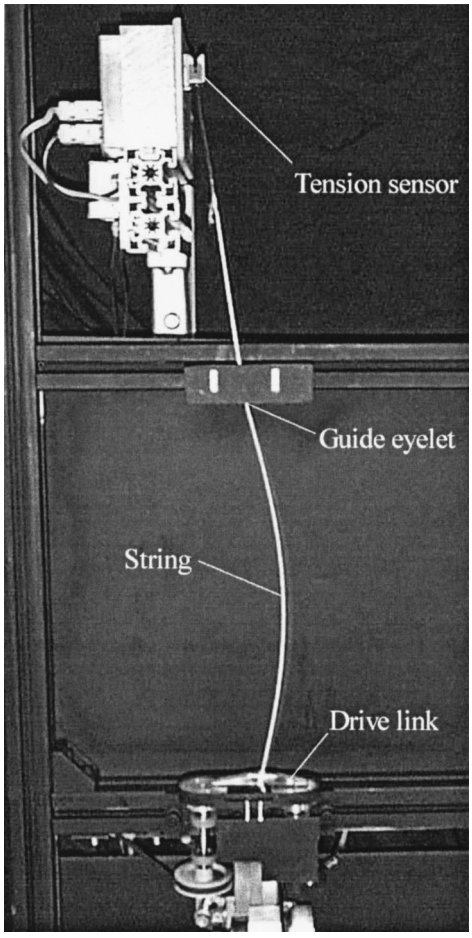


Fig. 2 Experimental setup

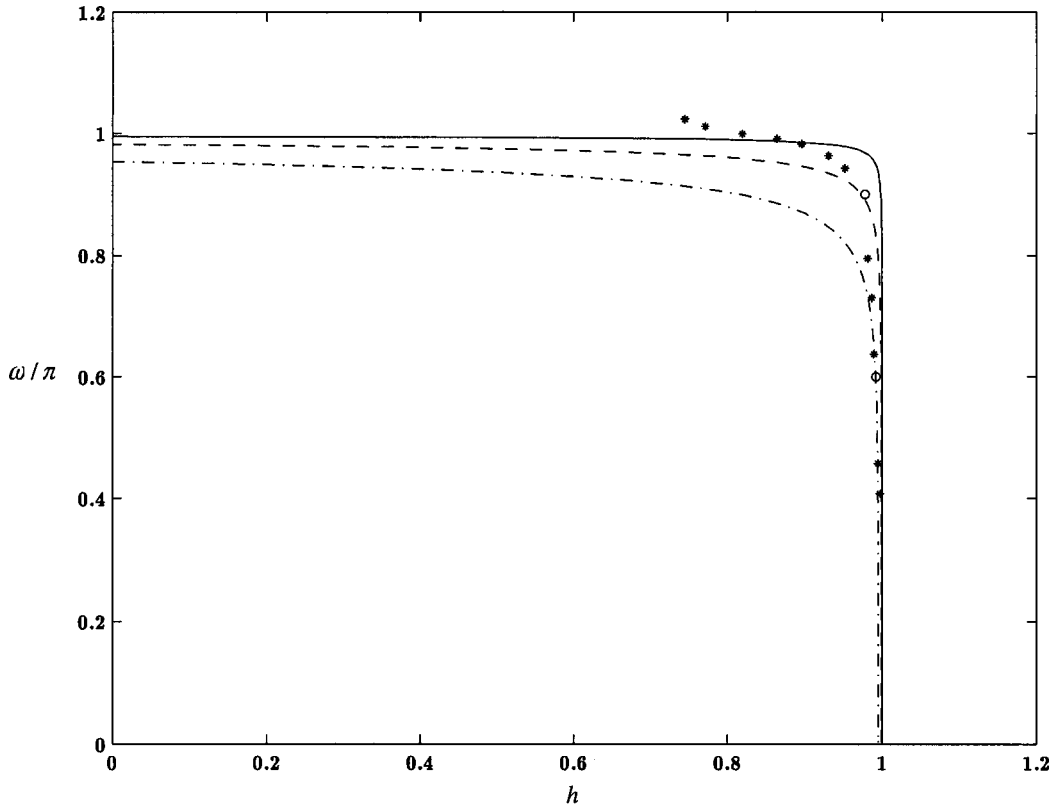


Fig. 3 The dependence on the nondimensional rotation speed ω on the nondimensional balloon height h and eyelet length d. Theoretical solid ($d=0.01$), dashed ($d=0.038$) and dash-dotted ($d=0.1$) curves and experimental data (*) are shown.

Assuming u_1 is $O(\epsilon^2)$ agrees with the inextensibility assumption Eq. (3). The remaining variables are ordered,

$$r_2=\epsilon r_2, \quad \beta=\epsilon\beta, \quad \alpha=1, \quad \lambda=\lambda/\epsilon. \tag{15}$$

Substitution of Eqs. (13) and (15) in Eqs. (9)–(11) produces

$$[p+\lambda\varepsilon]_{,s}=0+O(\epsilon^2) \tag{16}$$

$$u_{2,tt}-2\omega u_{3,t}+\omega^2(r_2-u_2)-[(p+\lambda\varepsilon)u_{2,s}]_{,s}-\lambda k\varepsilon-kp=0+O(\epsilon^3) \tag{17}$$

$$u_{3,tt}+2\omega u_{2,t}-\omega^2 u_3-[(p+\lambda\varepsilon)u_{3,s}]_{,s}=0+O(\epsilon^3). \tag{18}$$

We have assumed that both the steady-state and relative displacements are small. This produces the simple form of Eq. (16) and subsequent quasi-static elimination of $u_1(s,t)$.

Steady-State Equations. Solutions of Eqs. (16)–(18) depends on the steady-state tension p, curvature k, and displacement r_2. The steady-state equations result from substitution of $u_1=u_2=u_3=\varepsilon=0$ in the field Eqs. (16)–(17) to produce

$$p_{,s}=0 \tag{19}$$

$$\omega^2 r_2-kp=0. \tag{20}$$

Integration of Eq. (19) yields $p=1$. Substitution of $k=r_{2,ss}$ ($r_{1,ss}=\alpha_{,s}=0$) in Eq. (20) produces

$$r_{2,ss}+\omega^2 r_2=0 \tag{21}$$

with a solution satisfying the boundary conditions (12) as follows:

$$r_2(s)=\frac{d\sin\omega s}{\sin\omega} \tag{22}$$

$$k(s)=\frac{d\omega^2\sin\omega s}{\sin\omega}.$$

In this paper, only single loop balloons with $\omega<\pi$ are studied, ensuring $k>0$.

The balloon height is related to the steady-state displacement using inextensibility Eq. (3) as follows:

$$h=\int_0^1\sqrt{1-r_{2,s}^2}\,ds=\int_0^1\left(1-\frac{1}{2}r_{2,s}^2\right)ds+O(\epsilon^4). \tag{23}$$

Substitution of Eq. (22) into Eq. (23) yields

$$h=1-\frac{\omega d^2}{4\sin\omega}\left(\cos\omega+\frac{\omega}{\sin\omega}\right). \tag{24}$$

Equation (24) shows that changing Ω while keeping H, L, and D fixed does not change ω. The steady-state tension P_e increases quadratically with speed to make ω constant. In the experiments, we change ω by changing H while maintaining L, D, and Ω constant.

Quasi-static Equations of Vibration. After subtraction of the steady-state solution from Eq. (16), we have

$$\varepsilon_{,s}=0. \tag{25}$$

Thus, vibration in the axial direction is neglected under the assumption that the yarn stretches in a *quasi-static* manner ([12]). Integration of Eq. (25) and substitution into Eq. (14) yields

$$\varepsilon=\frac{1}{2}(u_{2,s}^2+u_{3,s}^2)+u_{1,s}-ku_2=g(t), \tag{26}$$

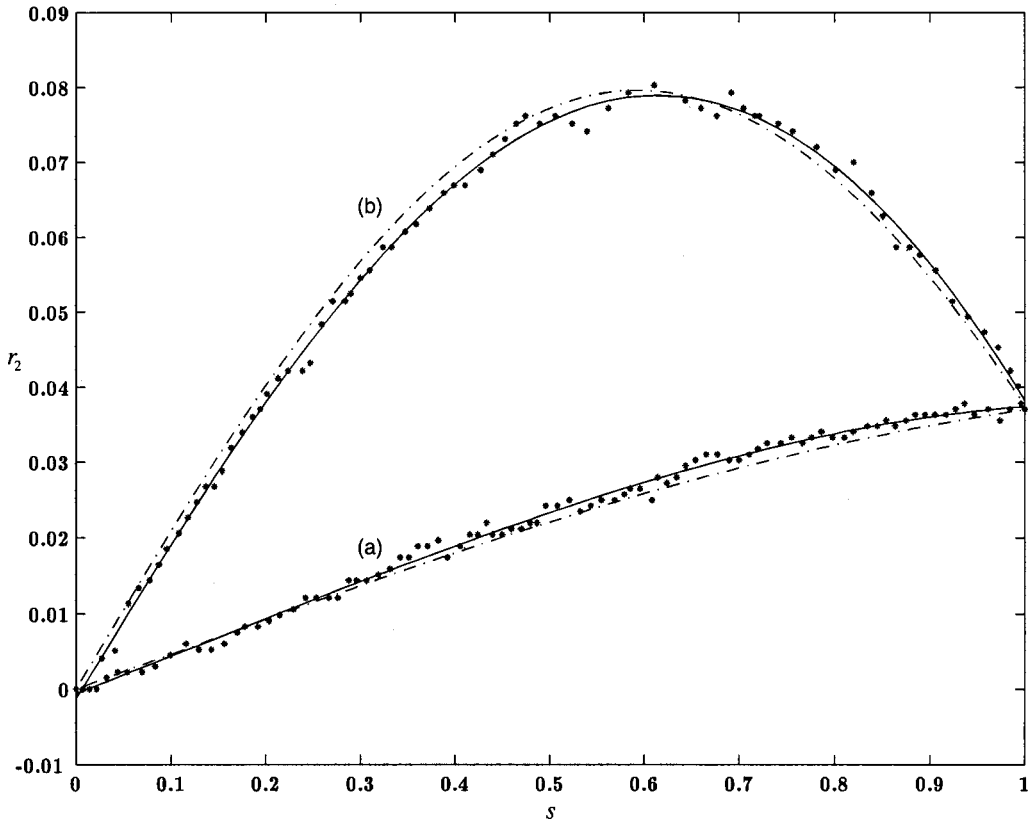


Fig. 4 Theoretical (dash-dotted and experimental data (*) and best fit line (solid) steady-state in-plane displacement: (*a*) $\omega=0.6\pi$, $d=0.038$, $\gamma=400$; (*b*) $\omega=0.9\pi$, $d=0.038$, $\gamma=400$. Experimental data correspond to circled points in Fig. 3.

where $g(t)$ is an arbitrary function of time. Integration and use of the boundary conditions (12) give

$$u_1(s,t)=\int_0^s\left[ku_2-\frac{1}{2}(u_{2,s}^2+u_{3,s}^2)\right]ds+g(t)s, \tag{27}$$

where

$$g(t)=\int_0^1\left[\frac{1}{2}(u_{2,s}^2+u_{3,s}^2)-ku_2\right]ds=\varepsilon(t). \tag{28}$$

Substitution of Eq. (28) into Eqs. (17)–(18) and elimination of the steady-state terms produces

$$u_{2,tt}-2\omega u_{3,t}-\omega^2u_2-\lambda k\varepsilon-u_{2,ss}(1+\lambda\varepsilon)=0 \tag{29}$$

$$u_{3,tt}-2\omega u_{2,t}-u_3\omega^2-u_{3,ss}(1+\lambda\varepsilon)=0. \tag{30}$$

Canceling nonlinear terms in Eqs. (29) yields the linearized equations

$$u_{2,tt}-2\omega u_{3,t}-\omega^2u_2+\lambda k\int_0^1 ku_2ds-u_{2,ss}=0 \tag{31}$$

$$u_{3,tt}+2\omega u_{2,t}-\omega^2u_3-u_{3,ss}=0. \tag{32}$$

Modal Analysis. Galerkin's method is used for numerical balloon vibration analysis. The displacements are represented by n term separable series of the form

$$u_2(s,t)=\sum_{i=1}^{n}a_{2i}(t)\sin(i\pi s) \tag{33}$$

$$u_3(s,t)=\sum_{j=1}^{n}a_{3j}(t)\sin(j\pi s) \tag{34}$$

that satisfy the pinned boundary conditions. Substitution of Eqs. (33) and (34) into Eqs. (31) and (32) and application of Galerkin's method provide the discretized equation of motion

$$\ddot{\mathbf{a}}+\mathbf{G}\dot{\mathbf{a}}+\mathbf{K}\mathbf{a}=\mathbf{0} \tag{35}$$

where the coordinate vector

$$\mathbf{a}=[a_{21}\ \ldots\ a_{2n}\ \ a_{31}\ \ldots\ a_{3n}]^T.$$

The stiffness matrix $\mathbf{K}=\mathbf{K}_1+\mathbf{K}_2$ with

$$\mathbf{K}_1=\begin{bmatrix}\mathbf{K}_{10} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix},\quad \mathbf{K}_{10}(i,j)=\frac{(-1)^{i+j}2ij\pi^2\lambda d^2\omega^4}{[\omega^2-(i\pi)^2][\omega^2-(j\pi)^2]},$$

$$\mathbf{K}_2=\begin{bmatrix}\mathrm{diag}\{(i\pi)^2\}-\omega^2\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathrm{diag}\{(i\pi)^2\}-\omega^2\mathbf{I}\end{bmatrix}, \tag{36}$$

and the gyroscopic matrix

$$\mathbf{G}=\begin{bmatrix}\mathbf{0} & -2\omega\mathbf{I}\\ 2\omega\mathbf{I} & \mathbf{0}\end{bmatrix}. \tag{37}$$

The eigenvalues λ_n of the state matrix associated with Equation (35)

$$\mathbf{A}=\begin{bmatrix}0 & \mathbf{I}\\ -\mathbf{K} & -\mathbf{G}\end{bmatrix} \tag{38}$$

are determined numerically using MATLAB.

Experiments

Experimental Setup. The experimental setup is shown in Fig. 2. The string attaches between a rotating eyelet on a four-bar linkage ($D=0.0139$ m) and an inductive tension sensor mounted

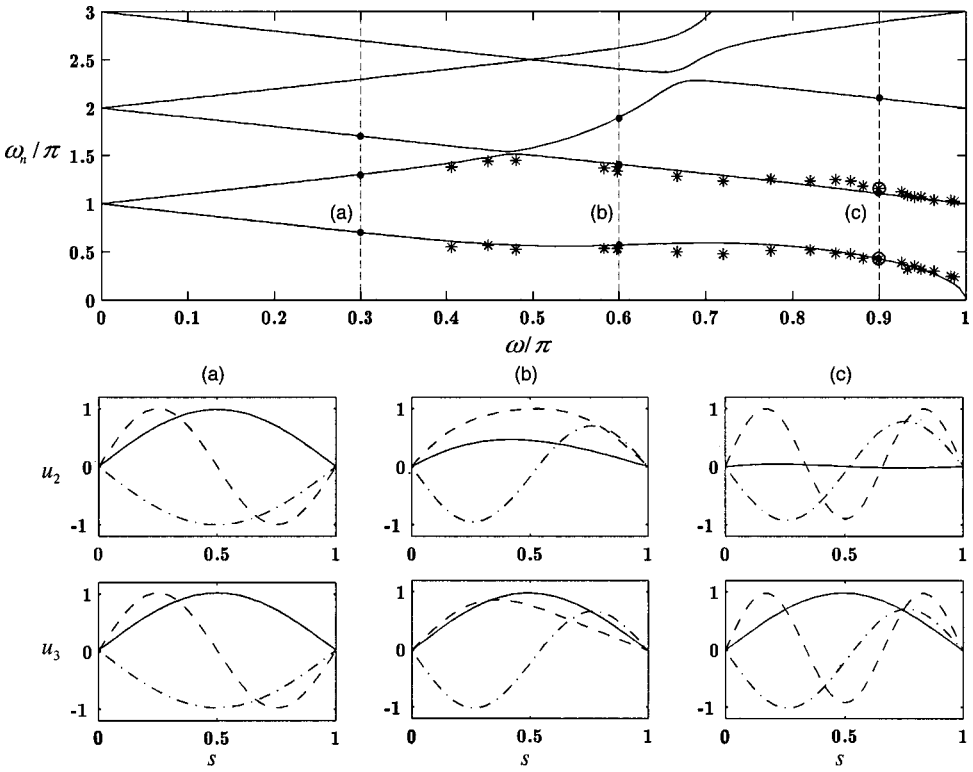


Fig. 5 Theoretical (solid) and experimental (*) natural frequencies ($d=0.038$, $\gamma=400$). Mode shapes (1st solid, 2nd dash-dotted, 3rd dashed): (*a*) $\omega=0.3\pi$, (*b*) $\omega=0.6\pi$, (*c*) $\omega=0.9\pi$.

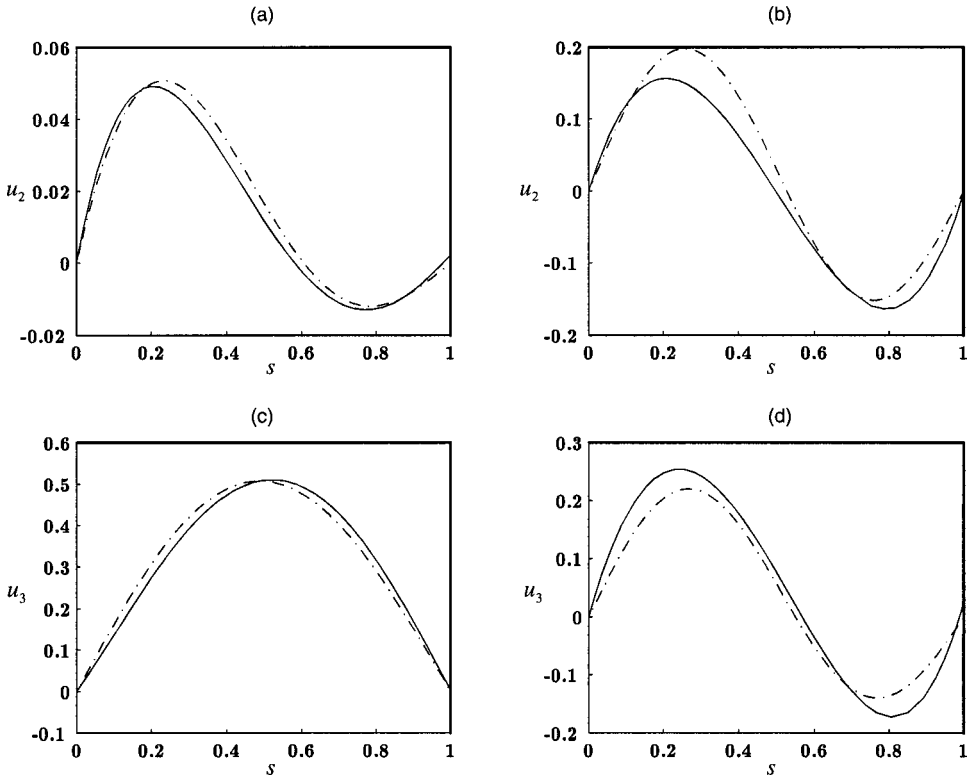


Fig. 6 Theoretical (dash-dotted) and experimental (solid) mode shapes ($\omega=0.9\pi$, $d=0.038$, $\gamma=400$): (*a*) first mode in-plane, (*b*) second mode in-plane, (*c*) first mode out-of-plane, (*d*) second mode out-of-plane. Experimental mode shapes correspond to the circled points in Fig. 5.

above an upper eyelet. The four-bar linkage is driven by a PC-controlled permanent magnet DC motor so that the eyelet in the rotating link undergoes the desired circular motion. An axially stiff ($EA=237$ N) three strand continuous filament polyester string with linear density $\rho A=0.001$ kg/m is used. The nominal string length and steady-state tension are 0.376 m and 0.091 N, respectively. The nondimensional parameter $\lambda=2604$ is large, justifying the quasi-static assumption. Steady-state and mode-shape displacement measurements are generated using a strobe and a digital camera. A small sinusoidal variation of the rotation speed excites vibration and allows determination of the natural frequencies and mode shapes. The sinusoidal input sweeps through a frequency range and the tension output is monitored to determine peak vibration amplitudes (resonances).

Experimental Results. Figure 3 plots the theoretical and experimental relationship between the nondimensional rotation speed ω and the nondimensional balloon height h and eyelet length d. The theoretical curves come from Eq. (24). The theory and experiment show that ω decreases with increasing h and/or d.

