Transactions of the ASME®

Technical Editor, **L. B. FREUND**

Division of Engineering
Brown University
Providence, R.I. 02912

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.) for $105 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017. Second class postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to The Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.
CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.

PRICES: To members, $24.00, annually; to nonmembers, $105.00.
Add $6.00 for postage to countries outside the United States and Canada.

INDEXED by the Engineering Information

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 54 • NUMBER 3 • SEPTEMBER 1987

(Contents continued on Inside Back Cover)

CONTENTS (CONTINUED)

BRIEF NOTES

DISCUSSION

A Doubly Asymptotic, Nonreflecting Boundary for Ground-Shock Analysis[1]

I. C. Mathews[2]
Consultant.

T. L. Geers[3]
Manager.
Fellow ASME

Applied Mechanics Laboratory,
Lockheed Palo Alto Research Laboratory,
Palo Alto, CA 94304

This paper desribes the formulation and implementation of a nonreflecting boundary *for use with existing finite-element codes to perform nonlinear ground-shock analyses of buried structures. The boundary is based on a first-order* doubly asymptotic approximation (DAA_1) *for disturbances propagating outward from a selected portion of the soil medium surrounding the structure of interest. The resulting set of first-order ordinary differential equations is then combined with the second-order equations of motion for the finite-element model so as to facilitate solution by a* staggered solution procedure. *This procedure is shown to be computationally stable as long as the time increment is smaller than a limiting value based on the finite-element mass matrix and the DAA-boundary stiffness matrix. Computational results produced by the boundary are compared with exact results for linear canonical problems pertaining to infinite-cylindrical and spherical shells.*

1 Introduction

The primary objective of this effort has been the implementation of a *nonreflecting boundary* for use with existing finite-element codes to perform nonlinear ground-shock analyses of buried structures. This boundary is based on the *first-order doubly asymptotic approximation* (DAA_1) for elastodynamic scattering (Geers and Yen, 1981; Underwood and Geers, 1981). In addition, a staggered solution procedure is utilized to partition the global equations in order to achieve both computational efficiency and software modularity (Felippa and Park, 1980).

This work extends that of Underwood and Geers (1981) for linear ground-shock problems, wherein the DAA surface is placed *on the surface* of the buried structure. Here, the DAA surface is moved some distance *out from the surface* of the structure, enclosing both the structure and a portion of the surrounding soil medium, which may be treated with nonlinear finite elements. Other extensions include formulation and implementation for general two-dimensional and three-dimensional problems, improved discretization of the DAA surface with higher-order interpolation functions, and utilization of a conditionally stable staggered solution procedure.

It is important to differentiate between doubly asymptotic approximations, which address quasistatic and wave-propagation effects simultaneously, and singly asymptotic approximations, which address these effects separately (see, e.g., various papers in Kalinowski, ed., 1981, Datta, ed., 1982, and Cohen and Jennings, 1983). For example, representation of the external medium by an elastic foundation, which may be quite satisfactory at low frequencies, does not account, at higher frequencies, for energy dissipation through outward propagation of scattered waves. On the other hand, representation of the external medium by a viscous boundary, which may be quite satisfactory for wave-propagation problems, does not provide elastic restoring forces in the static limit.

A response-averaging method originally proposed by Smith (1974) and extended by Cundall et al. (1978) also fails in the static limit. For example, consider the response of a rigid structure surrounded by an infinite, linear-elastic medium to an internal, quasi-static point force. A computational model for this problem might consist of the rigid structure surrounded by a portion of the medium enclosed by a nonreflecting boundary. If this boundary is that of Smith, the total response of the structure is the *average* of two responses, one associated with the structure and bounded portion of medium enclosed by a rigid boundary, and the other associated with the structure and bounded portion of medium floating freely in space. Unfortunately, the latter response grows indefinitely in the static limit because the freely floating system is not in static equilibrium. In contrast, doubly asymptotic approximations approach exactness in the static limit.

2 Governing Equations

This section presents the governing equations for the finite-element (FE) model of the structure along with a portion of the surrounding soil medium, and for the boundary-element model (BE) of the nonreflecting DAA surface. These equa-

[1]Presented at the ASME Winter Annual Meeting, Miami, Fla., November 17-21, 1985.
[2]Currently Lecturer, Department of Aeronautics, Imperial College, London SW7 2BY.
[3]Currently Professor and Chairman, Department of Mechanical Engineering, University of Colorado, Boulder, CO 80309.
Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May 6, 1986; final revision August 15, 1986.

tions are then partitioned, and a staggered-solution procedure is introduced to solve for transient response. Throughout the development, the dependence of excitation and response quantities on time is implicit.

2.1 Finite-Element/Boundary-Element Model. Let $\mathbf{x}$ be the *computational vector* of displacement response in global coordinates for the FE model of the structure and a portion of surrounding medium. The governing equations for the finite-element model are then (see, e.g., Zienkiewicz, 1977)

$$\mathbf{M}_s\ddot{\mathbf{x}} + \mathbf{D}_s\dot{\mathbf{x}} + \mathbf{K}_s\mathbf{x} = \mathbf{f}_e + \mathbf{f}_i \tag{2.1}$$

where $\mathbf{M}_s$, $\mathbf{D}_s$, and $\mathbf{K}_s$ are the mass, damping, and stiffness matrices, respectively, for the FE model, $\mathbf{f}_e$ is the computational vector of external medium forces imposed by the DAA surface, and $\mathbf{f}_i$ is the vector of internal nonlinear forces; as usual, a dot denotes differentiation in time. Compatibility of forces and displacements at the DAA surface may be expressed as (Underwood and Geers, 1981)

$$\begin{aligned} \mathbf{f}_e &= -\mathbf{G}\mathbf{g} \\ \mathbf{u} &= \mathbf{G}^t\mathbf{x} \end{aligned} \tag{2.2}$$

where the superscript t denotes transpose, where $\mathbf{g}$ and $\mathbf{u}$ are the global force and displacement vectors, respectively, for the BE model of the DAA surface, and where $\mathbf{G}$ is the force-transformation matrix when moving from BE to FE coordinates.

Now the force vector $\mathbf{g}$ and displacement vector $\mathbf{u}$ may be decomposed into incident-wave and scattered-wave components as

$$\begin{aligned} \mathbf{g} &= \mathbf{g}_I + \mathbf{g}_S \\ \mathbf{u} &= \mathbf{u}_I + \mathbf{u}_S \end{aligned} \tag{2.3}$$

where $\mathbf{g}_I$ is the *known* force vector associated with a free-field incident wave and $\mathbf{g}_S$ is the *unknown* force vector associated with the wave scattered by the structure. It is worth noting that this dual decomposition does not require constitutive linearity of the medium to be valid, for $\mathbf{g}_S$ and $\mathbf{u}_S$ may each be viewed as merely the difference between two vectors, one obtaining with the structure absent and the other obtaining with the structure present.

2.2 Doubly Asymptotic Approximation. A first-order DAA is used here to relate the scattered-force vector $\mathbf{g}_S$ and the scattered-displacement vector $\mathbf{u}_S$ (Geers and Yen, 1981; Underwood and Geers, 1981). this approximation approaches exactness in both the high and low-frequency limits, and effects a smooth transition between. The development of DAA_1 for a linear, isotropic external medium proceeds as follows.

At *high frequencies*, the *geometrical vector* of scattered-wave surface tractions for the DAA surface corresponding to normal and tangential motions of that surface is given by

$$t_s'(p) = \rho_m C_m \dot{u}_s'(p) \tag{2.4}$$

where p denotes a point on the surface, ρ_m is the mass density of the medium, and C_m is the diagonal sound-speed matrix corresponding to $\dot{u}_s'$, which is geometrical vector of normal and tangential scattered-wave velocities. For the component of $\dot{u}_s'$ normal to the DAA surface, the corresponding matrix component is the dilatational velocity, while for each component of $\dot{u}_s'$ tantential to the DAA surface, the corresponding matrix component is the shear velocity.

Now the local-coordinate vectors of equation (2.4) may be transformed into global-coordinate vectors as

$$u_s'(p) = Q(p)u_s(p), \quad t_s'(p) = Q(p)t_s(p) \tag{2.5}$$

to obtain, inasmuch as $Q^{-1} = Q^t$, where the superscripts -1 and t denote inverse and transpose, respectively,

$$t_s(p) = Q^t(p)\rho_m C_m Q(p)\dot{u}_s(p) \tag{2.6}$$

Hence boundary-element discretization of u_s as (see, e.g., Zienkiewicz, 1977)

$$u_s(p) = \mathbf{N}(p)\ \mathbf{u}_s \tag{2.7}$$

where $\mathbf{N}(p)$ is a matrix of shape-functions and $\mathbf{u}_s$ is a vector of displacement degrees of freedom, and defintion of the high-frequency scattered-wave force vector as

$$\mathbf{g}_s^h = \int \mathbf{N}^t(p)\ t_s(p)dS \tag{2.8}$$

yield, for high-frequency motions,

$$\mathbf{g}_s^h = \mathbf{D}_m\dot{\mathbf{u}}_s \tag{2.9}$$

in which

$$\mathbf{D}_m = \int \mathbf{N}^t Q^t \rho_m C_m Q\ \mathbf{N}\ dS \tag{2.10}$$

At *low frequencies*, the computational vector for scattered-wave forces is given by the quasi-static relation

$$\mathbf{g}_s^l = \mathbf{K}_m\mathbf{u}_s \tag{2.11}$$

where $\mathbf{K}_m$ is a full, nonsymmetric stiffness matrix for the boundary-element mesh, whose construction is described in the next section.

Finally, the first-order doubly asymptotic approximation DAA_1 is formed by the superposition of $\mathbf{g}_s^l$ and $\mathbf{g}_s^h$ to obtain

$$\mathbf{g}_s = \mathbf{D}_m\dot{\mathbf{u}}_s + \mathbf{K}_m\mathbf{u}_s \tag{2.12}$$

It is clear that, at high frequencies where $\dot{\mathbf{u}}_s \gg \mathbf{u}_s$, this equation approaches (2.9), and that, at low frequencies where $\mathbf{u}_s \gg \dot{\mathbf{u}}_s$, it approaches equation (2.11); hence equation (2.12) is doubly asymptotic. At intermediate frequencies, equation (2.12) constitutes an approximation whose accuracy may be explored by solving canonical steady-state problems; this has, in fact, been done for an acoustic medium by Geers (1978). Unfortunately such explorations have not proven very useful in assessing DAA performance in transient problems because the frequency content of the excitation dictates so strongly the frequency content of the response. Much more useful has been the comparison of DAA and exact solutions for canonical transient problems, as is done later in this paper.

Now the assumption embodied in DAA_1 of a constitutively linear medium for the scattered wave is justified within the framework of classical plasticity theory if the material point for every exterior location, i.e., every location in the medium outside the DAA surface, remains within its corresponding yield surface when and after the scattered wave arrives at the DAA surface. For incident waves with sufficiently rapid decay rates and for a DAA surface sufficiently removed from the surface of the structure, the scattered wave causes minor perturbations about an elastic state at each exterior location, thereby satisfying the preceding condition.

The assumption of material isotropy outside the DAA surface cannot be rigorously maintained if the material has suffered plastic excursions in response to the incident wave. However, it is likely that the resulting anisotropy is no more pronounced than that characterizing the ambient state, which is generally uncertain in practical cases. Hence, while an extension to material orthotropy may be theoretically possible, it may not be worth the trouble.

2.3 Response Equations. Introduction of the first of equations (2.2) and (2.3) into (2.1) and of the second of equations (2.2) and (2.3) into (2.12) yields the *doubly asymptotic equations of motion*

$$\begin{aligned} \mathbf{M}_s\ddot{\mathbf{x}} + \mathbf{D}_s\dot{\mathbf{x}} + \mathbf{K}_s\mathbf{x} &= -\mathbf{G}\{\mathbf{g}_I + \mathbf{g}_S\} + \mathbf{f}_i \\ \mathbf{g}_S &= \mathbf{D}_m\{\mathbf{G}^t\dot{\mathbf{x}} - \dot{\mathbf{u}}_I\} + \mathbf{K}_m\{\mathbf{G}^t\mathbf{x} - \mathbf{u}_I\} \end{aligned} \tag{2.13}$$

which may be numerically integrated in time to obtain the solution vectors $\mathbf{x}$ and $\mathbf{g}_S$. Because $\mathbf{M}_s$, $\mathbf{D}_s$, and $\mathbf{K}_s$ are typically large and banded, while $\mathbf{K}_m$ is relatively small and full, it is not computationally practical to introduce the second of these equations into the first to eliminate $\mathbf{g}_S$.