Figure 4 compares the theoretical and experimental balloon shapes for the two circled points in Fig. 3. Best fit lines through the experimental data are also shown. The results agree to within one percent for the small radius balloon in Fig. 4(a), corresponding to a small $\omega=0.6\pi$, and from Fig. 3, a large $h=0.99$. Large h implies that the balloon is stretched taut because the height is almost equal to the string length. For larger ω (smaller h), the balloon displacement is larger as shown in Fig. 4(b). Reduction of $h<0.7$ causes the balloon to collapse or flutter chaotically due to the system nonlinearities.

Several assumptions increase the difference between theory and experiment to seven percent for the large-amplitude displacement case (Fig. 4(b)). First, the relatively low speed of the experimental tests (635 rpm) means that gravity loading is significant. Figure 4(b) clearly shows gravity pulling the steady-state shape downward relative to the theoretical prediction. Second, air drag in the experiment pushes the balloon slightly out-of-plane. Finally, the large-amplitude experimental results are influence by nonlinear effects.

Figure 4 also explains why ω varies with h and d as shown in Fig. 3. As h increases, the balloon becomes taut and the steady-state tension increases. The nondimensional rotation speed is inversely related to steady-state tension and therefore decreases. If d increases dramatically, the balloon also becomes taut, increasing steady-state tension and decreasing ω.

Three nondimensional parameters (d, λ, and ω) govern the system response. In the experiment, the length, height, and rotation speed of the balloon can be easily controlled. Maintaining a constant balloon length while changing the balloon height ensures that d remains constant. The nondimensional variable

$$\gamma=\frac{\lambda}{\omega^2}=\frac{EA}{\rho A\Omega^2L^2} \tag{39}$$

is also constant if the experiments are run at constant speed. Thus, the experimental results are generated for constant d and γ and for varying ω by maintaining constant balloon length and rotation speed but changing the balloon height.

In Zhu et al. [7], the inextensibility assumption was found to affect the results by less than one percent if $\gamma/d^2>1000$. The experimental parameters yield $\gamma/d^2=2.8\times10^5\gg1000$, providing the validity of the inextensibility assumption.

Figure 5 shows how the natural frequencies and mode shapes change with nondimensional rotation speed ω for fixed $d=0.038$ and $\gamma=400$. At $\omega=0$ the balloon is taut and the in-plane and out-of-plane vibration frequencies repeat at the pinned string natural frequencies $\omega_n=n\pi$. As ω increases, the frequencies separate into forward and backward whirling modes. The forward whirling modes move faster relative to a ground observer and thus have higher frequency. The mode shapes for the first three modes at $\omega=0.3\pi$ shown in Fig. 5(a) have identical in-plane and out-of-plane amplitudes, indicating circular whirling. Above $\omega=0.4\pi$ the curvature begins to become significant and the stiffening integral term in Eq. (31) increases the frequency of modes that are not orthogonal to $k(s)$. As $k(s)$ is approximately a half-sine, the first two modes are most affected by increasing curvature. The other modes, being essentially sinusoids with spatial wavelengths equal to integral divisors of the first mode wavelength, are orthogonal to $k(s)$ and are not as sensitive to increasing curvature. At $\omega=0.6\pi$ (Fig. 5(b)), the first forward whirling mode and the second backward whirling mode exchange order due to an eigenvalue curve veering near $\omega=0.5\pi$. The first backward whirling mode frequency increases slightly from $\omega=0.5\pi$ to $\omega=0.7\pi$ due to curvature stiffening that reduces the in-plane modal amplitude, producing elliptical whirling. At $\omega=0.9\pi$ (Fig. 5(c)), the first backward whirling mode is almost entirely confined to the out-of-plane with the small in-plane component having a full sine shape. The first forward whirling mode has veered up off the graph. The second and third modes are the second and third backward whirling modes, respectively. At $\omega=\pi$, the first backward whirling mode buckles and the first forward whirling mode veers to infinity. The other forward and backward whirling modes change by $\pm\pi$ from their values at $\omega=0$, respectively.

The experimentally measured first two natural frequencies, plotted as asterisks on Fig. 5, match the theory curves to within 13 percent for the first mode and four percent for the second mode. Above $\omega=.97\pi$, the balloon becomes too large and tends to collapse into a double balloon or flutter. Thus the theory is reasonably accurate for $\omega<.97\pi$ or $r_2<0.4$. Below $\omega=0.5\pi$, the balloon is too taut for mode shape measurement. Figure 6 shows the theoretical and experimental mode shapes for the first two modes at $\omega=.9\pi$ in reasonably good agreement. Note that the first mode is shaped like a half-sine out-of-plane and small-amplitude full-sine in-plane as predicted by the theory.

Conclusions

The calculated natural frequencies and mode shapes from the small sag ($R_2<0.4L$), quasi-static ($P_e\ll EA$), single loop ($\Omega<\pi\sqrt{P_e/\rho AL^2}=\Omega_c$) balloon model match the experimental results over a broad range of speeds. The first backward whirling mode buckles at the critical speed $\Omega=\Omega_c$. The first forward whirling mode is sensitive to balloon curvature and veers upward in frequency with increasing ω. The remaining modes start at $\omega_n=n\pi$ and change approximately linearly as ω increases with slopes of $\pm\pi$ for forward and backward whirling modes, respectively.

Acknowledgment

This research was supported by the National Textile Center.

References

[1] Kong, X., Rahn, C., and Goswami, B., 1999, ''Steady State Unwinding of Yarn from Cylindrical Packages,'' Textile Res. J., **69**, No. 4, pp. 292–306.

[2] Reedman, A., 1998, ''Measurement of Lift-Off Position and Rotation Rate During Unwinding,'' M. S. thesis, Clemson University, Clemson, SC, Aug.

[3] Batra, S., Ghosh, T., and Zeidman, M., 1992, ''An Integrated Approach to Dynamic Analysis of the Ring Spinning Process—III,'' Textile Praxis Int., **47**, pp. 793–800.

[4] Fraser, W., 1993, ''On the Theory of Ring Spinning,'' Philos. Trans. R. Soc. London, Ser. A, **A342**, pp. 439–468.

[5] Zhu, F., Hall, K., and Rahn, C. D., 1998, ''Steady State Response and Stability of Ballooning Strings in Air,'' Int. J. Non-Linear Mech., **33**, No. 1, pp. 33–46.

[6] Stump, D., and Fraser, W., 1996, ''Transient Solutions of the Ring-Spinning Balloon Equations,'' ASME J. Appl. Mech., **63**, pp. 523–528.

[7] Zhu, F., Sharma, R., and Rahn, C. D., 1997, ''Vibrations of Ballooning Elastic Strings,'' ASME J. Appl. Mech., **64**, pp. 676–683.

[8] Zhu, F., and Rahn, C. D., 2000, ''Limit Cycle Prediction for Ballooning Strings,'' Int. J. Non-Linear Mech., **35**, No. 3, May, pp. 373–383.

[9] Clark, J. D., Fraser, W. B., Sharma, R., and Rahn, C. D., 1998, ''The Dynamic Response of a Ballooning Yarn: Theory and Experiment,'' Proc. R. Soc. London, Ser. A, **A454**, No. 1978, pp. 2767–2789.
[10] Perkins, N. C., 1992, ''Modal Interactions in the Non-Linear Response of Elastic Cables Under Parametric/External Excitation,'' Int. J. Non-Linear Mech., **27**, No. 2, pp. 233–250.
[11] Lee, C. L., and Perkins, N. C., 1995, ''Experimental Investigation of Isolated and Simultaneous Internal Resonances in Suspended Cables,'' ASME J. Vibr. Acoust., **117**, pp. 385–391.
[12] Irvine, H. M., and Caughey, T. K., 1974, ''The Linear Theory of Free Vibrations of a Suspended Cable,'' Proc. R. Soc. London, Ser. A, **A341**, pp. 299–315.

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 Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME International, Three Park Avenue, New York, NY 10016-5990, or to the Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

A New Lagrangian and a New Lagrange Equation of Motion for Fractionally Damped Systems

O. P. Agrawal
Professor, Mechanical Engineering and Energy Processes, Southern Illinois University, Carbondale, IL 62901

1 Introduction

All dynamic systems exhibit some degree of internal damping. Recent investigations have shown that a fractional derivative model provides a better representation of the internal damping of a material than an ordinary derivative model does. For a survey of fractional damping models and their applications to engineering systems, the readers are referred to Rossikhin and Shitikova [1] and the references therein. Traditionally, the Newton's law is used to model such nonconservative systems, and when a Lagrangian, Hamiltonian, variational, or other energy-based approach is used, it is modified so that the resulting equations match those obtained using the Newtonian's approach.

Several attempts have been made to include nonconservative forces in the Lagrangian and the Hamiltonian mechanics. Riewe [2,3] presented a succinct survey of research in this area. He also pointed out that a term proportional to $d^n x/dt^n$ in the Euler-Lagrange equation follows from a Lagrangian with a term proportional to $(d^{n/2}x/dt^{n/2})^2$. Hence, a frictional force of the form $c(dx/dt)$ may follow directly from a Lagrangian containing a term of the form $(d^{1/2}x/dt^{1/2})^2$. Using this as the starting point, he developed a new approach to mechanics that allows nonconservative terms (both ordinary and fractional damping) to be included in Lagrangians and Hamiltonians. This paper presents another form of a Lagrangian and the Lagrange equation that can be used to obtain equations of motion of systems whose damping forces are proportional to a fractional derivative of order j/n. With a minor change in the formulation, the resulting equations can be thought of as a state space representation of Riewe's formulation ([2,3]).

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 7, 1999; final revision, Apr. 24, 2000. Associate Editor: A. A. Ferri.

2 A New Lagrangian and a New Lagrange Equation of Motion

This section presents a new Lagrangian and a new Lagrange equation of motion for a fractionally damped system. There are several definitions of a fractional derivative. Here a fractional derivative is defined in the Riemann-Liouville sense ([4]):

$$D^\alpha x(t) = \frac{d^\alpha x(t)}{dt^\alpha} = \frac{1}{\Gamma(1-\alpha)}\frac{d}{dt}\int_0^t \frac{x(t-u)}{u^\alpha}\,du,$$

$$t>0,\ \ 0<\alpha<1. \tag{1}$$

This definition can be extended for $\alpha>1$. Here $x(t)$ represents a state space coordinate of the dynamical system. The Lagrangian and the Lagrange equation of motion are given as

$$L = \nu_1(\alpha)\tfrac{1}{2}(D^\alpha y)^T M D^\alpha y - \tfrac{1}{2}y^T K y + Q^T y, \tag{2}$$

and

$$\nu_1^{-1}(\alpha)\frac{d^\alpha}{dt^\alpha}\frac{\partial L}{\partial(D^\alpha y)} - \frac{\partial L}{\partial y} = 0 \tag{3}$$

where y is a state vector, M and K are the mass-like and the stiffness-like matrices, $\nu_1(\alpha)$ is a α-dependent coefficient, and Q is a vector of generalized forces. The purpose of $\nu_1(\alpha)$ is to make the formulation consistent with the variational (or Euler-Lagrange) approach ([2,3]). Since $\nu_1(\alpha)$ does not appear in the equations of motion, it will not be included, and its expression will not be given here (for its expression see [2,3]). The Lagrangian L defined in Eq. (2) is applicable for positive rational α only. The dimensions of y, M, and K depend on the denominator part of α. Therefore, in this setting, L cannot be developed for irrational α. Matrices M and K are not the traditional mass and stiffness matrices. It will be seen that M may contain the mass and the damping, and K may contain the mass, the damping, and the stiffness. In the case of zero damping, M and K reduce to the mass and the stiffness matrices. It is assumed that, for the functions considered here, the composition rule applies.

Substituting Eq. (2) into Eq. (3), we get the equation of motion as

$$MD^{2\alpha}y + Ky = Q. \tag{4}$$

To generate L, we need α, y, M, K, and Q. α is half of the lowest common fractional derivative order. Thus, for a force of the form $(d^{j/n}x/dt^{j/n})$, $\alpha = 1/(2n)$. α can be smaller. However, the smaller α introduces no new modeling capability. The elements of the

state vector y are defined as $y_i = D^{2\alpha} y_{i+1} = D^{1/n} y_{i+1}$, $(i = 1, \ldots, l-1)$, and $y_l = x(t)$, where $l = 2n$ is the dimension of the vector.

Now consider that mD^2x, $cD^{j/n}x$, and kx represent the inertia, the damping, and the spring forces, where m, c, and k are, respectively, the mass, the damping coefficient, and the stiffness of the system. In this case, M and K are given as

$$M = \begin{bmatrix} 0 & \cdots & 0 & \cdots & m \\ \vdots & \cdots & \vdots & \iddots & \vdots \\ 0 & \cdots & m & \cdots & c \\ \vdots & \iddots & \vdots & \iddots & \vdots \\ m & \cdots & c & \cdots & 0 \end{bmatrix},$$

$$K = \begin{bmatrix} 0 & \cdots & \cdots & 0 & -m & 0 \\ \vdots & \iddots & \iddots & \iddots & \iddots & 0 \\ \vdots & \iddots & \iddots & \iddots & -c & 0 \\ 0 & \iddots & \iddots & \iddots & \iddots & \vdots \\ -m & \iddots & -c & 0 & \iddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & k \end{bmatrix}. \tag{5}$$

In matrix M the off-diagonal contains m and the jth off-diagonal measured from the bottom right corner contains c. In matrix K the bottom right corner contains k, the elements left to the off-diagonal contain $-m$, and all except the first and the last elements of the $(j+1)$th off-diagonal measured from the bottom right corner contains $-c$. Note that M and K are symmetric. Structure of these matrices will be explained further using two examples. Finally, vector Q is given as

$$Q = [0, \cdots, 0, F] \tag{6}$$

where F is the generalized force.

The Lagrangians and the Lagrange equations of motion for two fractionally damped systems are given below.

Example 1. Damping force $= c\dot{x}$.

For this system, $\alpha = 1/2$. Vector y, matrices $M = M_1$ and $K = K_1$, the Lagrangian L, and the Lagrange equation are

$$y = [y_1 \;\; y_2]^T = [\dot{x} \;\; x]^T,$$

$$M_1 = \begin{bmatrix} 0 & m \\ m & c \end{bmatrix}, \quad K_1 = \begin{bmatrix} -m & 0 \\ 0 & k \end{bmatrix},$$

$$L = \tfrac{1}{2}(D^{1/2}y)^T M_1 D^{1/2} y - \tfrac{1}{2} y^T K_1 y,$$

$$\frac{d^{1/2}}{dt^{1/2}} \frac{\partial L}{\partial (D^{1/2}y)} - \frac{\partial L}{\partial y} = 0.$$

These terms give $M_1 D y + K_1 y = 0$, which is equivalent to $m\ddot{x} + c\dot{x} + kx = 0$.

Example 2. Damping force $= cD^{4/3}x$.

This example shows the locations of c when j in the force term $cD^{j/n}x$ is more than 1. For this system, $\alpha = 1/6$. Vector y, matrices $M = M_2$ and $K = K_2$, the Lagrangian L, and the Lagrange equation are

$$y = [y_1 \;\; y_2 \;\; y_3 \;\; y_4 \;\; y_5 \;\; y_6]^T$$
$$= [D^{5/3}x \;\; D^{4/3}x \;\; D^1 x \;\; D^{2/3}x \;\; D^{1/3}x \;\; x]^T,$$

$$M_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & m \\ 0 & 0 & 0 & 0 & m & 0 \\ 0 & 0 & 0 & m & 0 & c \\ 0 & 0 & m & 0 & c & 0 \\ 0 & m & 0 & c & 0 & 0 \\ m & 0 & c & 0 & 0 & 0 \end{bmatrix},$$

$$K_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & -m & 0 \\ 0 & 0 & 0 & -m & 0 & 0 \\ 0 & 0 & -m & 0 & -c & 0 \\ 0 & -m & 0 & -c & 0 & 0 \\ -m & 0 & -c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k \end{bmatrix},$$

$$L = \tfrac{1}{2}(D^{1/6}y)^T M_2 D^{1/6} y - \tfrac{1}{2} y^T K_2 y,$$

$$\frac{d^{1/6}}{dt^{1/6}} \frac{\partial L}{\partial (D^{1/6}y)} - \frac{\partial L}{\partial y} = 0.$$

These terms give $M_2 D^{1/3} y + K_2 y = 0$, which is equivalent to $m\ddot{x} + cD^{4/3}x + kx = 0$.

Note that M may contain m and c, and K may contain m, c, and k. Also, note that in the examples c can be set to zero to obtain the differential equations of motion of an undamped system in a higher dimension. Similarly, α in Example 1 can be set to 1/4 to obtain the differential equations of motion of a damped system in a higher dimension. However, such increase in dimensions adds no benefit. Further, as the order of the fractional derivative α moves from 1 towards 0 (1 towards 2) the damper behaves like a spring (mass).

Several techniques have been developed to solve the resulting set of fractional differential equations. Suarez and Shokooh [5] presented an eigenvector expansion method to solve these differential equations. Other methods to solve these fractional differential equations include, for example, Laplace transform and direct techniques similar to the techniques for ordinary differential equations ([6–8]), and the numerical techniques ([9]).

3 Additional Remarks

Riewe [2,3] developed a new approach to mechanics with fractional derivatives that includes the nonconservative Hamiltonian, Canonical transformations and the Jacobi theory. The approach can also be used to develop similar Hamiltonians and the Hamilton equations. To this end, we propose the following Hamiltonian and Hamilton equations for fractional systems.

$$H = (D^\alpha y)^T p - L = (D^\alpha y)^T p - \tfrac{1}{2}(D^\alpha y)^T M D^\alpha y + \tfrac{1}{2} y^T K y,$$

$$\frac{\partial H}{\partial p} = D^\alpha y, \quad \text{and} \quad \frac{\partial H}{\partial y} = -D^\alpha p,$$

where $p = MD^\alpha y$. It can be shown that the above equations lead to the correct equation of motion. Using these equations, the Hamiltonian can also be written as

$$H = \tfrac{1}{2}(D^\alpha p)^T M^{-1} D^\alpha p + \tfrac{1}{2} y^T K y,$$

which is similar to the total energy term for a conservative system.

Acknowledgment

The author would like to thank the reviewers for their suggestions to improve the quality of the paper.

References

[1] Rossikhin, Y. A., and Shitikova, M. V., 1997, ''Applications of Fractional Calculus to Dynamic Problems of Linear and Nonlinear Hereditary Mechanics of Solids,'' Appl. Mech. Rev., **50**, No. 1, pp. 15–67.

[2] Riewe, F., 1996, ''Nonconservative Lagrangian and Hamiltonian Mechanics,'' Phys. Rev. E, **53**, No. 2, pp. 1890–1899.

[3] Riewe, F., 1997, ''Mechanics With Fractional Derivative,'' Phys. Rev. E, **55**, No. 3, pp. 3581–3592.

[4] Oldham, K. B., and Spanier, J., 1974, *The Fractional Calculus*, Academic Press, New York.

[5] Suarez, L. E., and Shokooh, A., 1997, ''An Eigenvector Expansion Method for the Solution of Motion Containing Fractional Derivatives,'' ASME J. Appl. Mech., **64**, pp. 629–635.

[6] Miller, K. S., and Ross, B., 1993, *An Introduction to the Fractional Calculus*

and Fractional Differential Equations, John Wiley and Sons, New York.
[7] Mainardi, F., 1997, ''Fractional Calculus: Some Basic Problems in Continuum and Statistical Mechanics,'' *Fractals and Fractional Calculus in Continuum Mechanics*, A. Carpinteri and F. Mainardi, eds., Springer-Verlag, New York, pp. 291–348.
[8] Podlubny, I., 1999, *Fractional Differential Equations*, Academic Press, New York.
[9] Yuan, L., and Agrawal, O. P., 1998, ''A Numerical Scheme for Dynamical Systems Containing Fractional Derivatives, *Proceedings of the DETC98, ASME Design Engineering Technical Conferences*, Paper No. DETC98/MECH 5857, Sept. 13–16, Atlanta, GA.

On the Unification of Yield Criteria

S. C. Fan
Associate Professor, School of Civil & Structural Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798

M.-H. Yu
Professor, School of Civil Engineering & Mechanics, Xi'an Jiaotong University, Xi'an 710049, China

S.-Y. Yang
Post Doctoral Fellow, School of Civil & Structural Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798

A piecewise linear unified yield criterion called the twin-shear-unified was proposed. It is based on a kind of orthogonal dodecahedron stress element. The effects of intermediate principal stress are taken into account such that most available yield loci on the π-plane are embraced in a unified manner. Besides, it is capable to represent not only convex limit surfaces but also nonconvex limit surfaces. [DOI: 10.1115/1.1320451]

1 Introduction

For many engineering materials, two characteristic strength properties are crucial, i.e., initial and subsequent yield properties. The initial yield defines the critical state when the material under the complex stress state starts to yield. The subsequent yield deals with the post yield phenomena. It describes the material behavior beyond the initial yield. The initial yield provides the basis. As long as the initial yield property is defined, the remaining task is to define the different hardening/softening properties and to incorporate them into the initial yield provided. In this note, the yield criteria will be confined to the initial yield properties unless otherwise stated.