However, because $\mathbf{D}_m$ is banded and multiplies the highest-derivative terms in the second of equations (2.13), it is advantageous to apply the technique of *augmentation* (Park et al., 1977), which here merely involves introducing the second of equations (2.13) into the first, moving the term containing $\mathbf{D}_m$ to the left side of the resulting set of equations, and keeping $\mathbf{D}\,\mathbf{K}_m\mathbf{G}^t\mathbf{x}$ on the right. This yields the *augmented doubly asymptotic equations of motion*

$$\mathbf{M}_s\ddot{\mathbf{x}}+[\mathbf{D}_s+\mathbf{G}\mathbf{D}_m\mathbf{G}^t]\dot{\mathbf{x}}+\mathbf{K}_s\mathbf{x}=-\mathbf{G}\mathbf{g}_I+\mathbf{G}\mathbf{D}_m\dot{\mathbf{u}}_I+\mathbf{G}\mathbf{K}_m\mathbf{u}_I+\mathbf{f}_i-\mathbf{G}\mathbf{K}_m\mathbf{G}^t\mathbf{x} \tag{2.14}$$

which are highly amenable to staggered solution, as discussed in Section 4.

3 Medium Stiffness Matrix

This section describes the construction of the boundary-element stiffness matrix that relates the scattered-wave force and displacement vectors at low frequencies. The development is based on Somigliana's identities, which derive from Betti's reciprical work theorems and Kelvin's problem of a point load in an infinite elastic medium (see, e.g., Kupradze, 1964; Rizzo, 1967; Cruse, 1969; Lachat and Watson, 1976).

3.1 Elastostatic Boundary-Integral Equations. The surface behavior of an elastic medium, whether occupying an exterior or interior region, may be expressed as (Rizzo, 1967; Cruse, 1969)

$$c(p)u(p)+\int_\Gamma T(p,q)u(q)d\Gamma_q=\int_\Gamma U(p,q)t(q)d\Gamma_q \tag{3.1}$$

where p is a point on the boundary and q is the integration variable, and where $u(p)$ and $t(p)$ are $d\times 1$ vectors ($d=2$ or 3) of medium displacements and tractions in Cartesian coordinates on the boundary at p. The elements $T_{ij}(p,q)$ and $U_{ij}(p,q)$ of the $d\times d$ matrices $T(p,q)$ and $U(p,q)$ are fundamental solutions for the tractions and displacements at a location q in the direction i due to a point load at location p in direction j. With δ_{ij} as the Kronecker symbol, each element of the matrix c is defined as

$$c_{ij}(p)=\tfrac{1}{2}\delta_{ij} \tag{3.2}$$

if there exists a continuous tangent at p, or, with Γ_ϵ as the surface of a sphere of radius ϵ centered at p,

$$c_{ij}(p)=\lim_{\epsilon\to 0}\int_{\Gamma_\epsilon}T_{ij}(p,q)d\Gamma_q \tag{3.3}$$

if the tangent is not continuous.

Now an element of the *two-dimensional* displacement-kernel matrix $U(p,q)$ for plane-strain problems is given by

$$U_{ij}(p,q)=\frac{-1}{8\pi(1-\nu)G}[(3-4\nu)ln(r)\delta_{ij}-r_{,i}r_{,j}] \tag{3.4}$$

where G and ν are the shear modulus and Poisson's ratio, respectively, and $r=r(p,q)$ is the distance between the load point p and the field point q; the derivatives are taken with reference to the coordinates of q. With p_i and q_i as the coordinates of p and q, respectively,

$$r_i=q_i-p_i,\quad r=(r_ir_i)^{1/2},\quad r_{,i}=\frac{q_i-p_i}{r} \tag{3.5}$$

In contrast, an element of the *three-dimensional* displacement-kernel matrix $U(p,q)$ is given by

$$U_{ij}(p,q)=\frac{1}{16\pi(1-\nu)Gr}[(3-4\nu)\delta_{ij}+r_{,i}r_{,j}] \tag{3.6}$$

Finally, an element of the traction-kernel matrix $T(p,q)$ for *both two and three-dimensional* problems is given by

$$T_{ij}(p,q)=\frac{-1}{4\alpha\pi(1-\nu)r^\alpha}\{[(1-2\nu)\delta_{ij}+\beta r_{,i}r_{,j}]r_{,i}n_i-(1-2\nu)(r_{,i}n_j-r_{,j}n_i)\} \tag{3.7}$$

where n_i and n_j are direction-cosines for the surface normal at q. The two and three-dimensional forms are explicitly obtained by letting $\alpha=1$, 2 and $\beta=2$, 3, respectively.

3.2 Discretization. Numerical solution of the integral equation (3.1) requires discretization of the DAA surface, over each boundary element of which the displacement and traction vectors are approximated. The curved isoparametric elements of finite-element theory offer both the generality and the accuracy needed for this purpose. With this approach, the global Cartesian coordinates of any point in an element are taken as related to the nodal coordinates by (cf (2.7))

$$x(p)=\mathbf{N}(p)\ \mathbf{x} \tag{3.8}$$

i.e., the same shape functions are used to approximate element geometry, displacements, and tractions. This allows interpolated displacements and tractions along the DAA curve in two-dimensional space to be integrated over a normalized length in ξ-coordinate space, and similar quantities over the DAA surface in three-dimensional space to be integrated over a standard 2×2 normalized square in ξ_1,ξ_2-coordinate space.

On an element-by-element basis, equation (3.8) becomes

$$x^e(\xi^e)=\sum_k N_k(\xi^e)\ \mathbf{x}_k^e \tag{3.9}$$

where $x^e(\xi^e)$ is the $d\times 1$ vector of Cartesian coordinates of a point in element e, the $N_k(\xi^e)$ are the element shape functions, and $\mathbf{x}_k^e$ is the $d\times 1$ vector of Cartesian coordinates of the kth element node; also, $\xi^e=\xi^e$ in 2-D, but $\xi^e=\xi_1^e,\xi_2^e$ in 3-D. The elements used in this study are the three-noded, quadratic, curved element for 2-D anaysis and the eight-noded, quadratic, serendipity element for 3-D analysis. The shape functions for the three-noded quadratic element are

$$N_1=\tfrac{1}{2}\xi(\xi-1),\quad N_2=1-\xi^2,\quad N_3=\tfrac{1}{2}\xi(\xi+1) \tag{3.10}$$

where $\xi\in[-1,1]$; the nodes are located at $\xi=-1,0,1$. The shape functions for the eight-noded quadratic element are

$$\begin{aligned}
N_1&=-\tfrac{1}{4}(1-\xi_1)(1-\xi_2)(1+\xi_1+\xi_2)\\
N_2&=\tfrac{1}{2}(1-\xi_1^2)(1-\xi_2)\\
N_3&=\tfrac{1}{4}(1+\xi_1)(1-\xi_2)(\xi_1-\xi_2-1)\\
N_4&=\tfrac{1}{2}(1+\xi_1)(1-\xi_2^2)\\
N_5&=\tfrac{1}{4}(1+\xi_1)(1+\xi_2)(\xi_1+\xi_2-1)\\
N_6&=\tfrac{1}{2}(1-\xi_1^2)(1+\xi_2)\\
N_7&=\tfrac{1}{4}(1-\xi_1)(1+\xi_2)(-\xi_1+\xi_2-1)\\
N_8&=\tfrac{1}{2}(1-\xi_1)(1-\xi_2^2)
\end{aligned} \tag{3.11}$$

where $\xi_1\in[-1,1]$ and $\xi_2\in[-1,1]$, and all nodes lie at the intersections of the $\xi_1=-1,0,1$ and the $\xi_2=-1,0,1$ lines, except at 0, 0, where there is no node.

3.3 Matrix Assembly. With DAA-surface coordinates, displacements and tractions approximated as

$$x(p)=\mathbf{N}(p)\ \mathbf{x},\quad u(p)=\mathbf{N}(p)\ \mathbf{u},\quad t(p)=\mathbf{N}(p)\ \mathbf{t} \tag{3.12}$$

equation (3.1) may be expressed at a node P as

$$c(P)u(P) + \sum_{e=1}^{E} \int_{\Gamma_e} T(P,q[\xi^e]) \sum_k N_k(\xi^e)\mathbf{u}_k^e J(\xi^e)d\xi^e \tag{3.13}$$
$$= \sum_{e=1}^{E} \int_{\Gamma_e} U(P,q[\xi^e]) \sum_k N_k(\xi^e)\mathbf{t}_k^e J(\xi^e)d\xi^e$$

where E is the total number of elements on the DAA surface and $J(\xi^e)$ is the Jacobian for $\mathbf{x}^e$:ξ^e transformation; also, $d\xi^e = d\xi^e$ in 2-D, but $d\xi^e = d\xi_1^e d\xi_2^e$ in 3-D. Finally, coalescence of element contributions at common nodes is implicit in (3.13). The numerical techniques used to evaluate the integrals in this equation are discussed in Takahasi and Mori (1976), Burton (1976), Lachat and Watson (1976), and Mathews and Geers (1985).

The enforcement of equation (3.13) at every node on the DAA surface yields a set of simultaneous algebraic equations that can be expressed in the form

$$\mathbf{A}\,\mathbf{u} = \mathbf{B}\,\mathbf{t} \tag{3.14}$$

so that

$$\mathbf{t} = \mathbf{B}^{-1}\mathbf{A}\,\mathbf{u} \tag{3.15}$$

Now the nodal force vector $\mathbf{g}$ corresponding to a traction distribution t on the DAA surface is given by

$$\mathbf{g} = \int_\Gamma \mathbf{N}^t(p)\ t(p)\ d\Gamma \tag{3.16}$$

Introduction of the third of equations (3.12) and of (3.15) into this relation then yields

$$\mathbf{g} = \mathbf{K}_m\mathbf{u} \tag{3.17}$$

where the generally nonsymmetric medium stiffness matrix $\mathbf{K}_m$ is given by

$$\mathbf{K}_m = \left[\int_\Gamma \mathbf{N}^t\mathbf{N}d\Gamma\right]\mathbf{B}^{-1}\mathbf{A} \tag{3.18}$$

A symmetric form may be obtained as

$$\hat{\mathbf{K}}_m = \tfrac{1}{2}(\mathbf{K}_m + \mathbf{K}_m^t) \tag{3.19}$$

which is identical to that derived from energy considerations (Zienkiewicz et al., 1977). As indicated in the Appendix, however, the use of $\hat{\mathbf{K}}_m$ generally yields numerical results inferior to those produced by $\mathbf{K}_m$.

4 Staggered Solution Procedure

In the interest of computational efficiency, the augmented doubly asymptotic equations of motion given by equation (2.14) are solved with a staggered solution procedure. The procedure is conditionally stable, requiring that the time increment be smaller than the shortest *medium-boundary period* divided by π. This shortest period may be obtained by determining the highest natural frequency for the eigenproblem

$$\omega^2\mathbf{M}_s\mathbf{x} = \mathbf{G}\,\mathbf{K}_m\mathbf{G}^t\mathbf{x} \tag{4.1}$$

In cases where the surrounding soil does not appreciably stiffen the embedded structure beyond its inherent level, the highest *medium-boundary frequency* is substantially lower than the highest natural frequency characterizing the structure itself, thereby allowing the analyst to carry out stable calculations with a relatively large time increment. The remainder of this section describes the staggered solution procedure and the stability analysis that leads to equation (4.1).

4.1 Solution Algorithm. To construct the staggered solution procedure for equation (2.14), those equations are expressed at mid-step as

$$\mathbf{M}_s\ddot{\mathbf{x}}_{n+1/2} + \mathbf{D}_T\dot{\mathbf{x}}_{n+1/2} + \mathbf{K}_s\mathbf{x}_{n+1/2} = \mathbf{f}_{n+1/2} - \mathbf{K}_M\mathbf{x}_{n+1/2} \tag{4.2}$$

where the time step $n = t/\Delta t$, in which t and Δt are time and *fixed* time increment, respectivley, and where the *total damping matrix* $\mathbf{D}_T$, the *medium-boundary stiffness matrix* $\mathbf{K}_M$, and the *total force vector* $\mathbf{f}$ are given by

$$\begin{aligned}\mathbf{D}_T &= \mathbf{D}_s + \mathbf{G}\,\mathbf{D}_m\mathbf{G}^t\\ \mathbf{K}_M &= \mathbf{G}\,\mathbf{K}_m\mathbf{G}^t\\ \mathbf{f} &= -\mathbf{G}\mathbf{g}_I + \mathbf{G}\,\mathbf{D}_m\dot{\mathbf{u}}_I + \mathbf{G}\,\mathbf{K}_m\mathbf{u}_I + \mathbf{f}_i\end{aligned} \tag{4.3}$$

The integration algorithm utilized is the *trapezoidal rule* (see, e.g., Henrici, 1962), for which

$$\begin{aligned}\dot{\mathbf{x}}_{n+1/2} &= (\mathbf{x}_{n+1/2} - \mathbf{x}_n)/\delta\\ \ddot{\mathbf{x}}_{n+1/2} &= (\dot{\mathbf{x}}_{n+1/2} - \dot{\mathbf{x}}_n)/\delta\\ \mathbf{x}_{n+1} &= 2\mathbf{x}_{n+1/2} - \mathbf{x}_n\\ \dot{\mathbf{x}}_{n+1} &= 2\dot{\mathbf{x}}_{n+1/2} - \dot{\mathbf{x}}_n\end{aligned} \tag{4.4}$$

where $\delta = \Delta t/2$. Introduction of the first and then the last of these into the third yields the standard form