Generally speaking, yield criteria can be classified into two categories. The first category has originated from the concept of a single shear stress yield criterion such as the Tresca and the Mises. The Tresca is a maximum principal shear stress $\tau_{\max}$ criterion. The Mises is an octahedral shear stress τ_{oct} criterion. They both postulate that when $\tau_{\max}$ (or τ_{oct}) reaches a critical value, the material begins to yield. Both the Tresca and the Mises are applicable only to materials with (i) equal tensile and compressive strengths $\sigma_t=\sigma_c$; (ii) linear failure loci on the meridian plane (i.e., parallel to the hydrostatic axis which is represented by the line $\sigma_1=\sigma_2=\sigma_3$ in the principal stress space). Many derivatives of these two criteria have been proposed to model a wider range of engineering materials. Among them, the well-known ones are Mohr-Coulomb, Hoek-Brown, and Johnson (Tresca's derivatives), Matsuoka and Lade (Mises' derivatives) [1]. The Mohr-Coulomb is a $\tau_{\max}-\sigma_\tau$ criterion which considers the effect of $\tau_{\max}$ as well as the normal stress σ_τ on the same plane. The introduction of σ_τ makes it capable of simulating materials with different tensile and compressive strengths $\sigma_t\neq\sigma_c$ (called the strength differential effect or SD effect). The Hoek-Brown and Johnson are Mohr-Coulomb's derivatives for rock. They take into account the effect of nonlinear failure loci on the meridian plane while the Mohr-Coulomb does not. The Matsuoka and Lade are failure theories for soil (sand). They account for both the SD effect and the nonlinear failure loci on the meridian plane.

Unfortunately, recent research revealed that those theories do not necessarily represent the real failure/yield of materials under complex stress state. One prominent feature has been ignored, i.e., the effect of the intermediate principal stress (σ_2). The Tresca and its derivatives ignore this effect whereas the Mises and its derivatives average the effects of all the three principal stresses $\sigma_1,\sigma_2,\sigma_3$. Experiments show that the σ_2 effect varies from case to case ([2–5]). The extent of the σ_2 effect depends on the material type and the stress state.

The second category of yield criteria may be called curve-fitting multiparameter criteria such as the Argyris-Gudehus-Zienkiewicz criterion, the Willam-Warnke criterion, and some smooth models ([1]). These criteria usually have complex mathematical expressions because simple expressions usually cannot reflect the diversity of test results. The main advantage is that they can simulate accurately the yield properties in the particular range of complex stress state where most tests are conducted. Another deficiency is that they have little physical background.

2 Twin Shear Unified Strength Criterion (TS-Unified)

2.1 The TS-Unified Criterion Bridges Most Available Theories on π Plane. The TS-unified is an extension of the twin shear yield criterion (TS) ([6]) and the generalized twin shear yield criterion (GTS) ([7]). The TS is based on a kind of orthogonal dodecahedron (OD) stress element and it assumes that when a function of the two larger principal shear stresses (τ_{13},τ_{12}) or (τ_{13},τ_{23}) reaches a critical value, the material begins to yield, i.e.,

$$\begin{aligned}&\tau_{13}+\tau_{12}=\sigma_t \quad \text{when} \quad \tau_{12}\geqslant\tau_{23}\\ &\tau_{13}+\tau_{23}=\sigma_t \quad \text{when} \quad \tau_{12}\leqslant\tau_{23}\end{aligned} \tag{1}$$

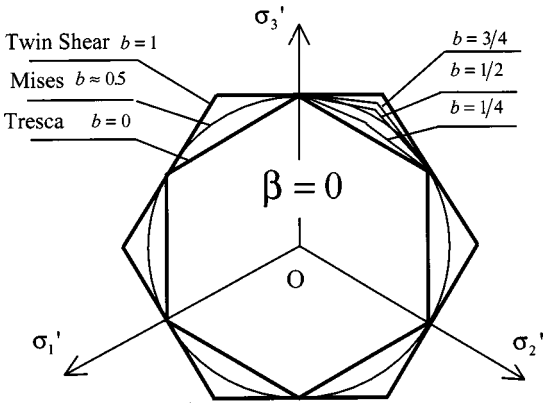


Fig. 1 Different yield loci on the π-plane for non-SD materials ($\sigma_t=\sigma_c$ or $\beta=0$)

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, Feb. 16, 1999; final revision, Nov. 1, 2000. Associate Editor: M.-J. Pindera.

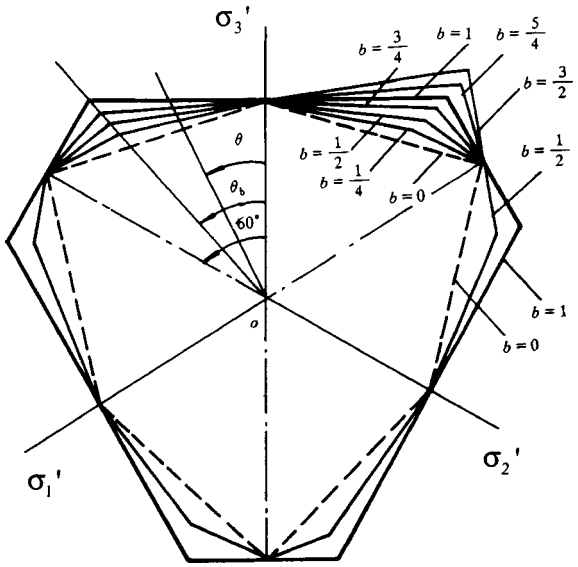


Fig. 2 Different yield loci on the π-plane for SD materials ($\sigma_t \neq \sigma_c$ or $\beta \neq 0$)

where $\tau_{13}=(\sigma_1-\sigma_3)/2$, $\tau_{23}=(\sigma_2-\sigma_3)/2$, $\tau_{12}=(\sigma_1-\sigma_2)/2$ are the principal shear stresses. σ_1, σ_2 and σ_3 ($\sigma_1 \geqslant \sigma_2 \geqslant \sigma_3$) are the principal normal stresses. σ_t is the uniaxial tensile strength.

The TS is an upper limit surface in the stress space complying with Drucker's convexity postulate and is used for non-SD materials ($\sigma_t=\sigma_c$, where σ_c is the uniaxial compressive strength) only. It is the counterpart to the Tresca, as shown in Fig. 1. The GTS extends the idea and assumes that the yield surface is a function of two larger principal shear stresses (τ_{13}, τ_{12}) or (τ_{13}, τ_{23}) and their corresponding normal stresses (σ_{13}, σ_{12}) or (σ_{13}, σ_{23}), i.e.,

$$\tau_{13}+\tau_{12}+\beta(\sigma_{13}+\sigma_{12})=\frac{2\sigma_t\sigma_c}{\sigma_c+\sigma_t} \quad \text{when} \quad \tau_{12}+\beta\sigma_{12} \geqslant \tau_{23}+\beta\sigma_{23}$$
$$\tau_{13}+\tau_{23}+\beta(\sigma_{13}+\sigma_{23})=\frac{2\sigma_t\sigma_c}{\sigma_c+\sigma_t} \quad \text{when} \quad \tau_{12}+\beta\sigma_{12} \leqslant \tau_{23}+\beta\sigma_{23}. \tag{2}$$

where $\sigma_{13}=(\sigma_1+\sigma_3)/2$, $\sigma_{12}=(\sigma_1+\sigma_2)/2$, $\sigma_{23}=(\sigma_2+\sigma_3)/2$. The parameter β reflects the effects of the normal stresses so that SD effect can be represented. Likewise, the GTS serves as the counterpart to the Mohr-Coulomb, as shown in Fig. 2. When $\beta=0$, the GTS is simplified to the TS. So the TS is a special case of the GTS.

In the TS and GTS, the largest stresses (both the shear and the normal stresses) have the same extent of influence as that of the second largest. When different weight parameters are employed to reflect the different effects of the largest stresses and the second largest, the GTS can be generalized to the TS-unified. The TS-unified can be expressed as

$$\tau_{13}+b\tau_{12}+\beta(\sigma_{13}+b\sigma_{12})=C \quad \text{when} \quad \tau_{12}+\beta\sigma_{12} \geqslant \tau_{23}+\beta\sigma_{23}$$
$$\tau_{13}+b\tau_{23}+\beta(\sigma_{13}+b\sigma_{23})=C \quad \text{when} \quad \tau_{12}+\beta\sigma_{12} \leqslant \tau_{23}+\beta\sigma_{23} \tag{3}$$

where b is a material parameter which represents the effect of the intermediate (the second largest) principal stresses. The value of b can be determined by material tests. β and C are also material parameters. If uniaxial tensile and compressive strength (σ_t and σ_c) are chosen as the basic test points, then β and C can be expressed as

$$\beta=\frac{\sigma_c-\sigma_t}{\sigma_c+\sigma_t}=\frac{1-\alpha}{1+\alpha} \quad C=\frac{(1+b)\sigma_t\sigma_c}{\sigma_c+\sigma_t}=\frac{1+b}{1+\alpha}\sigma_t. \tag{4}$$

where $\alpha=\sigma_t/\sigma_c$ is the ratio of the tensile to the compressive strengths. The ratio is an index of the material strength differential effect (SD effect).

The TS-unified is a series of piecewise linear yield criteria on the π-plane as shown in Figs. 1 and 2. The exact form of expression depends on the choice of parameter b, which in turn can be determined by some basic test results as illustrated in Section 2.3 of this note. The TS-unified has the following characteristics as shown in Table 1:

(a) With different choices of parameter b, the TS-unified can be simplified to the Tresca ($\beta=0$ and $b=0$), the linear approximations of Mises ($\beta=0$ and $b=1/2$ or $\beta=0$ and $b=1/(1+\sqrt{3})$), the Mohr-Coulomb ($\beta\neq 0$ and $b=0$), the TS ($\beta=0$ and $b=1$), the GTS ($\beta\neq 0$ and $b=1$) and a series of new strength criteria (other values of parameters β and b).

(b) In the stress space, the lower and upper bounds for the yield surfaces on the π-plane are special cases of the TS-unified, i.e., $b=0$ ($\beta=0$ for the Tresca or $\beta\neq 0$ for the Mohr-Coulomb) and $b=1$ ($\beta=0$ for the TS or $\beta\neq 0$ for the GTS), respectively.

(c) When the parameter b varies between 0 and 1, a series of yield surfaces between the two limiting surfaces can be obtained.

(d) When the parameter b varies beyond the range (i.e., $b<0$ or $b>1$), a series of nonconvex limit surfaces could be derived.

2.2 The TS-Unified Reflects Different σ_2 Effects. The Lode angle θ_σ is a parameter to represent the relative value of σ_2

Table 1 The characteristics of TS-unified

	$\beta=0$ ($\sigma_t=\sigma_c=\sigma_0$)	$\beta\neq 0$ ($\sigma_t\neq\sigma_c$)	Drucker's Convex Postulate
$b=0$ π plane loci	Tresca Regular hexagon	Mohr-Coulomb Irregular hexagon	Lower bound
Restriction for application	$\tau_0=\frac{\sigma_0}{2}$	$\tau_0=\frac{\sigma_t\sigma_c}{\sigma_c+\sigma_t}$	
$0<b<1$ π plane loci	New theories Irregular dodecagon	New theories Irregular dodecagon	Intermediate
Restriction for applications	$\tau_0=\frac{(b+1)}{(2+b)}\sigma_0$	$\tau_0=\frac{(b+1)\sigma_t\sigma_c}{\sigma_t+(b+1)\sigma_c}$	
$b=1$ π plane loci	Twin shear theory Regular hexagon	Twin shear theory Irregular hexagon	Upper bound
Restriction for applications	$\tau_0=\frac{2}{3}\sigma_0$	$\tau_0=\frac{2\sigma_t\sigma_c}{\sigma_t+2\sigma_c}$	

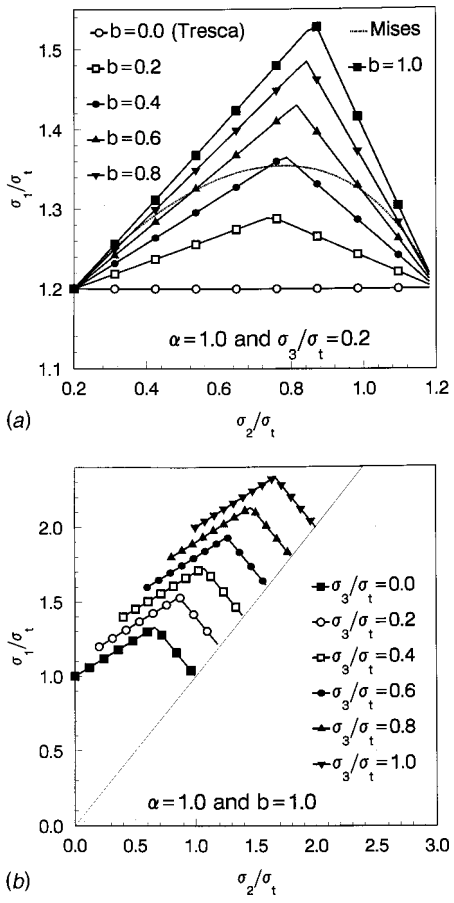


Fig. 3 TS-unified can reflect the σ_2 effect in a piecewise linear manner; (*a*) α=1 and σ_3/σ_t=0.2 case; (*b*) α=1 and b=1 case

with respect to σ_1 and σ_3. To reflect the σ_2 effect, a theory should embed this parameter in its expression either in explicit or in implicit forms.

The TS-unified is usually expressed in terms of the stress (deviatoric) invariants, that is the first invariant of stress tensor I_1, the second invariant of stress deviatoric tensor J_2, and the stress angle θ, as follows:

$$\frac{I_1}{3}(1-\alpha)+\left(1+\frac{\alpha}{2}\right)\frac{2\sqrt{J_2}}{\sqrt{3}}\cos\theta-\frac{\alpha(1-b)}{1+b}\sqrt{J_2}\sin\theta = \sigma_t \quad \text{when} \quad 0\ \text{deg}\leqslant\theta\leqslant\theta_b$$

$$\frac{I_1}{3}(1-\alpha)+\left(\frac{2-b}{1+b}+\alpha\right)\frac{\sqrt{J_2}}{\sqrt{3}}\cos\theta-\left(\alpha+\frac{b}{1+b}\right)\sqrt{J_2}\sin\theta = \sigma_t \quad \text{when} \quad \theta_b\leqslant\theta\leqslant 60\ \text{deg} \tag{5}$$

where θ_b is the stress angle when the two Eqs. (5) are equal:

$$\theta_b=\text{tg}^{-1}\frac{\sqrt{3}}{1+2\alpha}, \qquad (0\ \text{deg}<\theta_b<60\ \text{deg}). \tag{6}$$

The stress angle θ is defined to reflect the σ_2 effect as shown in Fig. 2, such that

$$\theta=\frac{1}{3}\cos^{-1}\frac{3\sqrt{3}}{2}\frac{J_3}{\sqrt{J_2^3}} \qquad (0\leqslant\theta\leqslant\pi/3). \tag{7}$$

It should be noted that the stress angle θ is different from the Lode angle θ_σ, and they are related as follows:

$$\theta=\frac{\pi}{6}-\theta_\sigma. \tag{8}$$

Equation (5) is the explicit expression for the TS-unified in terms of the stress angle θ. With different choices of the parameter b, the TS-unified can reflect different piecewise linear σ_2 effects. Two illustrations are given. Figure 3(*a*) shows the curves of σ_1/σ_t versus σ_2/σ_t for $\alpha=1$ and $\sigma_3/\sigma_t=0.2$. Obvious strength difference can be observed when different criteria are adopted (represented by different values of the parameter b). The maximum difference is about 25 percent. Figure 3(*b*) shows the curves for $\alpha=1$ and $b=1$. The same conclusion can be drawn as from Fig. 3(*a*). In other words, the TS-unified is capable of representing variable strengths of the same material under different stress states.

2.3 Application of the TS-Unified. If σ_t, σ_c and the shear strength τ_0 are chosen as the basic material parameters, through Eq. (3) for pure shear loading, the parameter b can be expressed as

$$b=\frac{(\sigma_c+\sigma_t)\tau_0-\sigma_t\sigma_c}{(\sigma_t-\tau_0)\sigma_c}. \tag{9}$$

The parameter b plays an important role in the TS-unified. It builds a bridge among different strength theories. It is this parameter that distinguishes one theory from another. On the other hand, the scope of application of each theory is also represented by this parameter. Hence, the TS-unified is a unified theory that can be applied to more than one kind of material. In practice, when basic material parameters are obtained by experiments, the value of b can be determined through Eq. (9). Whenever parameter b is obtained, the yield criterion is determined and the application is possible ([8]).

3 Conclusions

A piecewise linear unified yield criterion called the TS-unified was proposed. Besides the capability that the TS-unified can bridge most available yield loci on the π-plane for both SD and non-SD materials, the most prominent characteristics of the criterion are their capability to represent the effects of the intermediate principal stress σ_2 in piecewise linear forms. Illustrations were given. The determination of the parameters was also discussed.

Future research could focus on the pursuit for some criteria that can bridge different criteria both on the π-plane and on the meridian plane. It is obvious that the representation for nonlinear meridian loci can be obtained by adopting multiparameter criterion instead of the proposed two-parameter TS-unified.

References

[1] Chen, W. F., and Han, D. J., 1988, *Plasticity for Structural Engineers*, Springer-Verlag, New York.

[2] Mogi, K., 1967, ''Effect of Intermediate Principal Stress on Rock Failure,'' J. Geophys. Res., **72**, pp. 5117–5131.

[3] Michelis, P., 1987, ''True Triaxial Cyclic Behavior of Concrete and Rock in Compression,'' Int. J. Plast., **3**, No. 2, pp. 249–270.

[4] Faruque, M. O., and Chang, C. J., 1990, ''A Constitutive Model for Pressure Sensitive Materials With Particular Reference to Plain Concrete,'' Int. J. Plast., **6**, No. 1, pp. 29–43.

[5] Li, X. C., and Xu, D. J., 1990, ''Experimental Verification of Twin Shear Strength Theory—The Strength Properties of Granite Under True Triaxial Stress State,'' Wuhan Institute of Rock and Soil Mechanics, Chinese Academy of Science, Paper Yantu (90) 52 (in Chinese).

[6] Yu, M. H., 1983, ''Twin Shear Stress Yield Criterion,'' Int. J. Mech. Sci., **25**, No. 1, pp. 71–74.

[7] Yu, M. H., He, L. N., and Song, L. Y., 1985, ''Twin Shear Stress Theory and Its Generalization,'' Sci. Sinica A, **28**, No. 1, pp. 1174–1183.

[8] Yu, M. H., Yang, S.-Y., and Fan, S. C., 1999, ''Unified Elasto-Plastic Associated and Non-Associated Constitutive Model,'' Comput. Struct., **71**, No. 6, pp. 627–636.

Analytical Solution for W-N Criteria for the Prediction of Notched Strength of an Orthotropic Shell

R. Ramesh Kumar
Engineer, Structural Design and Analysis Division, Structural Engineering Group, Vikram Sarabhai Space Centre, Thiruvananthapuram 695 022, India

S. Jose
Senior Lecturer, Department of Mechanical Engineering, T.K.M. College of Engineering, Kollam 691 005, India

G. Venkateswara Rao
Group Director, Structural Design and Analysis Division, Structural Engineering Group, Vikram Sarabhai Space Centre, Thiruvananthapuram 695 022, India

Analytical solution for the tangential stress distribution ahead of a hole is needed for the theoretical prediction of notched strength of brittle laminate using the well-known W-N criteria. In the present study, tangential stress distribution in an orthotropic circular cylindrical shell under uniaxial loading with a circular hole is obtained intuitively with the use of a stress function. A good agreement is obtained for the stresses around and ahead of the circular hole in $(0\,deg_4 \pm 30\,deg)_s$ *and* $90\,deg$ *laminates with the finite element results.* [DOI: 10.1115/1.1320452]

Introduction

Prediction of failure strength of brittle laminate with a hole was very well established based on the W-N fracture criteria ([1–4]). Failure of the laminate occurs when the dominant stress near the hole or the average of the dominant stress over a region near the hole reaches the strength of the laminate. Konish and Whitney [5] developed an analytical solution for the stress distribution near a hole in an orthotropic plate. For the shell-type composite structures one has to essentially go for the finite element approach as there is no such analytical solution available in the literature for employing the W-N criteria. For the best numerical results a clear understanding of the maximum stress location is very essential. Unlike in metallic structures, for fiber-reinforced orthotropic shells, the maximum stress does not occur at the hole edge in a plane normal to the loading direction, but depends on the fiber orientations ([5]). Moreover, establishing convergence for the finite element model for the isotropic medium for a known problem does not ensure convergence for the orthotropic medium. Hence an analytical solution that can bring out the overall behavior of the orthotropic structure is very much required.