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{\Delta t}{2}(\dot{\mathbf{x}}_{n+1} + \dot{\mathbf{x}}_n) \tag{4.5}$$

Now the first two of equations (4.4) are introduced into the left side of equation (4.2) and $\mathbf{x}_{n+1/2}$ on the right side of equation (4.2) is *predicted* as $\mathbf{x}_{n+1/2}^p$ to obtain the set of algebraic equations

$$\mathbf{E}_l\mathbf{x}_{n+1/2} = \mathbf{e}_{n+1/2} - \mathbf{E}_r\mathbf{x}_{n+1/2}^p \tag{4.6}$$

where

$$\begin{aligned}\mathbf{E}_l &= \mathbf{M}_s + \delta\mathbf{D}_T + \delta^2\mathbf{K}_s\\ \mathbf{E}_r &= \delta^2\mathbf{K}_M\\ \mathbf{e}_{n+1/2} &= \delta^2\mathbf{f}_{n+1/2} + \mathbf{M}_s(\mathbf{x}_n + \delta\dot{\mathbf{x}}_n) + \delta\mathbf{D}_T\mathbf{x}_n\end{aligned} \tag{4.7}$$

Finally, the prediction $\mathbf{x}_{n+1/2}^p$ is based on the *one-term extrapolation*

$$\mathbf{x}_{n+1/2}^p = \mathbf{x}_n \tag{4.8}$$

The preceding staggered solution procedure leads to the following computational sequence to determine system response at time step $n + 1$:

(a) $\mathbf{f}_{n+1/2} = (\mathbf{f}_n + \mathbf{f}_{n+1})/2$

(b) $\mathbf{e}_{n+1/2} = \delta^2\mathbf{f}_{n+1/2} + \mathbf{M}_s(\mathbf{x}_n + \delta\dot{\mathbf{x}}_n) + \delta\mathbf{D}_T\mathbf{x}_n$

(c) $\mathbf{x}_{n+1/2}^p = \mathbf{x}_n$

(d) $\mathbf{x}_{n+1/2} = \mathbf{E}_l^{-1}[\mathbf{e}_{n+1/2} - \mathbf{E}_r\mathbf{x}_{n+1/2}^p]$

(e) $\mathbf{x}_{n+1} = 2\mathbf{x}_{n+1/2} - \mathbf{x}_n$

(f) $\dot{\mathbf{x}}_{n+1/2}^0 = (x_{n+1/2} - \mathbf{x}_n)/\delta$

(g) $\ddot{\mathbf{x}}_{n+1/2} = \mathbf{M}_s^{-1}(\mathbf{f}_{n+1/2} - \mathbf{D}_T\dot{\mathbf{x}}_{n+1/2}^0 - \mathbf{K}_T\mathbf{x}_{n+1/2})$

(h) $\dot{\mathbf{x}}_{n+1/2} = \dot{\mathbf{x}}_n + \delta\ddot{\mathbf{x}}_{n+1/2}$

(i) $\dot{\mathbf{x}}_{n+1} = 2\dot{\mathbf{x}}_{n+1/2} - \dot{\mathbf{x}}_n$

where the *total stiffness matrix* $\mathbf{K}_T = \mathbf{K}_s + \mathbf{K}_M$. To improve accuracy, an iterative loop has been introduced at (d), wherein $\mathbf{x}_{n+1/2}^p$ on the right is corrected to the previously calculated value of $\mathbf{x}_{n+1/2}$; two iterations generally produce satisfactory convergence. The calculation starts at $n = 0$ with $\mathbf{x}_0 = \dot{\mathbf{x}}_0 = 0$.

4.2 Stability Analysis. Park (1980) has performed a stability analysis of a generalized form of the staggered solution procedure just described. The result is that the procedure is computationally stable if no root of the characteristic equation

$$\det[[z^2(\mathbf{M}_s - \delta^2\mathbf{K}_M) + z\delta\mathbf{D}_T + \delta^2\mathbf{K}_T]] = 0 \tag{4.9}$$

has a positive real part. Verification of this condition is relatively straightforward when all of the matrices in (4.9) are symmetric; it is generally quite difficult when one or more is not. Unfortunately, as discussed in Section 3, the medium stiffness matrix $\mathbf{K}_m$ is nonsymmetric, which pollutes $\mathbf{K}_M$ and

$\mathbf{K}_T$. Fortunately, however, $\mathbf{K}_m$ constitutes a small perturbation of $\hat{\mathbf{K}}_m$, which *is* symmetric; hence it is appropriate to consider the characteristic equation

$$\det[[z^2(\mathbf{M}_s - \delta^2\hat{\mathbf{K}}_M) + z\delta\mathbf{D}_T + \delta^2\hat{\mathbf{K}}_T]] = 0 \tag{4.10}$$

where $\hat{\mathbf{K}}_M = \mathbf{G}\,\hat{\mathbf{K}}_m\mathbf{G}^t$ and $\hat{\mathbf{K}}_T = \mathbf{K}_s + \hat{\mathbf{K}}_M$.

As discussed on page 255 of Bellman (1970), no root of equation (4.10) has a positive real part if $(\mathbf{M}_s - \delta^2\hat{\mathbf{K}}_M)$, $\mathbf{D}_T$ and $\hat{\mathbf{K}}_T$ are all non-negative definite and either $(\mathbf{M}_s - \delta^2\hat{\mathbf{K}}_M)$ or $\hat{\mathbf{K}}_T$ is positive definite. On physical grounds, $\mathbf{D}_T$ and $\hat{\mathbf{K}}_T$ are both non-negative definite, but generally not positive definite. However, inasmuch as $\mathbf{M}_s$ is positive definite, $(\mathbf{M}_s - \delta^2\hat{\mathbf{K}}_M)$ is positive definite if δ is sufficiently small. The degree of smallness defines the stability requirement, as discussed next.