Savin [6] in his complex variable approach to the problem of stresses in an isotropic circular cylindrical shell with holes showed that the stress state in a shell is a sum of a plate solution and a function of curvature effect. Konish and Whitney's [5] and Kumar, Rao, and Mathew's [7] orthotropic equations (plate) were based on the sum of an isotropic plate solution and a function of higher order term of distance ahead of the hole in terms of orthotropic material constants. Thus one can conjecture that the orthotropic shell solution in a polar coordinate system (ρ,θ) can be obtained as a sum of an isotropic plate equation (first term in (1)), higher order term of distance ahead of a hole (third term in (1)), and functions of isotropic and orthotropic curvature effects (second and fourth terms in (1)).

$$\left.\frac{\sigma_\theta^{\mathrm{Sh}}(\rho,\theta)}{\sigma}\right|_{\mathrm{Ortho}} = \left.\frac{\sigma_\theta^{\mathrm{Pl}}(\rho,\theta)}{\sigma}\right|_{\mathrm{ISO}} + f(\beta)_{\mathrm{ISO}} + f\begin{pmatrix}\text{higher order terms in } \rho \text{ with}\\ \text{orthotropic coefficients}\end{pmatrix} + f\begin{pmatrix}\text{higher order terms in } \rho \text{ with}\\ \text{orthotropic coefficients for } \beta\end{pmatrix} \quad (1)$$

In the present work an analytical solution for the tangential stress distribution ahead of a circular hole in an orthotropic circular cylindrical shell under axial loading is obtained for use via the W-N criteria.

Analytical Solution for the Tangential Stress Distribution

The most general solution for the stress distribution near cutouts of any arbitrary shape under tension or internal pressures up to an accuracy of β^2 was given by Pirogov [8]. The equation with unknown coefficients (like A_0, B_0, E_K, F_K, etc.) is of the form

$$\mathrm{Im}(\phi_j) = \frac{1}{\pi}\left\{\begin{array}{l}\left[2A_0^{(j)}\ln\left(\dfrac{\gamma\beta}{\sqrt{2}}\right)+\left(A_2^{(j)}+\dfrac{B_2^{(j)}}{\rho^2}\right)\cos 2\theta\right]+\displaystyle\sum_{k=4,6,\ldots}^{\infty}\left(\frac{k-2}{2\rho^{k-2}}A_k^{(j)}+\frac{B_k^{(j)}}{\rho^k}\right)\cos k\theta \\ +\beta^2\left\{\begin{array}{l}2B_0^{(j)}\ln\left(\dfrac{\gamma\beta}{\sqrt{2}}\right)-\dfrac{\pi}{4}(2A_0^{(j)}+A_2^{(j)})\rho^2+\displaystyle\sum_{k=4,6,\ldots}^{\infty}\left(\frac{E_k^{(j)}}{\rho^{k-2}}+\frac{F_k^{(j)}}{\rho^k}\right)\cos k\theta \\ +\left[\dfrac{-\pi}{4}(A_0^{(j)}+A_2^{(j)})\rho^2+E_2^{(j)}+\dfrac{F_2^{(j)}}{\rho^2}\right]\cos 2\theta.\end{array}\right.\end{array}\right\} \quad (2)$$

Based on the above stress function, Savin obtained the tangential stress distribution around a circular hole of radius ''*a*'' in an isotropic shell of radius ''*R*'' and thickness ''*t*,'' under axial loading, as a sum of an isotropic plate solution $(\sigma_\theta^{\mathrm{Pl}}(\rho,\theta)/\sigma)$ and a function of isotropic curvature parameter, β. The solution is valid for a hole whose projected size (on a plane passing through the axis of shell and normal to the axis of hole) is close to the actual hole size.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 4, 1999; final revision, May 25, 2000. Associate Editor: J. W. Ju.

$$\left.\frac{\sigma_\theta^{\mathrm{Sh}}(\rho,\theta)}{\sigma}\right|_{\mathrm{ISO}} = \frac{1-\cos 2\theta}{2} + \frac{1}{2\rho^2} - \frac{3}{2\rho^4}\cos 2\theta \ (=\text{isotropic plate solution}) - \frac{\pi\beta^2}{8}\left(1+\frac{3}{\rho^4}\right)\cos 2\theta \ (=\text{function in isotropic } \beta,\ f(\beta)_{\mathrm{ISO}}). \quad (3)$$

Similarly for an orthotropic shell,

$$\left.\frac{\sigma_\theta^{\mathrm{Sh}}(\rho,\theta)}{\sigma}\right|_{\mathrm{Ortho}} = \left.\frac{\sigma_\theta^{\mathrm{Pl}}(\rho,\theta)}{\sigma}\right|_{\mathrm{Ortho}} + f(\beta)_{\mathrm{Ortho}} \quad (4)$$

where $\beta_{\mathrm{Ortho}}=[\sqrt[4]{3(1-\nu^2)(E_\theta/E_Y)}/2]a/\sqrt{Rt}$, ν=Poisson's ratio, E_θ and E_Y are the moduli of elasticity in θ and Y (loading direction) directions ([9]), and $(\sigma_\theta^{\mathrm{Pl}}(\rho,\theta)/\sigma)$ for the orthotropic plate is available in Kumar et al. [7].

$$\left.\frac{\sigma_\theta^{\mathrm{Pl}}(\rho,\theta)}{\sigma}\right|_{\mathrm{Ortho}} = \frac{1-\mathrm{Cos}\,2\theta}{2} + \frac{1}{2\rho^2} - \frac{3}{2\rho^4}\mathrm{Cos}\,2\theta + \sum_{j=2}^{\infty}\frac{-C_{2j}}{2}\left[\frac{4j-3}{\rho^{4j-2}} - \frac{4j-1}{\rho^{4j}}\right]\mathrm{Cos}\,2j\theta \quad (5)$$

Details of the orthotropic coefficients C_{2j} (j varying from 1 to ∞) are given in the Appendix.

It can be noticed from Pirogov's stress function given in (1) that the higher order terms for the plate solution and β are of similar trigonometric relationship. Therefore, the functions of β may be expressed as

$$f(\beta)_{\mathrm{Ortho}} = -\frac{\pi\beta^2}{2}\left\{\frac{1}{4}\left(1+\frac{3}{\rho^4}\right)\mathrm{Cos}\,2\theta + \sum_{j=2}^{\infty}\frac{C_{2j}}{2}\left[\frac{4j-3}{\rho^{4j-2}} - \frac{4j-1}{\rho^{4j}}\right]\mathrm{Cos}\,2j\theta\right\}. \quad (6)$$

Thus the final expression for the tangential stress distribution for an orthotropic circular cylindrical shell with a circular hole is given by

$$\left.\frac{\sigma_\theta^{\mathrm{Sh}}(\rho,\theta)}{\sigma}\right|_{\mathrm{Ortho}} = \frac{1-\cos 2\theta}{2} + \frac{1}{2\rho^2} - \frac{3}{2\rho^4}\cos 2\theta - \frac{\pi\beta^2}{8} \times\left\{\left(1+\frac{3}{\rho^4}\right)\cos 2\theta\right\} + \left(1+\frac{\pi\beta^2}{2}\right) \times\left\{\sum_{j=2}^{\infty}\frac{-C_{2j}}{2}\left\{\frac{4j-3}{\rho^{4j-2}} - \frac{4j-1}{\rho^{4j}}\right\}\cos 2j\theta\right\}. \quad (7)$$

In this work, circular cylindrical shells with $(0_4,\pm 30)_S$ and $(90)_{12}$ lay ups made of high modulus M55J/M18 carbon/epoxy laminate having material properties $E_X=328.949$ GPa, $E_Y=5.955$ GPa, $G_{XY}=4.414$ GPa and $\nu_{XY}=0.346$, with layer thickness of 0.1 mm is considered.

Finite Element Modeling

The laminated circular cylindrical shell ($R=48$ mm, $t=1.2$ mm and height=180 mm) with a hole of radius $a=5$ mm is modeled using eight-noded layered shell elements available in NISA2 finite element software. Due to the symmetry, one fourth of the shell (at its half height) is modeled. The region around the hole is modeled with a finer mesh with an element size of one tenth of the hole radius up to $\rho=0.5$, then one fifth up to $\rho=1$ and then one fourth up to $\rho=2$ in the radial direction, while in θ direction, it is at every five-degree interval along the circumference of the hole from 0 deg to 180 deg.

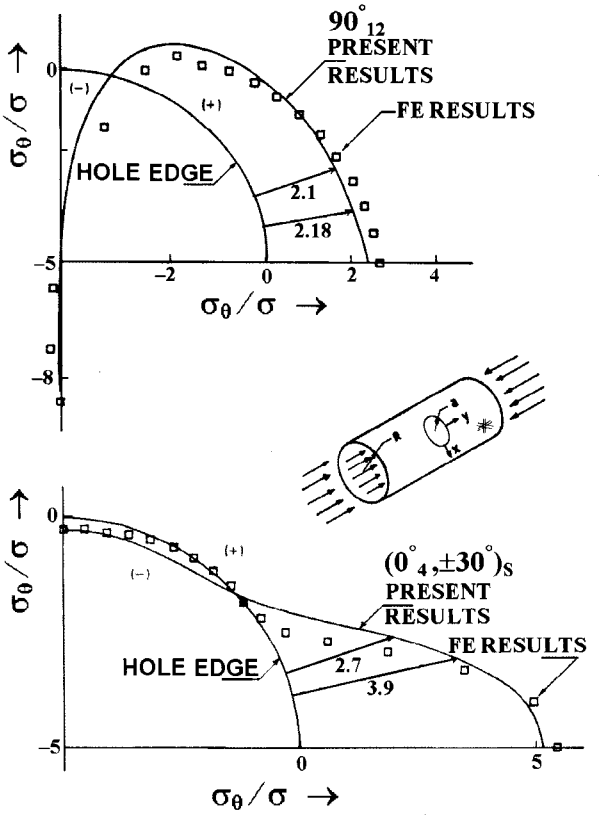


Fig. 1 Tangential stress distribution around a circular hole in an orthotropic shell

Boundary Conditions. Symmetric boundary conditions are applied at the symmetric planes.

Loading. A distributed load of 240 N is applied on the nodes at the top circumference. The load is distributed in each element in the ratio of 1:4:1 among the nodes in an element.

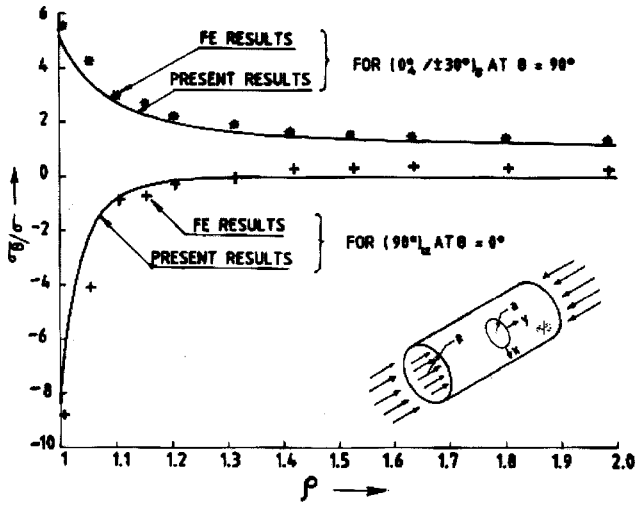


Fig. 2 Tangential stress distribution ahead of a circular hole in an orthotropic shell

Results and Discussion

Initially, convergence of (7) is established by progressively increasing the number of terms till a converged value is obtained. For the present cases terms up to C_{22} ($j=11$) are needed to achieve convergence within one percent. It may be noted that for $\beta=0$, (7) reduces to the orthotropic plate solution.

Using (7), the tangential stress distribution around a hole ($\rho=1$) is obtained and compared with the finite element results as shown in Fig. 1. The maximum stress concentration factor of -8.56 is observed at $\theta=0$ for $(90)_{12}$ laminate, as against the finite element result of -8.71. The deviation in the results is estimated as two percent.

In the case of $(0_4, \pm 30)_S$ lay up, the peek stress occurs at $\theta=90$ deg as expected, with a stress concentration factor value equal to 5.18 as against the finite element result of 5.44. It can be noticed that the deviation between the two values is about five percent. The stress distributions away from the hole for $\rho=1$ to 2, for the two types of lay-up sequence considered are shown in Fig. 2. From the finite element analysis, it is found that as a percentage of total stress, bending stress constitute within two percent for the types of lay-up sequences considered. It is concluded that the present solution shows a reasonably good agreement with the finite element results.

Conclusion

A new analytical solution for tangential stress distribution for an orthotropic circular cylindrical shell with a circular hole under axial loading is derived which gives good agreement with finite element results. The solution can be used with W-N criteria for the prediction of notched strength of an orthotropic shell with a circular hole.

Appendix

There is a standard technique ([10]) for determining the Fourier coefficients C_{2j} of a function as in (7):

$$A_0C_0+A_2C_2+A_4C_4=2a_0 \tag{A1a}$$

$$A_2C_0+(A_4+2A_0+)C_2+A_2C_4+A_4C_6=2a_2 \tag{A1b}$$

and for $j>3$ a recursive relationship exists in the form

$$C_{2j}=\frac{-1}{A_4}[A_2C_{2j-2}+2A_0C_{2j-4}+A_2C_{2j-6}+A_4C_{2j-8}] \tag{A1c}$$

where $a_0=4k(n+k-1)$; $a_2=-4k(n+k+1)$; $A_0=3+n^2-2k+3k^2$

$$A_2=4(1-k^2); \quad A_4=(1-n^2+2k+k^2) \tag{A2}$$

$$n=N/K, \quad N=\sqrt{2(K-\overline{\nu}_{XY})-\overline{E}_X/\overline{G}_{XY}}$$

$$k=1/K, \quad K=\sqrt{\overline{E}_X/\overline{E}_Y}$$

$\overline{E}_X, \overline{E}_Y, \overline{\nu}_{XY}, \overline{G}_{XY}$ are the overall orthotropic properties of the shell.

References

[1] Whitney, J. M., and Nuismer, R. J., 1974, ''Stress Fracture Criteria for Laminated Composites Containing Stress Concentrations,'' J. Compos. Mater., **8**, pp. 253–265.

[2] Rhodes, M. D., Mikulas, M. M., and Mcgowen, P. E., 1984, ''Effect of Orthotropy and Width on Compression Strength of Graphite-Epoxy Panels With Holes,'' AIAA J., **22**, pp. 1283–1292.

[3] Awerbuch, J., and Madhukar, M. S., 1985, ''Notched Strength of Composite Laminates: Prediction and Experiments—A Review,'' J. Reinf. Plast. Compos., **4**, pp. 3–159.

[4] Kumar, R. R., 1993, ''Experimental Investigation and Prediction on Notched Strength of Laminated Composite Circular Cylindrical Shells,'' DLR report IB 131-93/16, Germany.

[5] Konish, H. J., and Whitney, J. M., 1975, ''Approximate Stresses in an Orthotropic Plate Containing a Circular Hole,'' J. Compos. Mater., **9**, pp. 157–166.

[6] Savin, G. N., 1970, ''Stress Distribution Around Holes,'' NASA Technical Translation, NASA TTF-607.

[7] Kumar, R. R., Rao, G. V., and Mathew, K. J., 1999, ''A New Solution for Prediction of Notched Strength for Highly Orthotropic Plate,'' J. Aeronaut. Soc. India, **51**, No. 1, pp. 28–34.

[8] Pirogov, I. M., 1962, ''On Approximate Solution of Basic Differential Equations in the Theory of Shells,'' Izv. Vuzov. Mash, **11**.

[9] Lekhnitskii, S. G., 1947, *Anisotropic Plates*, Gordon and Breach, New York.

[10] Krylov, V. I., and Kruglikova, L. G., 1969, *Handbook of Numerical Harmonic Analysis*, Israel Program for Scientific Translations, Jerusalem.

Stress Wave Propagation in a Coated Elastic Half-Space due to Water Drop Impact

Hyun-Sil Kim[1]
e-mail: hskim@kimm.re.ir

Jae-Seung Kim
Mem. ASME

Hyun-Ju Kang

Sang-Ryul Kim

Acoustics Lab, Korea Institute of Machinery and Materials, Yusong, Taejon 305-600, Korea

Stress wave propagation in a coated elastic half-space due to water drop impact is studied by using the Cagniard-de Hoop method. The stresses have singularity at the Rayleigh wavefront whose location and singular behavior are determined from the pressure model and independent of the coating thickness, while reflected waves cause minor changes in amplitudes.

[DOI: 10.1115/1.1352060]

Introduction

High-speed impact of liquid drops has been known to cause damage and erosion of the structures such as steam turbine blades, skin of the aircraft, and missiles. Evans et al. [1] used a FDM in studying motion of the elastic bodies, where their pressure loading was obtained as if the water drop collided against a solid surface. Adler [2] performed a more comprehensive three-dimensional finite element method analysis by allowing interactions between a water drop and a deformable target. Blowers [3] studied stress wave propagation in an elastic half-space analytically by employing the Cagniard-de Hoop method ([4]). Although Blowers used a simplified pressure model, which is valid only in the early stage of the impact, his results provided important information about the role of the Rayleigh waves and later his method has been used extensively by others ([5]) to compute the stresses for various materials.

For damage and erosion protection, an idea of coating the surface with a thin elastic layer ([6]) has been frequently used. In order to select the proper coating material and thickness, it is essential to know the stresses inside the coating and the base material. In this paper, we study stress wave propagation in a coated elastic half-space analytically by using the generalized ray

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 27, 1999; final revision, July 21, 2000. Associate Editor: A. K. Mal.

method ([4]). We use the same pressure model as Blowers, which means that our solution is useful only in early stage of the impact process before lateral outflow jetting takes place. However, the results are of great importance, since high stresses and possible damages may occur in a very early stage of the impact.

Theory

As shown in Fig. 1, we consider a coated elastic half-space ($z>0$), where a thin elastic layer of uniform thickness h lies over the surface. On the surface, the stress generated by a water drop impact is given in a cylindrical coordinate by

$$\sigma_{zz}(r,z,t)=-P, \quad r<k_0\sqrt{t}, \quad \sigma_{zz}(r,z,t)=0, \quad r>k_0\sqrt{t}, \tag{1}$$

where P is a constant pressure and k_0 is a constant determined from the diameter and impact velocity of the water drop ([5]).

The potential functions in the coating and half-space satisfy the wave equations

$$\nabla^2\phi_i=\frac{1}{\alpha_i^2}\frac{\partial^2\phi_i}{\partial t^2}, \quad \nabla^2\psi_i=\frac{1}{\beta_i^2}\frac{\partial^2\psi_i}{\partial t^2}, \tag{2}$$

where the subscript $i=1$ is for coating, while $i=2$ for the half-space. The parameters α_i and β_i are the sound speeds of P and S waves in the ith medium. We nondimensionalize the parameters in a similar way as Blowers [3] did, and transform the wave equations by applying the Laplace and Hankel transform with respect to the nondimensionalized time T and radial distance R. In the coating, the solution consists of upgoing and downgoing waves due to the reflections at the interface and free surface, while in the half-space the solution contains only downgoing waves. The unknown coefficients can be determined from the boundary conditions.

We show, for instance, the Laplace transformed stress $\overline{S_{RR}^1}$ in the coating

$$\begin{aligned}\overline{S_{RR}^1}=\frac{2\,\mathrm{Re}}{\pi}\int_0^\infty\int_0^\infty\Bigg[&\left(\frac{\alpha_1^2}{\beta_1^2}-2-2w^2\right)\\&\times(D_\alpha^1 e^{-p\eta_{1\alpha}Z}+U_\alpha^1 e^{-p\eta_{1\alpha}(H-Z)})\\&+2w^2\left(\frac{\eta_{1\beta}}{\zeta}\right)(D_\beta^1 e^{-p\eta_{1\beta}Z}\\&-U_\beta^1 e^{-p\eta_{1\beta}(H-Z)})\Bigg]e^{iRpw}p^4\,dw\,dq,\end{aligned} \tag{3}$$

where $D_\alpha^1, D_\beta^1, U_\alpha^1, U_\beta^1$ are the coefficients to be determined from the boundary conditions, Z and H are the nondimensionalized depth and coating thickness, and p, ζ are the Laplace and Hankel transform variables. The parameters $\eta_{1\alpha}$ and $\eta_{1\beta}$ are $\eta_{1\alpha}=\sqrt{\zeta^2+1}$ and $\eta_{1\beta}=\sqrt{\zeta^2+\alpha_1^2/\beta_1^2}$.

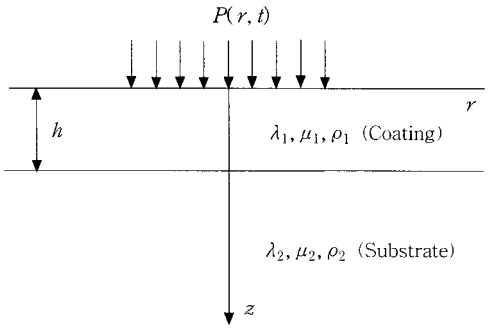


Fig. 1 A coated elastic half-space subject to water drop impact

It can be shown that the coefficients $D_\alpha^1, D_\beta^1, U_\alpha^1, U_\beta^1$ are expressed in terms of the infinite series, whose physical meaning is that the solution in the coating is composed of the reflected waves and the number of reflections can increase infinitely. The typical form $\overline{I_k}$ of the infinite series in the stress $\overline{S_{RR}^1}$ can be expressed as

$$\overline{I_k}=\mathrm{Re}\int_0^\infty\int_0^\infty R_k(w,q)e^{-pg_k(w,q)}\,dw\,dq, \tag{4}$$

where $R_k(w,q)$ and $g_k(w,q)$ are independent of the variable p, and g_k is of the form

$$g_k(w,q)=w^2+q^2-iwR+\sum_n b_n\sqrt{w^2+q^2+\alpha_1^2/c_n^2}, \tag{5}$$

in which c_n is α_i or β_i.

In order to apply the Cagniard-de Hoop method, we deform the integration path such that $g_k(w,q)=T$. The new integration path $w=w(T,q)$ intersects the imaginary w-axis at $w=i\nu_m$. We perform the integration along the new path in T and change the order of integration to find the inverse Laplace transform as

$$I_k(T)=\mathrm{Re}\left[\int_0^{q_m}R_k(w,q)\frac{\partial w}{\partial T}\,dq\right]H(T-T_m(0)), \tag{6}$$

where q_m and T_m satisfy the relation, $T_m(q_m)=T$, and $H(T)$ denotes a Heaviside step function. After summation of $I_k(T)$ over the rays, we can compute the stresses. Depending on the relative positions of $w=i\nu_m$ and the branch points, we may have to introduce additional integration path to detour the branch cut, which leads to the head wave.

When $Z=0$, we need to include the Rayleigh surface wave

$$I_R(T)=D_R\frac{H(T+a^2-aR)}{\sqrt{(T+a^2)^2-a^2R^2}}, \tag{7}$$

where D_R is the coefficient associated with the residue term and $a=\alpha_1/c_R$, in which c_R is the Rayleigh wave speed in the coating. The surface wave in Eq. (7) has singularity at $R=(T+a^2)/a$.

Numerical Example

For a numerical example, we consider a case that the diameter and velocity of the water drop are $d_0=2$ mm, $V_0=453$ m/s, and the thickness of the coating is 43 μm. The material properties of the coating are: Young's modulus $E_1=1.71\times10^{11}$ N/m^2, α_1

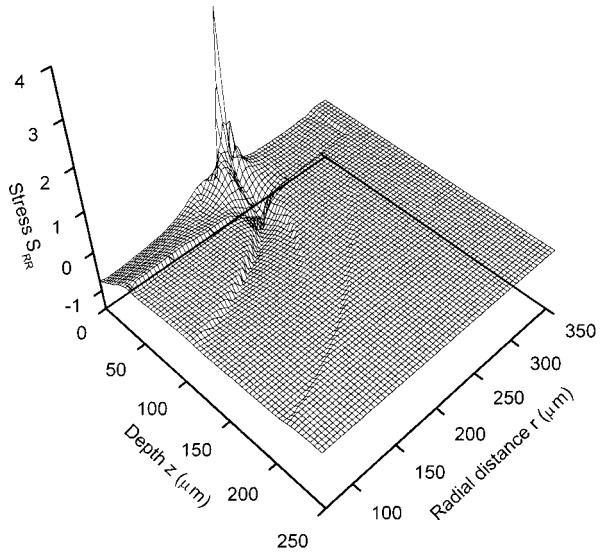


Fig. 2 Snapshot of the stress S_{RR} when $t=0.05\,\mu$s

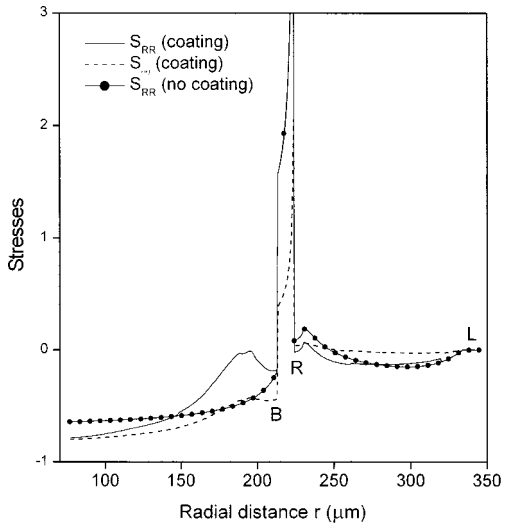


Fig. 3 Variation of the nondimensional stresses S_{RR} and $S_{\theta\theta}$ as functions of the radial distance r on the surface when $t=0.05\ \mu$s. Solid line with marks is S_{RR} when there is no coating and the substrate is filled with the same material as coating.

$=5910$ m/s, $\beta_1=3160$ m/s, density $\rho_1=6590$ kg/m^3; for substrate, $E_2=6.74\times10^{10}$ N/m^2, $\alpha_2=4150$ m/s, $\beta_2=2220$ m/s, $\rho_2=5270$ kg/m^3. The substrate material is Zinc-Selenide. In Blowers' paper, the only specified material property is Poisson's ratio ν, and we use the same value here as $\nu=0.3$ for both coating and substrate, for which case $a=2.017$.

In Fig. 2, we plot a snapshot of the nondimensional stress S_{RR} at 0.05 μs. There is a sharp peak near the surface, which corresponds to the Rayleigh wavefront at $r=224\ \mu$m $(R=(T+a^2)/a=5.84)$. The boundary of the impact is $r=213\ \mu$m $(R=2\sqrt{T}=5.56)$. In Fig. 3, we show the stresses S_{RR} and $S_{\theta\theta}$ at the surface as functions of the radial distance, in which the symbols ''B,'' ''R,'' and ''L'' mean the boundary of impact area, Rayleigh wavefront, and longitudinal wavefront, respectively. For comparison, we also show the stress S_{RR} when there is no coating and the half-space is filled with the same material as the coating.

Concluding Remarks

It was shown that the pressure model in Eq. (1) produces an annular strip of the high tensile stresses outside the contact area due to the Rayleigh wave, which has been observed experimentally by Hand et al. [7]. The location and singular behavior of the Rayleigh wavefront are determined from the pressure model and independent of the coating thickness. The region directly under the contact area is in pure compression. Since the stresses cannot have infinite magnitude in real impact situations, the singularity in the present study may be due to the abrupt change of pressure model at the impact boundary.

Acknowledgment

The authors would like to thank the Ministry of Science and Technology, Korea, for the financial support by a grant from the Critical Technology 21 Project.

References

[1] Evans, A. G., Ito, Y. M., and Rosenblatt, M., 1980, ''Impact Damage Thresholds in Brittle Materials Impacted by Water Drops,'' J. Appl. Phys., **51**, No. 5, pp. 2473–2482.
[2] Adler, W. F., 1995, ''Waterdrop Impact Modeling,'' Wear, **186–187**, pp. 341–351.
[3] Blowers, R. M., 1969, ''On the Response of an Elastic Solids to Droplet Impact,'' J. Inst. Math. Appl., **5**, No. 2, pp. 167–193.
[4] van der Hijden, J. H., 1987, *Propagation of Transient Elastic Waves in Stratified Anisotropic Media*, North-Holland, Amsterdam.
[5] Adler, W. F., 1977, ''Liquid Drop Collisions on Deformable Media,'' J. Mater. Sci., **12**, pp. 1253–1271.
[6] Springer, G. S., 1976, *Erosion by Liquid Impact*, John Wiley and Sons, New York.
[7] Hand, R. J., Field, J. E., and Townsend, D., 1991, ''The Use of Liquid Jets to Simulate Angled Drop Impact,'' J. Appl. Phys., **70**, No. 11, pp. 7111–7118.

Closed-Form Representation of Beam Response to Moving Line Loads

Lu Sun
Department of Civil Engineering, The University of Texas at Austin, ECJ Hall 6.10, Austin, TX 78712
e-mail: lusun@mail.utexas.edu

Fourier transform is used to solve the problem of steady-state response of a beam on an elastic Winkler foundation subject to a moving constant line load. Theorem of residue is employed to evaluate the convolution in terms of Green's function. A closed-form solution is presented with respect to distinct Mach numbers. It is found that the response of the beam goes to unbounded as the load travels with the critical velocity. The maximal displacement response appears exactly under the moving load and travels at the same speed with the moving load in the case of Mach numbers being less than unity. [DOI: 10.1115/1.1352064]

1 Introduction

The response of beams to moving loads has been studied extensively over the past several decades (Fryba [1]). The investigation of Bernoulli-Euler beams with moving loads includes the work of Kenney [2], Steele [3], Huang [4], Choros and Adams [5], Jezequel [6], Elattary [7], Lee [8], Sun and Deng [9], Sun [10], Sun and Greenberg [11], and Benedetti [12]. It is found that the moving load is often treated as a concentrated load. Since the concentrated loading condition is only an idealized model of the tire load, it is preferable to use a distributed line load model to characterize the wheel load more realistically.

Denote $y(x,t)$ as the deflection of the beam in y-direction, in which x represents the traveling direction of the pavement structure, and t represents time. The well-known governing equation of a Bernoulli-Euler beam on a Winkler foundation is (Sun [10])

$$EI\frac{\partial^4 y}{\partial x^4}+Ky+m\frac{\partial^2 y}{\partial t^2}=P\frac{H[r_0^2-(x-vt)^2]}{2r_0} \tag{1}$$

where EI is the rigidity of the beam, E is Young's modulus of elasticity, I is the moment of inertia of the beam, K is the modulus of subgrade reaction, m is the unit mass of the beam, r_0 is the half-width of the line load, P is the magnitude of the applied load, and $H(\cdot)$ is the Heaviside step function.

The Green's function of the beam is defined as the solution of Eq. (1) given that the right-hand side external load is characterized by $\delta(x-x_0)\delta(t-t_0)$. Taking two-dimensional Fourier transform and inverse on both sides of Eqs. (1) gives the Green's function

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, Jan. 25, 2000; final revision, Oct. 26, 2000. Associate Editor: A. K. Mal.

$$G(x,t;x_0,t_0)=(2\pi)^{-2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{i[\xi(x-x_0)+\omega(t-t_0)]\}}{EI\xi^4+K-m\omega^2}\times d\xi d\omega. \quad (2)$$

The solution of (1) given $F(x,t)$ is given by $y(x,t)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(x_0,t_0)G(x,t;x_0,t_0)dx_0dt_0$. Substituting Eqs. (1) and (2) into it gives

$$y(x,t)=\frac{\bar{P}}{2\pi}\int_{-\infty}^{\infty}\frac{\sin r_0\xi\exp[i\xi(x-vt)]}{r_0\xi(\xi^4-\bar{m}v^2\xi^2+\bar{K})}d\xi \quad (3)$$

where $\bar{P}=P/EI$, $\bar{K}=K/EI$ and $\bar{m}=m/EI$.

2 Closed-Form Representation of the Solution

Expression (3) can be further developed using complex function theory. To do so, one needs to identify the roots of the characteristic equation of this type $\xi^4+\bar{K}-\bar{m}v^2\xi^2=0$. Define a new variable $u=\xi^2$ so we have a quadratic equation

$$u^2-\bar{m}v^2u+\bar{K}=0. \quad (4)$$

Denote the critical velocity as $v_{\text{critical}}=(4\bar{K}/\bar{m}^2)^{1/4}$. Define dimensionless velocity (i.e., the Mach number) $M=v/v_{\text{critical}}$.

(a) Subsonic case ($M<1$).

Two roots of Eq. (4) are $u_1=[\bar{m}v^2(1+i\sqrt{M^{-4}-1})]/2$ and $u_2=[\bar{m}v^2(1-i\sqrt{M^{-4}-1})]/2$. Further, we have four complex valued roots $\xi_{1,2}=(\bar{m}v^2/2M^2)^{1/2}\exp[i(2j\pi+\theta)/2]$ and $\xi_{3,4}=(\bar{m}v^2/2M^2)^{1/2}\exp[i(2j\pi-\theta)/2]$ in which $\tan\theta=(M^{-4}-1)$ and $j=0$ and 1. In the case of $x-vt\geq 0$, we select the closed contour in the upper half ξ-plane and, if $x-vt<0$, in the lower half ξ-plane. To shorten the length of the paper, only the case $x-vt\geq 0$ is considered in the following. Applying the theorem of residue we obtain

$$y(x,t)=\frac{p}{2\pi EIr_0}\left\{2\pi i\sum_{\text{Im}\,\xi>0}\text{res}\left\{\frac{\sin r_0\xi\exp[i\xi(x-vt)]}{\xi(\xi^4-\bar{m}v^2\xi^2+\bar{K})}\right\}+\pi i\sum_{\text{Im}\,\xi=0}\text{res}\left\{\frac{\sin r_0\xi\exp[i\xi(x-vt)]}{\xi(\xi^4-\bar{m}v^2\xi^2+\bar{K})}\right\}\right\}. \quad (5)$$

After identifying the residues in Eq. (5), it is straightforward to see

$$y(x,t)=\frac{iP}{EIr_0}\sum_{l=1,4}\frac{\sin r_0\xi_l\exp[i\xi_l(x-vt)]}{5\xi_l^4-\bar{m}v^2\xi_l^2+K}. \quad (6)$$

(b) Sonic case ($M=1$).

In this case two duplicated roots of Eq. (5) are $u_{1,2}=\bar{m}v^2/2$. Thus four real valued roots $\xi_{1,2}=\xi_{3,4}=\pm(\bar{m}v^2/2)^{1/2}$. Since these two poles are of the second order, this means that a singularity occurs when integrating (4) along the contour. Using the same procedure as in the case of $M<1$, it is found that dynamic response in this case becomes infinite and the singularity is of the order $O(\varepsilon^{-1})$. This result indicates the existence of a resonance phenomenon as $v=v_{\text{critical}}$.

(c) Supersonic case ($M>1$).

Two roots of Eq. (4) are $u_1=[\bar{m}v^2(1+\sqrt{1-M^{-4}})]/2$ and $u_2=[\bar{m}v^2(1-\sqrt{1-M^{-4}})]/2$. Therefore, we have real valued roots $\xi_{1,2}=\pm R_1$ and $\xi_{3,4}=\pm R_2$ where $R_1=\{\bar{m}v^2[1+(1-M^{-4})^{1/2}]/2\}^{1/2}$ and $R_2=\{\bar{m}v^2[1-(1-M^{-4})^{1/2}]/2\}^{1/2}$. Apparently, since the distribution of the roots of the characteristic equation depends heavily on the range of the Mach number, one may expect that dynamic response of the beam to a moving load will also be distinct for different Mach number.

The poles of a system without damping can be thought of as the limit situation of poles of a system with damping while the damping is approaching zero. The poles of a physical system with tiny damping can be determined by seeking the roots of a new characteristic algebraic equation in which a negative infinitesimal imaginary term is added into the previous characteristic algebraic equation, i.e., $\xi^4-\bar{m}v^2\xi^2+\bar{K}-i\varepsilon=0$ where ε is a positive infinitesimal real number. Since $u=\xi^2$. This new characteristic equation becomes

$$u^2-\bar{m}v^2u+\bar{K}-i\varepsilon=0. \quad (7)$$

The square root of the discriminant of Eq. (7) is $\Delta^{1/2}=\bar{m}v^2(1-M^{-4})^{1/2}\exp(i\psi/2)$, in which $\tan\psi=\beta(1-M^{-4})^{-1}$ and $\beta=4\bar{m}^{-2}v^{-4}\varepsilon$. Since $\beta\to 0^+$ as $\varepsilon\to 0^+$, we have $\psi\to 0^+$ and $u_1=1/2\bar{m}v^2[1+(1-M^{-4})^{1/2}]\exp(i\lambda_1)$ and $u_2=1/2\bar{m}v^2[1-(1-M^{-4})^{1/2}]\exp(i\lambda_2)$ as two roots, in which $\tan\lambda_1=(1-M^{-4})^{1/2}\sin\psi/2/1+(1-M^{-4})^{1/2}\cos\psi/2$ and $\tan\lambda_2=-(1-M^{-4})^{1/2}\sin\psi/2/1-(1-M^{-4})^{1/2}\cos\psi/2$, respectively. Thus $\xi_{1,2}=R_1\exp[i(\lambda_1+2j\pi)/2]$ and $\xi_{3,4}=R_2\exp[i(\lambda_2+2j\pi)/2]$, $(j=0,1)$ where $R_{1,2}=\{\bar{m}v^2[1\pm(1-M^{-4})^{1/2}]/2\}^{1/2}$. Realize $\lambda_1=\lim_{\psi\to 0^+}\arctan(\tan\lambda_1)=0^+$ and $\lambda_2=\lim_{\psi\to 0^+}\arctan(\tan\lambda_2)=0^-$, as $\psi\to 0^+$, the roots ξ_1, ξ_2, ξ_3, and ξ_4, respectively, approach their own limits $\arg\xi_{1,2}=\{^{\lambda_1/2}_{\pi+\lambda_1/2}=\{^{0^+}_{\pi+0^+}$ and $\arg\xi_{3,4}=\{^{\lambda_2/2}_{\pi+\lambda_2/2}=\{^{0^-}_{\pi+0^-}$.

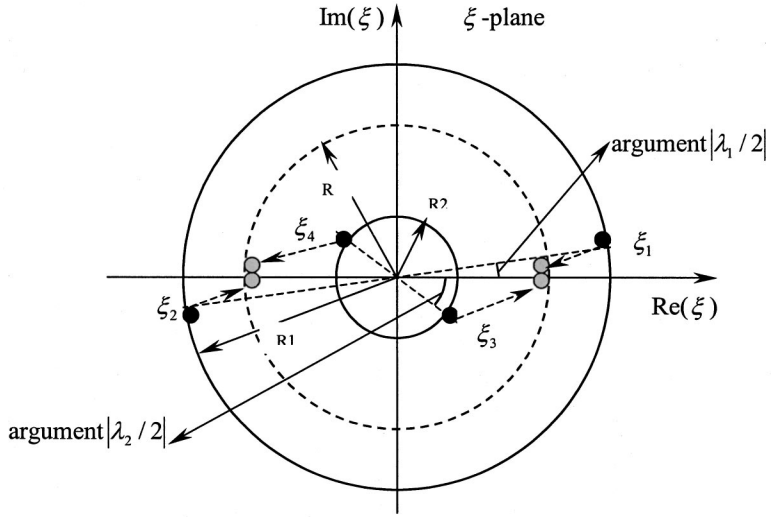


Fig. 1 Poles of the beam on an elastic foundation with tiny amount of damping

Figure 1 depicts the distribution of these poles in the complex ξ-plane. As v approaches critical velocity v_{critical} (i.e., $M \to 1$), four poles (black points) will move towards those poles (gray points) of the case $v = v_{\text{critical}}$. Each pair of gray points on one side of the imaginary axis will get more and more close to each other as $M \to 1$, and becomes a single pole of the second order. The dynamic response of the bean is given by

$$y(x,t) = \frac{iP}{2EIr_0} \sum_{l=1,4} \frac{\sin r_0 \xi_l \exp[i\xi_l(x-vt)]}{5\xi_l^4 - \bar{m}v^2\xi_l^2 + K} \quad \text{for } x - vt \geqslant 0. \tag{8}$$

In the case of $x - vt < 0$, we just need to replace $l = 1,4$ in Eq. (6) or (8) by $l = 2,3$.