Consider the following *first eigenproblem*:

$$\hat{\mathbf{Q}}\mathbf{x} = \lambda\mathbf{x} \tag{4.11}$$

where $\hat{\mathbf{Q}} = \mathbf{M}_s^{-1}\hat{\mathbf{K}}_M$. This problem yields nonnegative real eigenvalues and real eigenvectors. These eigenvectors may be assembled into a modal transformation matrix $\boldsymbol{\Psi}$ that diagonalizes $\hat{\mathbf{Q}}$ as $\boldsymbol{\Psi}^t\hat{\mathbf{Q}}\boldsymbol{\Psi} = \hat{\mathbf{Q}}^d$ and normalizes as $\boldsymbol{\Psi}^t\boldsymbol{\Psi} = \mathbf{I}$, the identity matrix. Hence the introduction into equation (4.11) of a transformation from physical to generalized coordinates as $\mathbf{x} = \boldsymbol{\Psi}\mathbf{y}$ and subsequent premultiplication through by $\boldsymbol{\Psi}^t$ yield the diagonal eigenvalue matrix

$$\boldsymbol{\Lambda}_Q = \hat{\mathbf{Q}}^d \tag{4.12}$$

Consider next the following *second eigenproblem*:

$$\hat{\mathbf{K}}_M\mathbf{x} = \lambda\mathbf{M}_s\mathbf{x} \tag{4.13}$$

whose eigenvalues and eigenvectors are the same as those of the first eigenproblem. Hence the transformation from physical to generalized coordinates and premultiplication through by $\boldsymbol{\Psi}^t$ yields

$$\boldsymbol{\Lambda}_{K/M} = (\mathbf{M}_s^d)^{-1}\hat{\mathbf{K}}_M^d \tag{4.14}$$

where $\mathbf{M}_s^d = \boldsymbol{\Psi}^t\mathbf{M}_s\boldsymbol{\Psi}$ and $\hat{\mathbf{K}}_M^d = \boldsymbol{\Psi}^t\hat{\mathbf{K}}_M\boldsymbol{\Psi}$; $\boldsymbol{\Lambda}_{K/M}$ is, of course, identical to $\boldsymbol{\Lambda}_Q$.

Finally, consider the following *third eigenproblem*:

$$(\mathbf{M}_s - \delta^2\hat{\mathbf{K}}_M)\mathbf{x} = \lambda\mathbf{x} \tag{4.15}$$

Transformation and premultiplication through as before yields

$$\begin{aligned}\boldsymbol{\Lambda}_{M-K} &= \mathbf{M}_s^d - \delta^2\hat{\mathbf{K}}_M^d \\ &= \mathbf{M}_s^d[\mathbf{I} - \delta^2(\mathbf{M}_s^d)^{-1}\hat{\mathbf{K}}_M^d] \\ &= \mathbf{M}_s^d[\mathbf{I} - \delta^2\boldsymbol{\Lambda}_{K/M}] \\ &= \mathbf{M}_s^d[\mathbf{I} - \delta^2\boldsymbol{\Lambda}_Q]\end{aligned} \tag{4.16}$$

Hence the eigenvalues of $(\mathbf{M}_s - \delta^2\hat{\mathbf{K}}_M)$ are all positive, and thus $(\mathbf{M}_s - \delta^2\hat{\mathbf{K}}_M)$ is positive definite, if δ^2 times the largest eigenvalue $\lambda_Q^{\max}$ is less than unity. With $\lambda_Q^{\max} = (\omega_Q^{\max})^2$, this yields the *stability requirement*

$$\Delta t < \frac{2}{\omega_Q^{\max}} \tag{4.17}$$

which is stated in slightly different terms at the beginning of this section.

Establishment of the stability requirement (4.17) for a symmetric medium stiffness matrix facilitates the estimation of a similar requirement for a nonsymmetric one. Clearly, no root of equation (4.9) has a positive real part if δ is vanishingly small, as $\mathbf{M}_s$ is symmetric and positive definite, and $\mathbf{D}_T$ is symmetric and nonnegative definite. Also, on physical grounds, the eigenvalues of $(\mathbf{M}_s)^{-1}\,\mathbf{K}_M$ must be real and nonnegative. Finally, the eigenvalues for the three eigenproblems above differ only slightly from their counterparts when $\hat{\mathbf{K}}_M$ is replaced by $\mathbf{K}_M$ because $\mathbf{K}_m$ constitutes a small perturbation of $\hat{\mathbf{K}}_m$. Hence, as δ is increased from zero, all the roots of equation (4.9) contain negative real parts until the stability requirement (4.17) is approached, where $\omega_Q^{\max}$ now pertains to the use of $\mathbf{K}_m$.

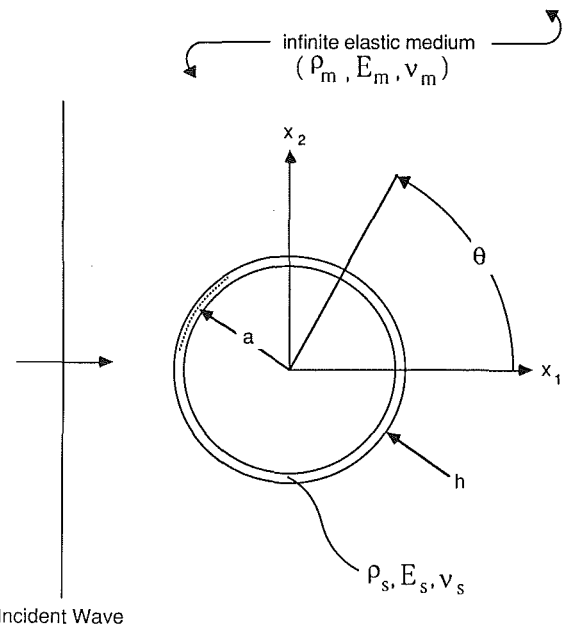


Fig. 1 Geometry and notation for canonical problems

5 Implementation and Computation

This section describes the techniques used to implement in software the approach delineated above, and presents numerical results generated by that software. Modern software-engineering techniques are used (Felippa, 1981), in order to facilitate extension to large-scale production analysis. The numerical results pertain to linear canonical problems involving plane, dilatational step-waves that envelope infinite-cylindrical and spherical shells (Fig. 1). These problems possess known analytical solutions.

5.1 Software Implementation. The approach described in Sections 2, 3, and 4 is embodied in an assembly of four software entities:

1. *Structural Matrix Generator.* The structural mass and stiffness matrices, $\mathbf{M}_s$ and $\mathbf{K}_s$ in equation (2.14), are generated by the finite-element code DIAL (Ferguson and Cyr, 1984); $\mathbf{D}_s$ is neglected. The structural matrices and related data are read into a NICE global database (Felippa, 1982).
2. *Medium Matrix Generator.* The medium damping and stiffness matrices, $\mathbf{D}_m$ and $\mathbf{K}_m$ in equation (2.14), are generated by software developed as part of this study in the manner described above; the force-transformation matrix $\mathbf{G}$ is constructed as a correspondence table. These data are read into the NICE global database.
3. *Incident Field Generator.* The incident-wave displacement, velocity and force vectors, $\mathbf{u}_I$, $\dot{\mathbf{u}}_I$, and $\mathbf{g}_I$ in equation (2.14) are also generated by software developed as part of this study in the manner described below; as these are time-dependent vectors, they are calculated dynamically as the calculation proceeds. $\mathbf{f}_i$ is taken as zero.
4. *Staggered Solution Procedure.* The solution algorithm described in Subsection 4.1 is implemented as a NICE procedure using a command language interpreter (Felippa, 1983). The matrix operations embedded in the algorithm are performed with a matrix utility processor for data in unblocked skyline format (Felippa, 1978).