3 Maximum Response

Define $\eta = x - vt$. The derivative of $y(x,t)$ with respect to η suggests that $\eta = 0$ correspond to the extreme point. The maximum response in the case of $M < 1$ can be obtained by substituting $x - vt = 0$ into Eq. (7). Define new parameters. The maximum response is $y_{\max}^{\text{moving}}(x = vt) = iP/EIr_0\{\sin[r\exp(i\theta/2)]/S_1 + iW_1 + \sin\{r\exp[i(-\theta+\pi)/2]\}/S_2 + iW_2\}$ in which $S_1 = 5\phi^4 \cos 2\theta - 3\bar{m}v^2\phi^2 \cos\theta + K$, $S_2 = 5\phi^4 \cos 2\theta + 3\bar{m}v^2\phi^2 \cos\theta + K$, $W_1 = 5\phi^4 \sin 2\theta - 3\bar{m}v^2\phi^2 \sin\theta$, and $W_2 = 5\phi^4 \sin 2\theta + 3\bar{m}v^2\phi^2 \sin\theta$, $\phi = (\bar{m}v^2/2M^2)^{1/2}$ and $r = \phi r_0$. Using Maclaren series to expand $y_{\max}^{\text{moving}}(x = vt)$ and taking only the real part

$$y_{\max}^{\text{moving}}(x = vt) = -\frac{P}{EIr_0}\left\{\frac{A_1S_1 - B_1W_1}{S_1^2 + W_1^2} + \frac{A_2S_2 - B_2W_2}{S_2^2 + W_2^2}\right\} \tag{9}$$

where

$$A_1 = \sum_{n=0}^{\infty} \frac{(-1)^n r^{2n+1} \sin[\theta(2n+1)/2]}{(2n+1)!},$$

$$B_1 = \sum_{n=0}^{\infty} \frac{(-1)^n r^{2n+1} \cos[\theta(2n+1)/2]}{(2n+1)!},$$

$$A_2 = \sum_{n=0}^{\infty} \frac{r^{2n+1} \cos[\theta(2n+1)/2]}{(2n+1)!} \quad \text{and}$$

$$B_2 = \sum_{n=0}^{\infty} \frac{r^{2n+1} \sin[\theta(2n+1)/2]}{(2n+1)!}.$$

It should be noted that, although $\eta = 0$ can make $d/d\eta y = 0$ satisfied, it is a sufficient condition rather than a necessary condition. Actually, the response of the beam at the center of the moving load is the maximal response in the case of $M < 1$, while the response of that location remains quiescent in the case of $M > 1$. It also should be pointed out that similar method is applicable to moving load problem with damping considered in the governing equation. Given the limit of the content, the result is not presented herein.

References

[1] Fryba, L., 1977, *Vibration of Solids and Structures Under Moving Loads*, Noordhoff, Groningen, The Netherlands.
[2] Kenney, J. T., 1954, ''Steady State Vibrations of Beam on Elastic Foundation for Moving Loads,'' ASME J. Appl. Mech., **21**, pp. 359–364.
[3] Steele, C. R., 1967, ''The Finite Beam With a Moving Load,'' ASME J. Appl. Mech., **34**, p. 111.
[4] Huang, C. C., 1977, ''Traveling Loads on a Viscoelastic Timoshenko Beam,'' ASME J. Appl. Mech., **44**, pp. 183–184.
[5] Choros, J., and Adams, G. G., 1979, ''A Steadily Moving Load on an Elastic Beam Resting on a Tensionless Winkler Foundation,'' ASME J. Appl. Mech., **46**, No. 1, pp. 175–180.
[6] Jezequel, L., 1981, ''Response of Periodic Systems to a Moving Load,'' ASME J. Appl. Mech., **48**, pp. 613–618.
[7] Elattary, M. A., 1991, 'Moving Loads on an Infinite Plate Strip of Constant Thickness,'' J. Phys. D: Appl. Phys., **24**, No. 4, 541–546.
[8] Lee, H. P., 1994, ''Dynamic Response of a Beam With Intermediate Point Constraints Subject to a Moving Load,'' J. Sound Vib., **171**, No. 3, pp. 361–368.
[9] Sun, L., and Deng, X., 1997, ''Transient Response for Infinite Plate on Winkler Foundation by a Moving Distributed Load,'' Chin. J. Appl. Mech., **14**, No. 2, pp. 72–78.
[10] Sun, L., 1998, ''Theoretical Investigations on Vehicle-Ground Dynamic Interaction,'' Final Report prepared for National Science Foundation of China, Southeast University, Nanjing, China.
[11] Sun, L., and Greenberg, B., 2000, ''Dynamic Response of Linear Systems to Moving Stochastic Sources,'' J. Sound Vib., **229**, No. 4, pp. 957–972.
[12] Benedetti, G. A., 1974, ''Dynamic Stability of a Beam Loaded by a Sequence of Moving Mass Particles,'' ASME J. Appl. Mech., **41**, pp. 1069–1071.

An Analytic Algorithm of Stresses for Any Double Hole Problem in Plane Elastostatics

Lu-qing Zhang
Engineering Geomechanics Laboratory, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China

Ai-zhong Lu
Professor, Department of Civil Engineering, Shandong University of Science and Technology, Tai'an 271019, China

This paper gives an analytic algorithm to plane elastostatic problem of an infinite medium containing two holes of arbitrary shapes and arrangement, using Schwarz's alternating method, and finds that the method has a very quick convergence speed even for solving a complex double hole problem.

[DOI: 10.1115/1.1352065]

1 Introduction

There are a large number of papers in plane elastostatics dealing with regions containing two circular holes ([1–4]). It seems that only Hasebe et al. [5] provided one analysis method for the problem of two complex holes in which one hole is of complex profile and the other is a crack. However, the method is only suitable for a symmetrical double hole problem.

The crucial ingredient in solving a double hole problem by means of Schwarz's alternating method is the repeated solution of a single hole problem, which can be well solved by Muskhelishvili's method ([6]) via a conformal transformation of mapping the given hole shape into a unit circle. The iterative solutions for the Schwarz's alternating method needs many repeated transformations between the physical and mapped planes. In order to conduct the iterative solutions easily, two mapping functions of two holes, $z_1 = \omega_1(\zeta_1)$ and $z_2 = \omega_2(\zeta_2)$, and two corresponding inverse mapping functions, $\zeta_1 = \omega_1^{-1}(z_1)$ and $\zeta_2 = \omega_2^{-1}(z_2)$, are introduced. In the process of iterative solutions every iteration refers to the completion of solutions for two single hole problems.

2 Basic Formulas for Stress Analysis of Any Double Hole Problem

In Fig. 1 z_1 and z_2 are the complex coordinates in $x_1o_1y_1$ and $x_2o_2y_2$ local coordinate systems, respectively; c is the relative

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 22, 2000; final revision, Sept. 20, 2000. Associate Editor: J. R. Barber.

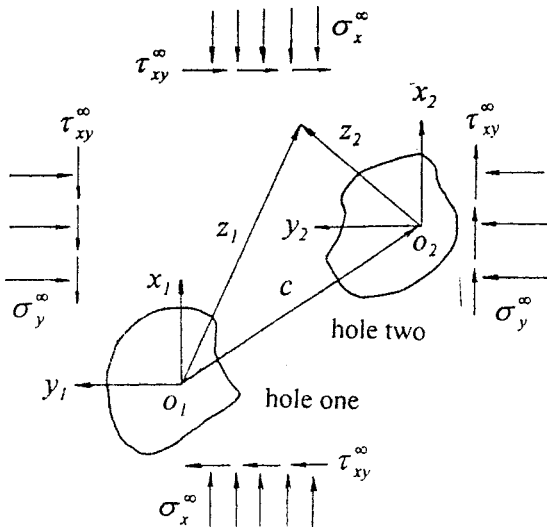


Fig. 1 The calculating model for any double hole problem

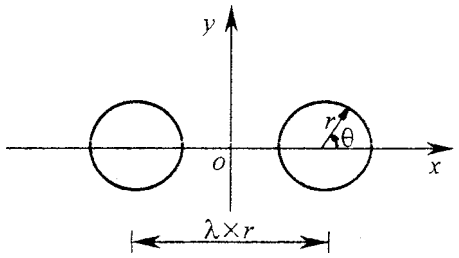


Fig. 2 The problem for two circular holes

position vector of two holes in $x_1o_1y_1$ coordinate system; σ_x^∞, σ_y^∞, and τ_{xy}^∞ are external loads uniformly distributed at infinity.

In the process of the first iteration, the presence of hole one will lead to a single hole problem, whose solution of stresses can be written in terms of two complex stress functions, $\varphi_{11}(\zeta_1)$ and $\psi_{11}(\zeta_1)$, of the complex variable ζ_1. The stress boundary condition for the presence of hole one is

$$\varphi_{11}(\sigma_1)+\frac{\omega_1(\sigma_1)}{\overline{\omega_1'(\sigma_1)}}\overline{\varphi_{11}'(\sigma_1)}+\overline{\psi_{11}(\sigma_1)}=f_1(\sigma_1) \tag{1}$$

where a prime on a function denotes differentiation with respect to its argument, and a bar on a function indicates its conjugate; σ_1 is the value of ζ_1 on the unit circle; $\varphi_{11}(\sigma_1)$ and $\psi_{11}(\sigma_1)$ are the values of $\varphi_{11}(\zeta_1)$ and $\psi_{11}(\zeta_1)$ on the unit circle, respectively; $f_1(\sigma_1)$ is the principal vector of surface forces at the edge of hole one.

The stress functions $\varphi_{11}(\zeta_1)$ and $\psi_{11}(\zeta_1)$ can be used as the loading functions for solving another single hole problem induced by the presence of hole two. At this stage the boundary conditions are satisfied at the edge of hole one, however, there exist redundant surface forces at the edge of hole two. The redundant surface forces are obtained directly by three coordinate transformations between coordinates ζ_2, z_2, z_1 and ζ_1 and a formula

$$f_{12}(\sigma_2)=\varphi_{11}(\gamma_1)+\frac{\omega_1(\gamma_1)}{\overline{\omega_1'(\gamma_1)}}\overline{\varphi_{11}'(\gamma_1)}+\overline{\psi_{11}(\gamma_1)} \tag{2}$$

where σ_2 is the value of boundary point t_2 of hole two in ζ_2 plane; $f_{12}(\sigma_2)$ is the principal vector of the redundant surface forces with respect to σ_2; γ_1 is the coordinate of σ_2 in ζ_1 plane via mapping transformation $t_2=\omega_2(\sigma_2)$, coordinate translation $T_1=t_2+c$ and inverse mapping transformation $\gamma_1=\omega_1^{-1}(T_1)$.

The distribution of $f_{12}(\sigma_2)$ at the edge of hole two can be approximated by complex series $\sum_{k=-L}^{L} D_k\sigma_2^k$, in which D_k is the complex coefficient of σ_2^k. In order to eliminate the redundant surface forces, the reversed forces of $\sum_{k=-L}^{L} D_k\sigma_2^k$, $-\sum_{k=-L}^{L} D_k\sigma_2^k$, are imposed at the edge of hole two, yielding the other single hole problem in the first iteration. The solution for the presence of hole two can be expressed by two complex stress functions $\varphi_{22}(\zeta_2)$ and $\psi_{22}(\zeta_2)$. The corresponding stress boundary condition is

$$\varphi_{22}(\sigma_2)+\frac{\omega_2(\sigma_2)}{\overline{\omega_2'(\sigma_2)}}\overline{\varphi_{22}'(\sigma_2)}+\overline{\psi_{22}(\sigma_2)}=f_2(\sigma_2)-\sum_{k=-L}^{L} D_k\sigma_2^k \tag{3}$$

Table 1 The comparison of the maximum tensile stresses at the edge of the right hole from two methods

λ	solution	all-around unit tension		horizontal unit tension		vertical unit tension	
		$\theta=0$	$\theta=\pi$	$\theta=+\pi/2$	$\theta=-\pi/2$	$\theta=0$	$\theta=\pi$
2	exact solution	2.894	∞	2.569	2.569	3.869	∞
	iterative solution	2.659375	35.570570	2.557303	2.557303	3.601490	40.281910
3	exact solution	2.255	2.887	2.623	2.623	3.151	3.264
	iterative solution	2.254839	2.887496	2.623146	2.623145	3.150994	3.264019
4	exact solution	2.158	2.411	2.703	2.703	3.066	3.020
	iterative solution	2.157655	2.410827	2.702637	2.702636	3.066025	3.020181
6	exact solution	2.080	2.155	2.825	2.825	3.020	2.992
	iterative solution	2.080380	2.154591	2.825038	2.825037	3.020114	2.992179
10	exact solution	2.033	2.049	2.927	2.927	3.004	2.997
	iterative solution	2.032807	2.048810	2.926672	2.926672	3.004336	2.996998
16	exact solution	2.014	2.018	2.970	2.970	3.001	2.999
	iterative solution	2.013795	2.017700	2.969799	2.969798	3.001042	2.999149
∞	exact solution	2.000	2.000	3.000	3.000	3.000	3.000
	iterative solution	2.000000	2.000000	3.000000	3.000000	3.000000	3.000000

Note: To see the difference from the exact solution with three decimal numbers, every iterative one is given by six decimal numbers.

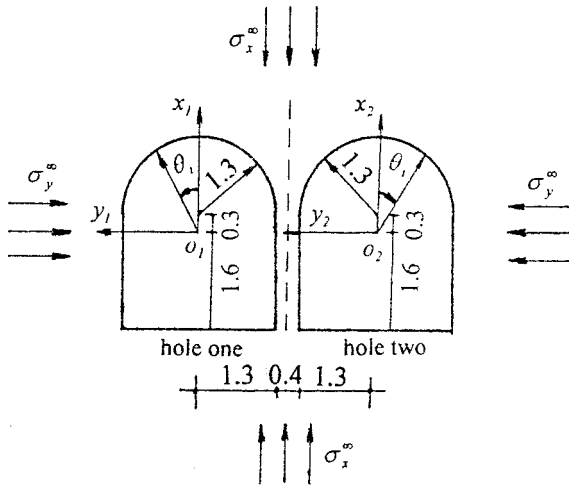


Fig. 3 The problem for two complex holes

where $\varphi_{22}(\sigma_2)$ and $\psi_{22}(\sigma_2)$ are the values of $\varphi_{22}(\zeta_2)$ and $\psi_{22}(\zeta_2)$ on the unit circle in ζ_2 plane, respectively; $f_2(\sigma_2)$ is the principal vector of surface forces at the edge of hole two.

The superposition of $\varphi_{11}(\zeta_1)$, $\psi_{11}(\zeta_1)$ and $\varphi_{22}(\zeta_2)$, $\psi_{22}(\zeta_2)$ is the solution for the first iteration of the Schwarz's alternating method. At this stage, the boundary conditions only at the edge of hole two are satisfied. Of course, the second and later iterations can be operated.

Taking $\varphi(\zeta)$ and $\psi(\zeta)$ as the superposition of two stress functions for all required iterative solutions, the stress components can be obtained readily.

3 Discussions on the Convergence Accuracy of Iterative Solutions

3.1 Comparison With the Exact Solution for the Problem of Two Circular Holes. Now let us consider a linearly elastic medium containing two equal circular holes, as plotted in Fig. 2. Three fundamental loading cases are discussed in some detail, namely, the all-around, horizontal and vertical tensions applied at infinity. Owing to the symmetry of the problem, Table 1 only

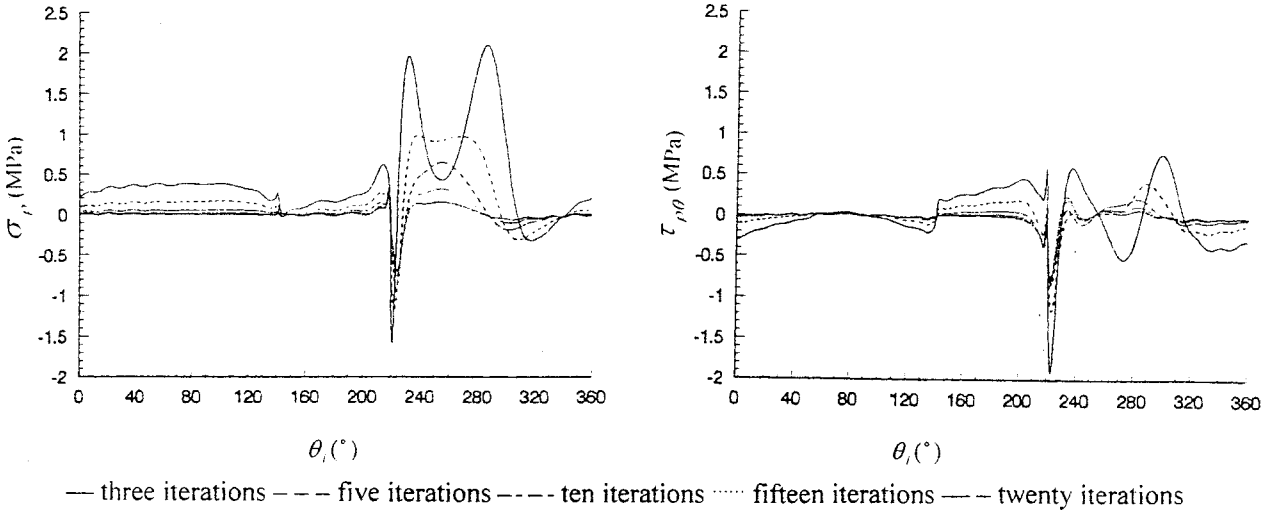


Fig. 4 The redundant surface forces σ_ρ and $\tau_{\rho\theta}$ for different iterations

Table 2 The maximum compressive stresses at the edges of two holes

arrangement		hole two / hole one, $c = 0.0 - 4.0i$		hole one / hole two, $c = 2\sqrt{2} - 2\sqrt{2}i$		hole two / hole one, $c = 4.0 + 0.0i$	
loads	position	hole one	hole two	hole one	hole two	hole one	hole two
$\sigma_x^\infty = 2$ $\sigma_y^\infty = 1$	$\theta = 141°$	14.713	14.722	15.742	13.297	14.181	11.657
	$\theta = 219°$	13.653	13.667	9.724	23.706	14.181	11.657
$\sigma_x^\infty = 1.5$ $\sigma_y^\infty = 1.5$	$\theta = 141°$	14.425	14.436	16.030	10.339	14.978	15.010
	$\theta = 219°$	11.291	11.336	12.446	25.258	14.978	15.010
$\sigma_x^\infty = 1$ $\sigma_y^\infty = 2$	$\theta = 141°$	14.566	14.574	16.817	7.859	16.208	18.801
	$\theta = 219°$	9.538	9.587	15.688	28.087	16.208	18.801

gives the maximum tensile stresses on the boundary of the right hole in which the iterative solutions are obtained by the Schwarz's alternating method for ten iterations and the exact one given by Ling [3].

3.2 Accuracy Analysis for the Problem of Two Complex Holes. Let us consider the problem of an infinite and linearly elastic region, containing two complex holes, only under the action of compressive stresses at infinity ($\sigma_x^\infty = 10$ MPa and $\sigma_y^\infty = 20$ MPa) (see in Fig. 3). If the solution is terminated at some iteration, the boundary condition of zero surface forces along hole two will be satisfied exactly and along hole one approximately. Figure 4 plots the distribution of redundant surface forces along hole one for 3, 5, 10, 15, and 20 iterative solutions, seen from which the redundant surface forces are gradually reduced to zeroes as the further iteration.

4 The Maximum Stresses Around Two Holes for Different Loads and Arrangements

This paper still takes two holes in Fig. 3 as examples, only changing the loads at infinity and arrangement of the two holes. Three loading cases and three arrangement cases are investigated, and the maximum stresses at the edges of two holes are presented in Table 2.

Acknowledgment

This paper is supported by the Chinese National Natural Science Foundation (No. 49772166).

References

[1] Howland, R. C. J., and Knight, R. C., 1939, ''Stress Functions for a Plate Containing Groups of Circular Holes,'' Philos. Trans., **238**, pp. 357–392.
[2] Green, A. G., 1940, ''General Biharmonic Analysis for a Plate Containing Circular Holes,'' Proc. R. Soc. London, Ser. A, **176**, pp. 121–139.
[3] Ling, Chin-bing, 1948, ''On the Stresses in a Plate Containing Two Circular Holes,'' J. Appl. Phys., **19**, pp. 77–82.
[4] Ukadgaonker, V. G., 1982, ''Stress Analysis of a Plate Containing Two Circular Holes Having Tangential Stresses,'' AIAA J., **20**, pp. 125–128.
[5] Hasebe, N., Yoshikawa, K., Ueda, M., and Nakamura, T., 1994, ''Plane Elastic Solution for the Second Mixed Boundary Value Problem and Its Application,'' Archive of Applied Mechanics, **64**, pp. 295–306.
[6] Muskhelishvili, N. I., 1953, ''Some Basic Problems of Mathematical Theory of Elasticity,'' P. Noordhoff, Groningen, The Netherlands.

The Rotating Tautochrone

T. J. Osler
Mathematics Department, Rowan University, Glassboro, NJ 08028
e-mail: osler@rowan.edu

E. Flores
Chemistry and Physics Department, Rowan University, Glassboro, NJ 08028
e-mail: flores@rowan.edu

In a recent paper by Flores and Osler, the authors investigated tautochrone curves in the xy-plane under an arbitrary potential V(y). In this paper we imagine that the xy-plane of the tautochrone curve is rotating about the y-axis with constant angular momentum. We find the differential equation of the tautochrone curves. While this differential equation is difficult to solve analytically, several exact solutions (in terms of elementary functions) are obtained in an indirect manner. Intuitive motivation for tautochrone motion is given. [DOI: 10.1115/1.1352066]

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Apr. 2, 2000; final revision, Oct. 9, 2000. Associate Editor: A. A. Ferri.

1 Introduction

Consider a bead of unit mass that moves on a frictionless wire described by the curve $x=x(y)$ in the xy-plane. Assume that the bead starts at time $t=0$ at the point $(x(Y),Y)$ with no initial velocity and that the curve $x=x(y)$ terminates on the x-axis at the point $(x_0,0)$. The motion of the bead is governed by a potential $V(y)$ as it moves along the curve $x=x(y)$. This curve is called a *tautochrone* if the time T required for the motion from the starting point at $(x(Y),Y)$ to the final point $(x_0,0)$ is *independent of Y,* (the starting height on the curve). The problem of determining the shape of the tautochrone curve under the gravitational potential $V(y)=gy$ was solved by Huygens and by Abel. The authors studied this problem under arbitrary potentials $V(y)$ in a recent paper ([1]) using the fractional calculus. In this paper we assume that the xy-plane containing our curve $x=x(y)$ is rotating with constant angular momentum L about a shaft centered on the y-axis.

In our previous study ([1]) (angular velocity $\omega=0$) we determined that the time T for the bead to descend from $y=Y$ to $y=0$ is given by

$$\int_0^Y \frac{ds}{\sqrt{V(Y)-V(y)}} = \sqrt{2}T. \tag{1}$$

Here s measures the distance along the curve $x=x(y)$ starting from $(x_0,0)$ to the point $(x(y),y)$. Using the fractional calculus we determined that when the curve is a tautochrone, then the potential and the arc length are related by

$$V(y) = \frac{\pi^2}{8T^2}s^2. \tag{2}$$

We also determined that the differential equation satisfied by the tautochrone curve is

$$1+x'(y)^2 = \frac{2T^2}{\pi^2}\frac{V'(y)^2}{V(y)}. \tag{3}$$

For (3) to be valid, we require that $V(0)=0$. (This can always be achieved by simply adding a constant to the potential.) The solution for our tautochrone curve in terms of the given potential is

$$x(y) = \int_0^y \sqrt{\frac{2T^2}{\pi^2}\frac{V'(u)^2}{V(u)} - 1}\,du + x_0. \tag{4}$$

We will use the above results when we solve our rotating tautochrone problem.

2 The Rotating Versus Nonrotating Tautochrones

The sum of the kinetic and potential energies for the rotating parts and for our bead of unit mass on the wire $x=x(y)$ in the rotating xy-plane (angular velocity $\omega(y)$) is

$$\frac{1}{2}I\omega(Y)^2 + \frac{1}{2}\omega(Y)^2x(Y)^2 + V(Y)$$
$$= \frac{1}{2}\left(\frac{ds}{dt}\right)^2 + \frac{1}{2}I\omega(y)^2 + \frac{1}{2}\omega(y)^2x(y)^2 + V(y). \tag{5}$$

The following are important features of (5):

A The wire and rotating ''parts'' are rigid and rotating freely (without torque) about the y-axis. (See Fig. 1.)

B On the left of (5) we see the energy at the moment the bead is released at the point (X,Y) and on the right we see the energy when the bead is at the arbitrary point (x,y) on the wire.

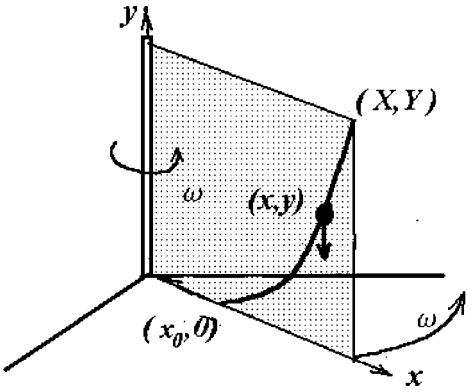


Fig. 1 The rotating frictionless wire with supports

C The wire is given an initial angular velocity and the bead is started with zero velocity relative to the rotating frame. In particular, at time $t=0$ the system is rotating with angular velocity $\omega(Y)$ about the y-axis.

D As a consequence of C, at time $t=0$ the bead has no component of velocity tangent to the curve but it does have a component of velocity perpendicular to the xy-plane given by $\omega(Y)x(Y)$.

E During the motion the angular velocity of the bead (and the system of parts) given by $\omega(y)$ will vary. It will change so that angular momentum is always conserved. (See relation (12)).

F The moment of inertia of the wire, rotating shaft and supports is I and its kinetic energy is the term $1/2I\omega(y)^2$.

G The term ds/dt is the magnitude of the velocity component tangent to the curve. The arc length s is measured from the terminal point $(x_0,0)$ to the moving point $(x(y),y)$.

If we call $V_*(y)$ the terms

$$V_*(y)=V(y)+\frac{1}{2}I\omega(y)^2+\frac{1}{2}\omega(y)^2x(y)^2, \tag{6}$$

we can abbreviate the writing of (5) as simply

$$V_*(Y)=V_*(y)+\frac{1}{2}\left(\frac{ds}{dt}\right)^2. \tag{7}$$

Solving (7) for ds/dt we get

$$\frac{ds}{dt}=-\sqrt{2}\sqrt{\{V_*(Y)\}-\{V_*(y)\}}. \tag{8}$$

The minus sign in (8) is due to the assumption that the arc length s is decreasing. This requires that the initial angular speed $\omega(Y)$ be small enough that when the bead is released with zero relative velocity, the bead falls downward instead of flying outward. For example, in the case of a gravity-potential, it assumes that

$$\frac{dy}{dx}>\frac{x(Y)\omega(Y)^2}{g}.$$

We can now write

$$\sqrt{2}dt=-\frac{ds}{\sqrt{V_*(Y)-V_*(y)}}. \tag{9}$$

Integrating from the beginning of the motion to the end we get

$$\sqrt{2}T=\int_0^Y\frac{ds}{\sqrt{V_*(Y)-V_*(y)}}. \tag{10}$$

Notice that Eqs. (1) and (10) have the same form, thus, they have the same solution. This implies that for a given curve $x=x(y)$, the time T for the rotating case under the potential $V(y)$ is the same as the time for the nonrotating case under the potential $V_*(y)$. This last statement is important for our work. If we know that the curve $x=x(y)$ is a nonrotating tautochrone under the potential $V_*(y)$, then we know that the same curve is a tautochrone rotating with angular velocity $\omega(y)$ under the potential

$$V(y)=V_*(y)-\frac{1}{2}(I+x(y)^2)\omega(y)^2. \tag{11}$$

We can eliminate $\omega(y)$ from (11) by using the conservation of angular momentum expressed as

$$L=(I+x(Y)^2)\omega(Y)=(I+x(y)^2)\omega(y). \tag{12}$$

We require that this angular velocity vary with the starting height Y so as to keep the angular momentum L constant.

Solving (12) for $\omega(y)$ and substituting into (11) we get

$$V(y)=V_*(y)-\frac{1}{2}\frac{L^2}{(I+x(y)^2)}. \tag{13}$$

We recall from our previous paper that all the potentials we use are required to satisfy $V(0)=V_*(0)=0$ so that relations (1) to (4) are all valid. Since $x(0)=x_0$ we must add the constant term $L^2/2(I+x_0^2)$ to the right side of (13) so that all potentials are zero when $y=0$. We get

$$V(y)=V_*(y)-\frac{1}{2}\frac{L^2}{(I+x(y)^2)}+\frac{L^2}{2(I+x_0^2)}. \tag{14}$$

We will use (14) to find several rotating tautochrones in the next section.

3 Finding Exact Rotating Tautochrones

In our previous paper we found exact nonrotating tautochrones indirectly. We started with a curve $x=x(y)$ for which we could calculate the arc length $s=s(y)$ exactly. We then used relation (2) to find the potential that would make this curve a tautochrone. Nine such curves were selected for their ease of calculation. All nine can be easily modified using (14) to give us rotating tautochrones. The results are shown in Table 1.

While all the resulting potentials are bizarre, we believe it is important to collect exact solutions of mechanics problems whenever they are possible. When exact analytic solutions cannot be found, perturbation or numerical solutions are usually possible. The latter tell us much less than the exact solutions.

4 The Differential Equation for the Rotating Tautochrone

We will now find the differential equation satisfied by the rotating tautochrone. Relation (3) is the differential equation for the nonrotating tautochrone. Substituting $V_*(y)$ from (14) for $V(y)$ in (3) we get

$$1+x'(y)^2=\frac{2T^2}{\pi^2}\frac{\left\{V'(y)-\dfrac{L^2x(y)x'(y)}{(I+x(y)^2)^2}\right\}^2}{\left\{V(y)+\dfrac{L^2}{2(I+x(y)^2)}-\dfrac{L^2}{2(I+x_0^2)}\right\}}. \tag{15}$$

This is our differential equation for the rotating tautochrone. It is much more complex and difficult to solve than (3). We can rewrite (15) in the form

$$\pi^2\{2(I+x_0^2)(I+x(y)^2)V(y)-L^2x(y)^2\}(I+x(y)^2)^3(1+x'(y)^2)$$
$$=4(I+x_0^2)T^2\{(I+x(y)^2V'(y)-L^2x(y)x'(y)\}^2. \tag{16}$$

We can use (15) or (16) to check our solutions in Table 1, but solving exactly for a given potential appears to be difficult. Of course, numerical solutions are possible.

5 The Case Where I is Large

We now discuss the approximation where the moment of inertia of the bead, x^2, is much smaller than I, the moment of inertia of

Table 1 Examples of Potentials and Corresponding Rotating Tautochrone Curves

	Tautochrone Curve	Arc Length s	Potential $V(y)$
1	$x=R\sin\theta+x_0$, $y=R-R\cos\theta$ circle: center (x_0,R), radius R	$s=R\theta=R\cos^{-1}\left(\frac{R-y}{R}\right)$	$\frac{\pi^2R^2}{8T^2}\left\{\cos^{-1}\left(\frac{R-y}{R}\right)\right\}^2+\frac{L^2}{2(I+x_0^2)}$ $-\frac{L^2}{2(I+(\sqrt{2Ry-y^2}+x_0)^2)}$
2	$x=R-R\cos\theta+x_0$ $y=R\sin\theta$ circle: center $(R+x_0,0)$, radius R	$s=R\theta=R\sin^{-1}(y/R))$	$\frac{\pi^2R^2}{8T^2}\{\sin^{-1}(y/R)\}^2+\frac{L^2}{2(I+x_0^2)}$ $-L^2/(2(I+(R-\sqrt{R^2-y^2}+x_0)^2))$
3	$x=R\sin(\alpha+\theta)-R\sin\alpha+x_0$ $y=R\cos\alpha-R\cos(\alpha+\theta)$, R and α fixed circle: radius R center $(x_0-R\sin\alpha,R\cos\alpha)$	$s=R\theta$ $\theta=\cos^{-1}\left(\frac{R\cos(\alpha)-y}{R}\right)-\alpha$	$\frac{\pi^2R^2}{8T^2}\left\{\cos^{-1}\left(\frac{a-y}{R}\right)-\alpha\right\}^2+\frac{L^2}{2(I+x_0^2)}$ $-\frac{L^2}{2(I+(\sqrt{R^2-(R\cos\alpha-y)^2}-R\sin\alpha+x_0)^2)}$ where $a=R\cos(\alpha)$
4	$x=R(\theta+\sin\theta)+x_0$ $y=R(1-\cos\theta)$ inverted cycloid: base line $y=2R$	$s=4R\sin(\theta/2)$ $s=2\sqrt{2Ry}$	$\frac{\pi^2R}{T^2}y+\frac{L^2}{2(I+x_0^2)}$ $-\frac{L^2}{2\left(I+\left\{R\cos^{-1}\left(\frac{R-y}{R}\right)+\sqrt{2Ry-y^2}+x_0\right\}^2\right)}$
5	$x=R\theta-R\sin\theta+x_0$ $y=R-R\cos\theta$ cycloid	$s=4R(1-\cos(\theta/2))$ $s=4R\left(1-\sqrt{1-\frac{y}{2R}}\right)$	$\frac{\pi^2R}{T^2}y+\frac{L^2}{2(I+x_0^2)}$ $-\frac{L^2}{2\left(I+\left\{R\cos^{-1}\left(\frac{R-y}{R}\right)-\sqrt{2Ry-y^2}+x_0\right\}^2\right)}$
6	$x=ay+x_0$	$s=\sqrt{1+a^2}y$	$\frac{\pi^2(1+a^2)}{8T^2}y^2-\frac{L^2}{2(I+(ay+x_0)^2)}+\frac{L^2}{2(I+x_0^2)}$
7	$x=2\sqrt{ay^3}/3+x_0$	$s=\frac{2}{3a}[(1+ay)^{3/2}-1]$	$\frac{\pi^2}{18a^2T^2}[(1+ay)^{3/2}-1]^2$ $-\frac{L^2}{2(I+(2\sqrt{ay^3}/3+x_0)^2)}+\frac{L^2}{2(I+x_0^2)}$
8	$x=ay^2/2+x_0$	$s=(ay\sqrt{1+a^2y^2}+\ln(ay+\sqrt{1+a^2y^2}))/2a$	$\frac{\pi^2}{32a^2T^2}(ay\sqrt{1+a^2y^2}+\ln(ay+\sqrt{1+a^2y^2}))^2$ $-\frac{L^2}{2(I+(ay^2/2+x_0)^2)}+\frac{L^2}{2(I+x_0^2)}$
9	$y=a\cosh((x-x_0)/a)-a+x_0$	$s=\sqrt{(y+a)^2-a^2}$	$\frac{\pi^2}{8T^2}[(y+a)^2-a^2]+\frac{L^2}{2(I+x_0^2)}$ $-L^2/\left(2\left(I+\left\{a\cosh^{-1}\left(\frac{y+a}{a}\right)+x_0\right\}^2\right)\right)$

the wire, shaft and support. We will also assume that $x_0=0$, so that the wire is attached directly to the shaft at the bottom. Starting with (14) we can write

$$V_*(y)=V(y)+\frac{L^2}{2I(1+x(y)^2/I)}-\frac{L^2}{2I}.$$

Expanding the second term on the right in powers of $x(y)^2/I$ we obtain to a first-order approximation

$$V_*(y)=V(y)-\frac{1}{2}\frac{L^2}{I^2}x(y)^2.$$

The conservation of angular momentum (12), to a first-order approximation becomes simply $L=I\omega$ where ω is now a constant. Thus the above expression becomes

$$V_*(y)=V(y)-\frac{1}{2}\omega^2x(y)^2.$$

The differential Eq. (15) is now much simpler,

$$1+x'(y)^2=\frac{2T^2}{\pi^2}\frac{\{V'(y)-\omega^2x(y)x'(y)\}^2}{\left\{V(y)-\frac{1}{2}\omega^2x(y)^2\right\}}.$$

As an example consider the potential of a simple harmonic oscillator $V(y)=ky^2/2$ and try a straight line solution $x(y)=ay$, the differential equation becomes (after simplifying)

$$1+a^2=\frac{4T^2}{\pi^2}\{k-a^2\omega^2\}.$$

Solving for a we get $a=\sqrt{4T^2k-\pi^2/4T^2\omega^2+\pi^2}$. Thus, under a harmonic oscillator potential the tautochrone curve is a straight line with slope a.

In general the potential for the rotating tautochrone, to a first-order approximation, is

$$V(y)=V_*(y)+\frac{1}{2}\omega^2x(y)^2+C.$$

The first term is the potential of a nonrotating tautochrone $V_*(y)$, the second term is the potential of a simple harmonic oscillator, and C is a constant. The second term could be produced by an ideal spring stretched along the x-axis with one end attached to the bead and at the other to the shaft. As the bead moves down so does the spring. The tension in the spring serves to exactly cancel the centrifugal force ω^2x. Thus, the motion of the bead along the wire, as seen by an observer rotating with the system, will be identical to the motion of the bead of a nonrotating tautochrone of potential $V_*(y)$.

6 Physical Intuition Behind Tautochrones

There is a simple, beautiful explanation for all tautochrone motion. *All tautochrones act like simple linear springs*. Suppose a mass m attached to a spring is displaced a distance A from equilibrium. The time required for the mass to return to the equilibrium position is

$$T=\frac{\pi}{2}\sqrt{\frac{m}{k}} \tag{17}$$

where k is the spring constant. This time is independent of the displacement A and thus we have tautochrone motion in every simple spring. The restoring force from the spring is proportional to the distance s that the spring is stretched: $F=ks$. Solving (17) for k we can write this force as

$$F=\frac{m\pi^2}{4T^2}s. \tag{18}$$

This formula is also valid for any nonrotating tautochrone ([2]).

Let us now examine the forces on the bead (of unit mass) on our rotating tautochrone curve. Starting with relation (2) and replacing the potential by $V_*(y)$ we get

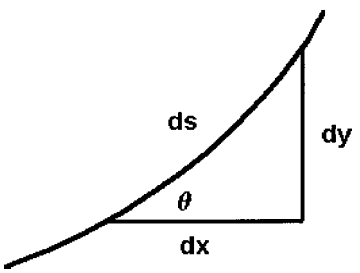


Fig. 2 Geometric meaning of differentials

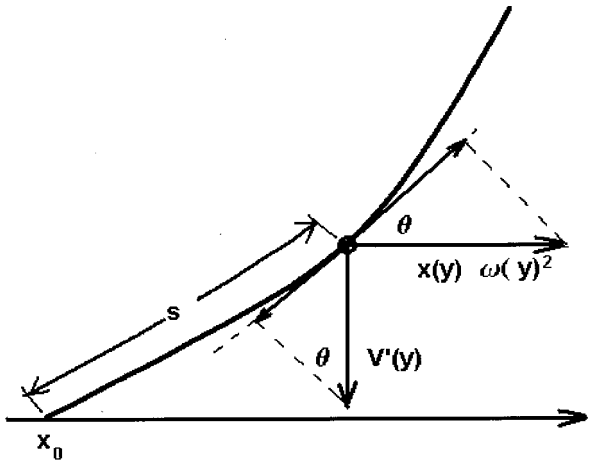


Fig. 3 Why all tautochrones act like simple springs

$$V_*(y)=\frac{\pi^2}{8T^2}s^2.$$

Next use (14) to replace $V_*(y)$ to obtain

$$V(y)+\frac{1}{2}\frac{L^2}{(I+x(y)^2)}-\frac{L^2}{2(I+x_0^2)}=\frac{\pi^2}{8T^2}s^2.$$

Differentiate with respect to y and get

$$V'(y)-\frac{L^2x(y)x'(y)}{(I+x(y)^2)^2}=\frac{\pi^2}{4T^2}s\frac{ds}{dy}.$$

From (12) we see that $\omega(y)^2=L^2/(I+x(y)^2)^2$ and thus the above expression can be written as

$$V'(y)-x(y)\omega(y)^2\frac{dx}{dy}=\frac{\pi^2}{4T^2}s\frac{ds}{dy}.$$

Finally we multiply by dy/ds to get

$$V'(y)\frac{dy}{ds}-x(y)\omega(y)^2\frac{dx}{ds}=\frac{\pi^2}{4T^2}s.$$

From Fig. 2 we see that the derivatives in the above expression can be replaced by $dx/ds=\cos\theta$ and $dy/ds=\sin\theta$.

$$V'(y)\sin\theta-x(y)\omega(y)^2\cos\theta=\frac{\pi^2}{4T^2}s. \tag{19}$$

Using Fig. 3 we see that the term $V'(y)\sin\theta$ is the component of the force generated by the potential in the direction tangent to the tautochrone curve. We also see that the term $x(y)\omega(y)^2\cos\theta$ is the component of centripetal acceleration along the tautochrone curve.