The FE and BE models are constructed independently, although the element grids match at their common boundary.

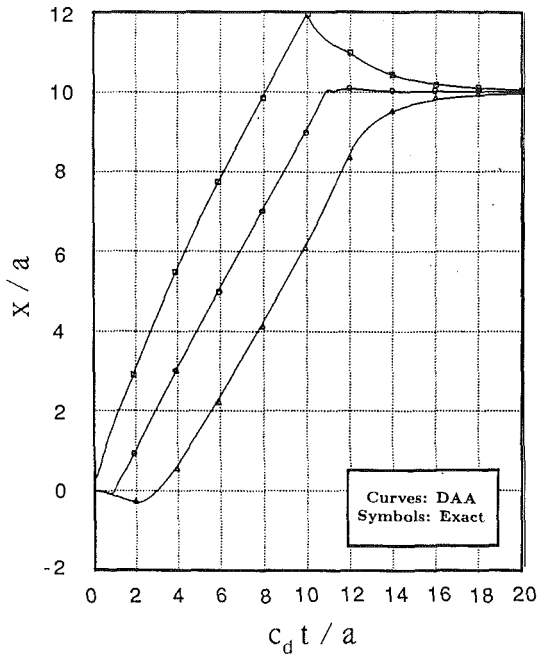


Fig. 2 Displacement response histories for the infinite cylinderical shell (DAA boundary on shell surface)

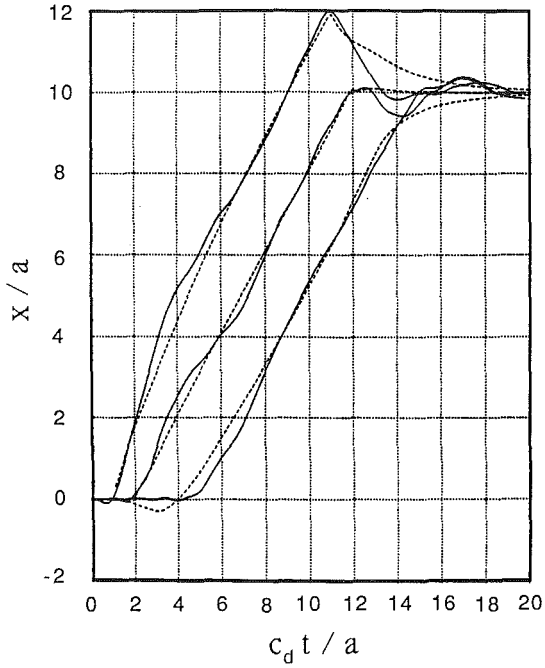


Fig. 4 Displacement response histories for the infinite cylinderical shell (solid curves: finite elements around shell; dashed curveds: DAA Boundary on shell surface)

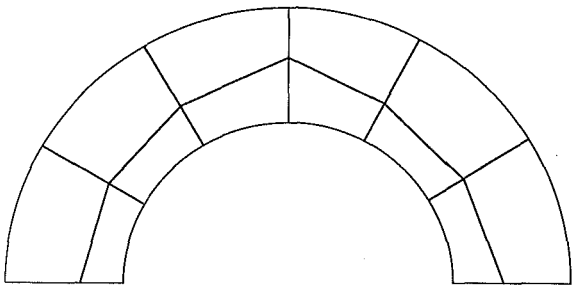


Fig. 3 Half-model grid for the infinite cylinderical shell problem (finite elements extending to $r = 2a$)

Geometrical symmetry is exploited in both canonical problems.

5.2 Incident-Wave Vectors. A plane, dilational step-wave characterized by a velocity jump V_o and propagating in the x_1 direction may be described in terms of a scalar potential as

$$\phi^I = -\frac{V_o}{2c_d}(c_d t - x_1 - a)^2 H(c_d t - x_1 - a) \tag{5.1}$$

where c_d is the dilatational speed in the elastic medium, H is the Heaviside operator, and $-a$ is the point on the x_1 axis where the wave front is located at $t = 0$. The application of classical continuum formulas (Achenbach, 1973) yields for the components of the geometrial displacement and velocity vectors for the incident wave

$$u_i^I = \delta_{1i}\frac{V_o}{c_d}(c_d t - x_1 - a)H(c_d t - x_1 - a)$$
$$\dot{u}_i^I = \delta_{1i} V_o H(c_d t - x_1 - a) \tag{5.2}$$

Hence the elements of the computational vectors $\mathbf{u}_I$ and $\dot{\mathbf{u}}_I$ are given by equations (5.2) evaluated at the surface nodes.

Similarly, the components of the incident-wave stress tensor and geometrical surface-traction vector are given by (Achenbach, 1973)

$$\sigma_{ij}^I = -\delta_{ij}\frac{V_o}{c_d}(\lambda + 2\mu\delta_{1i})H(c_d t - x_1 - a)$$
$$t_j^I = \sigma_{ij} n_i \tag{5.3}$$

where λ and μ are the Lamé constants and the n_i are the direction-cosines for the surface normal. Hence the computational vector $\mathbf{g}_I$ is determined from equation (3.16).

5.3 Infinite Cylindrical Shell. The first canonical problem is that of an infinite cylindrical shell embedded in an elastic medium and excited by a transverse, plane, dilatational wave (Garnet and Crouzet-Pascal, 1966). The parameter ratios for this problem are $E_s/E_m = 2.5$ (Young's modulus), $h/a = 0.01$ (shell thickness-to-radius), $\rho_s/\rho_m = 1.156$ (mass density), $\nu_m = 0.25$ and $\nu_s = 0.2$ (Poisson's); these pertain to a concrete shell in slow granite. The duration of the rectangular incident-wave pulse is $c_d t/a = 10$. A curved, three-noded shell element is used to model the shell, so that the FE/BE discretization employs *conforming elements*.

The first computational model for this problem places the DAA boundary directly on the shell in the manner of Underwood and Geers (1981). The use of six curved quadratic elements over the half-model yields results that are virtually identical to those of Underwood and Geers (1981), which were generated with twenty linear elements over the half-model. Figure 2 shows DAA and exact displacement-response histories; agreement is seen to be excellent.

Such agreement is not produced by singly asymptotic approximations. The elastic-foundation approximation generates response histories that oscillate markedly about the corresponding exact responses and the viscous-boundary approximation generates response histories that grossly exceed their exact counterparts (Geers and Yen, 1981).