Thus we see that (19) tells us that when the forces acting on our bead are resolved in the direction tangent to the tautochrone curve, then the sum of these forces is proportional to the distance s from the final point on the tautochrone. This is similar to (18) that describes the simple spring. (Notice that with $m=1$ the constants in (18) and (19) are identical.) *The sum of the forces tangent to a tautochrone curve on a moving bead is proportional to the distance measured from the final point to the bead.*

References

[1] Flores, E. V., and Osler, T. J., 1999, ''The Tautochrone Under Arbitrary Potentials Using Fractional Derivatives,'' Am. J. Phys., **67**, pp. 718–722.

[2] McKinley, J. M., 1979, ''Brachistochrones, Tautochrones, Evolutes, and Tesselations,'' Am. J. Phys., **47**, No. 1, pp. 81–86.

Smooth Asymmetric Two-Dimensional Indentation of a Finite Elastic Beam

M. Zhou
Graduate Research Assistant, Civil and Environmental Engineering Department, University of Alabama, Huntsville, AL 35899

W. P. Schonberg
Professor and Chair,
Civil Engineering Department, University of Missouri-Rolla, Rolla, MO 65409
Mem. ASME
e-mail: wschon@umr.edu

Standard methods of beam indentation analysis use a beam theory solution to obtain the load-displacement relationship and a Hertz solution to calculate local stresses. However, when the contact length exceeds the thickness of the beam point contact can no longer be assumed and Hertzian relations are no longer valid. This paper presents an improved superposition solution technique that uses a true elasticity solution to obtain the load-displacement relationship in non-Hertzian indentation problems.
[DOI: 10.1115/1.1352068]

Introduction

In this paper, we present an improved solution to the two-dimensional problem of a finite beam of length L and thickness h that is loaded *asymmetrically* on its upper surface by a frictionless cylindrical indenter (see Fig. 1). Standard methods of indentation analysis use a beam theory solution to obtain an overall load-displacement relationship and then a Hertzian contact solution to calculate local stresses under the indenter. However, previous modeling efforts have shown that the stress distribution in the region of contact will differ significantly from a Hertzian one when the contact length exceeds the thickness of the beam. In such cases, point contact can no longer be assumed and Hertzian relations are not valid.

Problems of this type were solved previously by Keer and Miller [1] and Peck and Schonberg [2] using a GLOBAL/LOCAL approach that superposed beam theory and elasticity expressions. The technique developed was also applied to cantilever beam indentation by Keer and Schonberg [3,4] and subsequently modified to include beam rotation effects by Zhou and Schonberg [5]. A review of the superposition procedure used by Keer, Schonberg, et al. reveals that it has a problematic aspect: it uses an approximate solution (instead of an elasticity solution) to establish the load-displacement relationship at the contact site.

The improved superposition technique presented in this paper addresses this issue by using a static finite layer solution that is a true elasticity solution. Also, in a manner similar to that used by Zhou and Schonberg [5] to model cantilever beam indentation, the rotation of the beam under the indenter is included in the mixed boundary condition at the contact site. As a result, the final solution takes into consideration all of the prescribed boundary and end conditions and describes more accurately the local surface deformation due to contact pressure and the global flexural deformation of the beam. The validity of the solutions presented is assessed by comparing the results obtained to the predictions of modified beam theory solutions.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, May 31, 2000; final revision, Oct. 10, 2000. Associate Editor: M.-J. Pindera.

Generalized Elasticity Solution for a Finite Layer

In this section we present the solution for a finite elastic layer of thickness that is subjected to an arbitrary upper surface pressure distribution. This solution is achieved by the superposition of an elasticity solution for an infinite layer loaded on its upper surface with an elasticity solution for a finite layer subjected to asymmetric bending.

A suitable elasticity solution that represents normal loading on the upper surface of an infinite elastic layer in plane strain with no loading on its lower surface is given by Keer and Miller [1]. To complete the finite elastic layer solution, we introduce the following elasticity solution for the asymmetric bending of a finite layer (thickness h and length L) having end moments M_0 and M_1:

$$\sigma_{yy}=0, \tag{1}$$

$$\sigma_{xy}=\frac{V(x)}{2I}y(h-y), \tag{2}$$

$$\sigma_{xx}=\frac{M(x)}{I}\left(y-\frac{h}{2}\right) \tag{3}$$

$$u_x=\frac{1}{D}\left\{b_0-a_1y+\frac{2-\nu}{12(1-\nu)}\frac{M_1-M_0}{L}[2(h-y)^3-3h(h-y)^2]+\left[M_0x+\frac{M_1-M_0}{2L}(L_0+x)^2\right]\left(y-\frac{h}{2}\right)\right. \tag{4}$$

$$u_y=\frac{1}{D}\left[a_0+a_1x-\frac{1}{2}M_0x^2-\frac{M_1-M_0}{6L}(L_0+x)^3\right]-\frac{\nu}{2(1-\nu)D}M(x)(y^2-hy) \tag{5}$$

$$\overline{\theta}(x)\equiv\frac{1}{h}\int_0^h\frac{\partial u_y}{\partial x}dy=\frac{1}{D}\left[a_1-M_0x-\frac{M_1-M_0}{2L}(L_0+x)^2\right]+\frac{\nu h^2}{12(1-\nu)D}\frac{M_1-M_0}{L} \tag{6}$$

$$M(x)=M_0+\frac{M_1-M_0}{L}(L_0+x), \tag{7}$$

$$V(x)=\frac{M_1-M_0}{L} \tag{8}$$

where L_0, L, and L_1 are as shown in Fig. 1, $D=\mu h^3/6(1-\nu)$, $I=h^3/12$, M_0, M_1 are end bending moments at $x=-L_0$ and at $x=L_1$, respectively, and a_0, b_0, a_1 are (as yet) unknown constants. This solution is a superposition of the Airy stress function solutions for asymmetric bending and for pure shear.

Superposing Eqs. (1)–(8) with the corresponding expressions for stresses, displacements, etc., in Keer and Miller [1] yields the generalized elasticity solution for a finite length isotropic elastic layer subjected to an arbitrary upper surface loading. In the next section, we apply appropriate end conditions to this solution to obtain the desired beam indentation problem solutions.

Asymmetric Beam Indentation Model

The mixed boundary value problems to be solved in this section are those of a simply supported and a fixed-fixed elastic beam that are indented by a cylindrical punch on their upper surface (see again Fig. 1). The solutions of these problems are achieved by application of the mixed boundary conditions at the contact sites

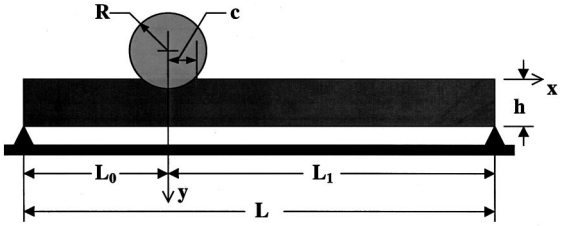


Fig. 1 Indentation of a finite layer

and the appropriate end support conditions in the generalized elasticity solution presented in the preceding section. As in Zhou and Schonberg [5], the boundary condition for both types of beams at contact sites is written as follows:

$$u_y(x,0)=\Delta+\theta_0 x-\frac{x^2}{2R} \quad |x|<c \tag{9}$$

where Δ and θ_0 are the beam upper surface displacement and rotation under the indenter. The two types of end support conditions that need to be satisfied are (1) simple supports (zero moment and displacement at the beam ends), and (2) fixed ends (zero slope and displacement at the fixed ends). By applying these end support conditions to appropriate expressions in the generalized elasticity solution, we obtain a system of four equations for the four unknowns a_0, a_1, M_0, and M_1 for each beam type. Once solved, the expressions for these four quantities are used in applying the mixed boundary condition at contact site given by Eq. (9). Following the approach used by Keer, Schonberg, et al. we obtain two coupled Fredholm integral equations of the second kind. These equations have the following forms:

$$\frac{h^3}{6}\psi(x)+\int_0^c \psi(t)K_1(x,t)dt+\int_0^c \phi(t)K_2(x,t)dt+f(x)=-\frac{Dx}{R} \tag{10}$$

$$\frac{h^3}{6}\phi(x)+\int_0^c \phi(t)K_3(x,t)dt+\int_0^c \psi(t)K_4(x,t)dt+g(x)=0 \tag{11}$$

where the kernels K_1 through K_4 and the functions $f(x)$, $g(x)$ are given as follows:

Simply Supported Ends.

$$K_1(x,t)=-\int_0^\infty\left[\frac{h^3}{6}\left(\frac{\beta+sh\beta ch\beta}{\beta^2-sh^2\beta}+1\right)\xi x J_0(\xi x)+\frac{x}{2}\frac{\cos(\xi L_0)+\cos(\xi L_1)}{\xi^2}\right]J_0(\xi t)d\xi+\frac{\pi(L_1-L_0)^2}{4L}x \tag{12}$$

$$K_2(x,t)=\frac{\pi(L_1-L_0)}{4L}xt \tag{13}$$

$$K_3(x,t)=\frac{h^3}{6}\left(\frac{1}{t}\right)+\frac{h^3}{6}\int_0^\infty\left(\frac{\beta+sh\beta ch\beta}{\beta^2-sh^2\beta}+1\right)\times[J_0(\xi x)-\xi x J_1(\xi x)-1]J_1(\xi t)d\xi-\frac{3\pi x^2}{8L}t \tag{14}$$

$$K_4(x,t)=-\frac{3\pi x^2}{8L}(L_0-L_1) \tag{15}$$

$$f(x)=0, \tag{16}$$

$$g(x)=0 \tag{17}$$

Fixed-Fixed Ends.

$$K_1(x,t)=-\int_0^\infty\left\{\frac{h^3}{6}\left(\frac{\beta+sh\beta ch\beta}{\beta^2-sh^2\beta}+1\right)\xi x J_0(\xi x)-\frac{h^3}{6L}\frac{x}{\beta-sh\beta}[\sin(\xi L_0)+\sin(\xi L_1)]\right\}J_0(\xi t)d\xi-\frac{3-2\nu}{1-\nu}\frac{\pi h^2}{12L}x \tag{18}$$

$$K_2(x,t)=\frac{h^3}{6L}\int_0^\infty\frac{x}{\beta-sh\beta}[\cos(\xi L_0)-\cos(\xi L_1)]J_1(\xi t)d\xi \tag{19}$$

$$K_3(x,t)=\frac{h^3}{6}\left(\frac{1}{t}\right)+\int_0^\infty\frac{h^3}{6}\left(\frac{\beta+sh\beta ch\beta}{\beta^2-sh^2\beta}+1\right)\times[J_0(\xi x)-\xi x J_1(\xi x)-1]J_1(\xi t)d\xi \tag{20}$$

$$K_4(x,t)=0 \tag{21}$$

$$f(x)=\frac{x}{2L}(L_1-L_0)(M_1-M_0), \tag{22}$$

$$g(x)=-\frac{3}{4}\frac{x^2}{L}(M_1-M_0) \tag{23}$$

The functions $\psi(t)$ and $\phi(t)$ are related to the symmetric and anti-symmetric components of the surface loading as follows:

$$P=-\pi\int_0^c \psi(t)dt, \tag{24}$$

$$M=-\pi\int_0^c t\phi(t)dt. \tag{25}$$

Once Eqs. (10),(11) are solved to obtain $\psi(x)$ and $\phi(x)$, all necessary quantities can also be readily obtained. The actual solution of Eqs. (10),(11) is performed numerically in nondimensional form. This transformation is obtained using the following nondimensional parameters: L_0/h, L/h, t/c, x/c, y/h, $Rh^3\psi(x)/Dc$, and $Rh^3\phi(x)/Dc$, and $R\Delta/h^2$. To assess the validity of the elasticity solutions, their predictions for beam displacement are compared with the predictions of beam theory solutions that use as input the contact pressure generated by the elasticity solutions. These solutions are given as follows:

Simply Supported Ends.

$$\Delta_{SS}=\frac{L_0L_1}{3DL}[L_0L_1P+(L_1-L_0)M]+\frac{1}{D}\int_0^c\left(\frac{\pi L_0L_1}{4L}-\frac{t}{9}\right)t^2\psi(t)dt+\frac{\pi(L_1-L_0)}{32DL}\int_0^c t^3\phi(t)dt \tag{26}$$

Fixed-Fixed Ends.

$$\Delta_{FF}=\Delta_{SS}-\frac{L_0^2L_1^2}{3DL^3}(L^2-L_0L_1)P-\frac{L_0L_1}{6DL^3}(L_1-L_0)\times[L^2+(L_1-L_0)^2+L_0L_1]M-\frac{\pi L_0L_1}{12DL^3}[L^2+2(L_1-L_0)^2+2L_0L_1]\int_0^c t^2\psi(t)dt+\frac{\pi L_0L_1(L_1-L_0)}{16DL^3}\int_0^c t^3\phi(t)dt \quad (27)$$

where Δ_{SS} is given by Eq. (26), and P, M are again given by Eqs. (24),(25), respectively.

Results and Discussion

Solutions to the two types of indentation problems were obtained for $c/h=0.25$, 0.5, 1.0 and 2.0, $L/h=10.0$ and 20.0, and for each L/h value, $2L_0/L=1.0$, 1.5, and 1.7 (where $2L_0/L=1.0$ corresponds to the case of central indentation studied by Keer and Miller [1]). We note that for $2L_0/L=1.7$ and $L/h=10$, $c/h=2.0$ would imply that the contact length would extend past the support; hence, $c/h=2.0$ was not considered in this case. For the fixed-fixed beam indentation problem, all calculations are performed with a Poisson's ratio of 0.3. The results of this parametric study were compared with values obtained using the models developed by Keer and Miller [1] and Peck and Schonberg [2] that did not include upper-surface rotation effects.

For small values of c/h (i.e., $c/h\leqslant 0.5$), the predictions of peak contact stress values by the various solutions agreed quite well. However, for $c/h\geqslant 1$ the peak stress predictions were found to occasionally differ significantly. This occurred because as c/h increases, the effects of beam upper surface rotation effects become more pronounced. While the current model incorporates those effects, the previous models do not. Therefore, the various models will match more closely for smaller contact areas and indenter locations that result in minimal rotation effects. For larger contact areas and for indenter locations near beam ends, peak contact stress values predicted by the various models differed by as much as 10–15 percent for both types of beam supports. Finally, we note that while there were differences in peak contact stress values, the overall shapes of the contract stress distributions were very similar to those presented by Keer and Miller [1] and by Peck and Schonberg [2].

A comparison between the predictions of beam displacement under the indenter generated by the improved solutions developed herein and the predictions of the beam theory solutions showed that the results of the two solutions agreed very well, and that the agreement improved as L/h increased and c/h decreased. This can be explained by the following considerations. First, as L/h increases, the effects of shear deformation on beam response become negligible. Second, as c/h decreases, the local effects of beam upper surface rotation become negligible. In both cases, while the elasticity solution incorporates those effects, the beam theory solutions do not. Therefore, the two solutions will match more closely for indenter locations that result in minimal shear deformations and for smaller c/h values.

Differences between the displacement predictions of the elasticity solution developed herein and the predictions obtained using the GLOBAL/LOCAL models developed previously by Keer and Miller [1] and Peck and Schonberg [2] showed fairly minimal differences for small c/h values and for most $2L_0/L$ values (typically less than five percent). This was expected because (1) the governing equations of motion for the solution developed herein and the solutions developed by Keer and Miller [1] and Peck and Schonberg [2] are identical for the case of simply supported ends, and (2) beam upper surface rotation effects on beam response are minimal when for small c/h values. Furthermore, in the case of fixed-fixed beams, as $2L_0/L\rightarrow 1$ and L/h increases, the condition of zero rotation angle at the fixed end supports has less of an effect on the contact zone. Therefore, the effects of second order shearing effects on the average beam rotation θ for the new model become negligible. While the current model incorporates those effects, the previous models do not. Therefore, the newly developed and previous models will match more closely for smaller contact area and indenter locations that result in minimal rotation and shearing effects.

Figure 2 shows the nondimensional load-displacement curves for $L/h=10$ and 20 fixed-fixed beams where $2L_0/L=1.7$. The

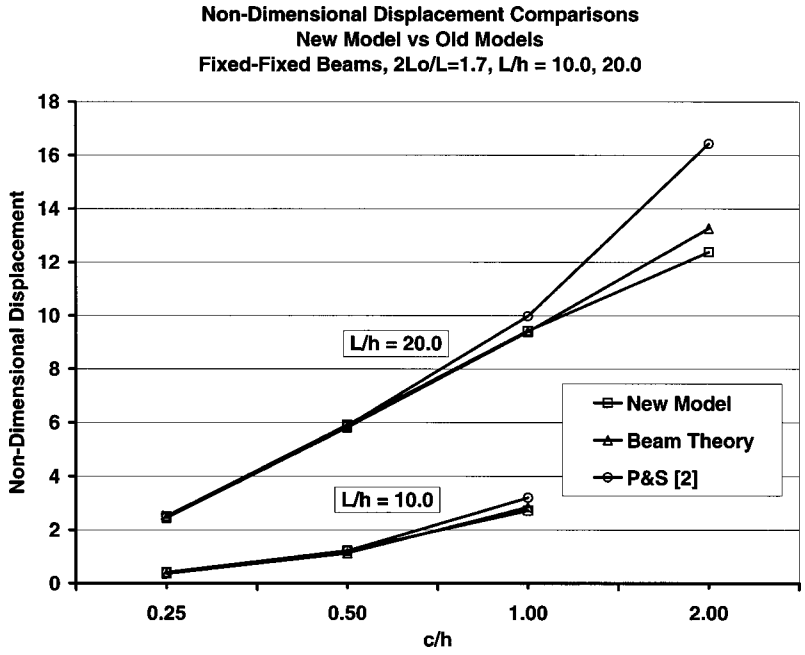


Fig. 2 Nondimensional displacement comparisons, new model versus old models (Keer and Miller [1]; Peck and Schonberg [2]), fixed-fixed beams, $2L_0/L=1.7$, $L/h=10.0$ and 20.0

larger differences in some of the indentation scenarios shown in this figure can probably be attributed to limitations in the numerical integrations of the infinite integrals that appeared in the work performed by Peck and Schonberg [2]. According to Peck and Schonberg, it was not possible to integrate the kernels in the fixed-fixed beam equations to the same accuracy as those in the simply supported cases. This was due to the fact that the kernels for simply supported beam equations converged at a rate of $1/\xi^2$, while those for the fixed-fixed beams converged at the much slower rate of $1/\xi$. In our solution, we have overcome this numerical integration limitation by using exact integration results for expressions involving Bessel functions. This allowed the kernels in the fixed-fixed beam equations to converge at the much faster rate of $1/sh\xi$.

Interestingly enough, the value of R (the radius of the indenter) did not appear to have any bearing on the solution of the governing equations for this problem. That is, the nondimensional stress and displacement values calculated using the model developed herein would appear to be valid for all values of R! Mathematically, this occurs because of the nondimensional scheme: in this scheme, all traces of R are removed by the nondimensionalization process. However, *R does* have an effect when the nondimensional values of stress and displacement are transformed into real values. While all of the nondimensional predictions of the model may be mathematically possible, they may *not* be physically attainable for some values of R. For example, the case of $c/h=2$ is clearly more readily attainable for very large values of R (as compared to h), and significantly less so (if not impossible) for small values of R (as compared to h).

Summary and Conclusions

The static and dynamic indentation of beams and plates continues to be an intriguing problem, especially for scenarios in which large area contact can be expected to occur. The solution scheme presented herein is a refinement of the approach originally developed for the relatively simple problem of central beam indentation. We found that the results provided by this refinement are, in most cases, not significantly different from those that can be obtained using existing GLOBAL/LOCAL analysis techniques. However, the changes introduced into the modeling process should allow this refined solution technique to more accurately predict internal stress fields due to upper surface indentations. This capability is critical in the case of composite beams or plates where internal damage can appear prior to any evidence of damage on the external beam or plate surface.

References

[1] Keer, L. M., and Miller, G. R., 1983, ''Smooth Indentation of a Finite Layer,'' J. Eng. Mech., **109**, pp. 706–717.

[2] Peck, J. A., and Schonberg, W. P., 1993, ''Asymmetric Indentation of a Finite Elastically Supported Beam,'' J. Appl. Mech., **60**, pp. 1039–1045.

[3] Keer, L. M., and Schonberg, W. P., 1986, ''Smooth Indentation of an Isotropic Cantilever Beam,'' Int. J. Solids Struct., **22**, pp. 87–106.

[4] Keer, L. M., and Schonberg, W. P., 1986, ''Smooth Indentation of a Transversely Isotropic Cantilever Beam,'' Int. J. Solids Struct., **22**, pp. 1033–1053.

[5] Zhou, M., and Schonberg, W. P., 1995, ''Rotation Effects in the Global/Local Analysis of Cantilever Beam Contact Problems,'' Acta Mech., **108**, pp. 49–62.