The second computational model introduces eight-noded medium finite elements between the shell and the DAA boundary, which is located one shell radius out from the shell surface (Fig. 3). The displacement-response histories thus produced are shown in Fig. 4 as solid lines, along with their DAA counterparts from Fig. 3, which are shown as dashed lines. It

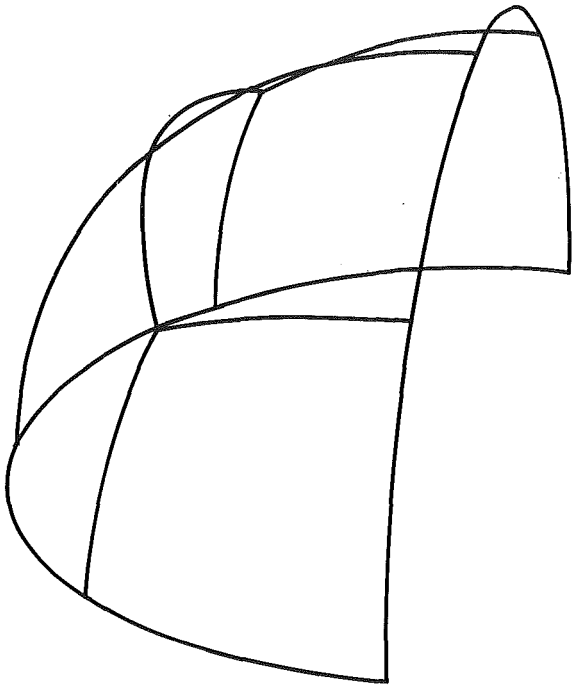


Fig. 5 Quarter-model grid for the spherical shell problem (DAA boundary on shell surface)

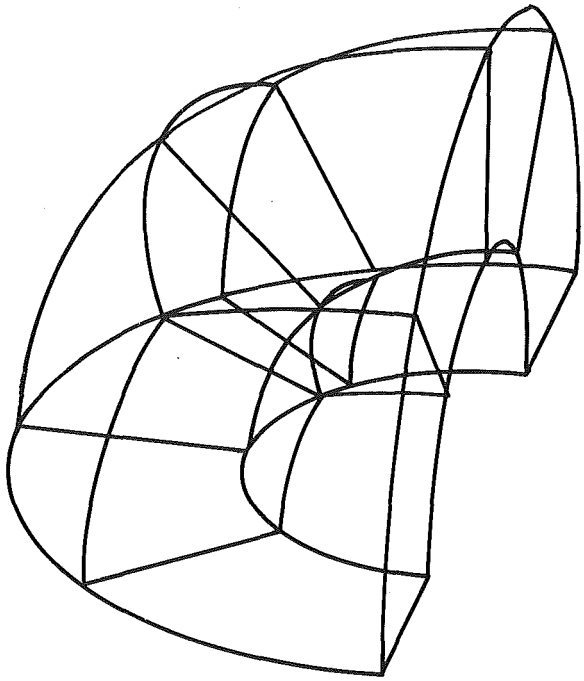


Fig. 7 Quarter-model grid for the spherical shell problem (Finite elements extending to $r = 2a$)

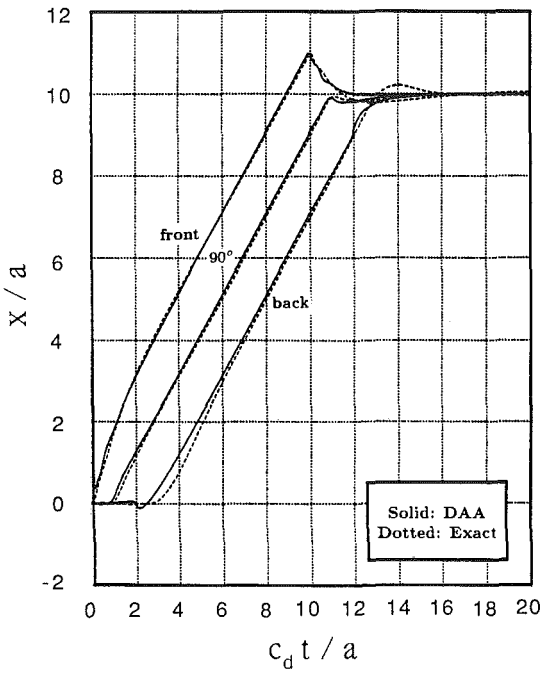


Fig. 6 Displacement response histories for spherical shell (DAA boundary on shell surface)

is seen that the use of medium finite elements degrades solution accuracy somewhat by introducing spurious oscillations caused by *ringing of the mesh*. A third computational model, which locates the DAA boundary three shell radii out from the shell surface, yields results that are even more oscillatory, although peak-response values are still satisfactory.

5.4 Spherical Shell. The second canonical problem is that of a spherical shell embedded in an elastic medium and excited by a plane dilatational wave (Grafton and Fox, 1965; Geers and Yen, 1981). The parameter ratios for this problem are the same as those for the infinite cylindrical shell, and the duration of the rectangular incident-wave pulse is also $c_d t/a = 10$. An eight-noded Ahmad shell element is used to model the shell, so that this FE/BE discretization also employs *conforming elements*.

As previously, the first computational model for this problem places the DAA boundary directly on the shell; six eight-noded quadratic elements are used over the quarter-model of the shell (Fig. 5). DAA-based displacement-response histories are compared with their exact counterparts in Fig. 6, the latter having been generated in the manner of Geers and Yen (1981). Here too, agreement is seen to be excellent; and, the singly asymptotic approximations also fail, generating response histories that exhibit behavior similar to that described above for the infinite cylindrical shell.

The second computational model introduces twenty-noded medium finite elements between the shell and the DAA boundary, which is located one shell radius out from the shell surface (Fig. 7). The displacement-response histories thus produced are shown in Fig. 8 as solid lines, along with their DAA counterparts from Fig. 6, which are shown as dashed lines. Here too, it is seen that the use of medium finite elements degrades solution accuracy by introducing spurious oscillations caused by *ringing of the mesh*.

6 Conclusion

This paper has documented the formulation and implementation of a nonreflecting boundary for use with existing finite-element codes to perform nonlinear ground-shock analyses of buried structures. The boundary is based on a first-order doubly asymptotic approxiation (DAA_1) for disturbances propagating outward from a selected portion of the soil medium surrounding the structure of interest. The resulting set of first-order ordinary differential equations is then combined with the second-order equations of motion for the finite-element model so as to facilitate solution by a staggered solution procedure. This procedure is shown to be computatioally stable as long as the time increment is smaller than a limiting value based on the finite-element mass matrix and the DAA-boundary stiffness matrix. Computational results produced by the boundary are compared with exact results for linear

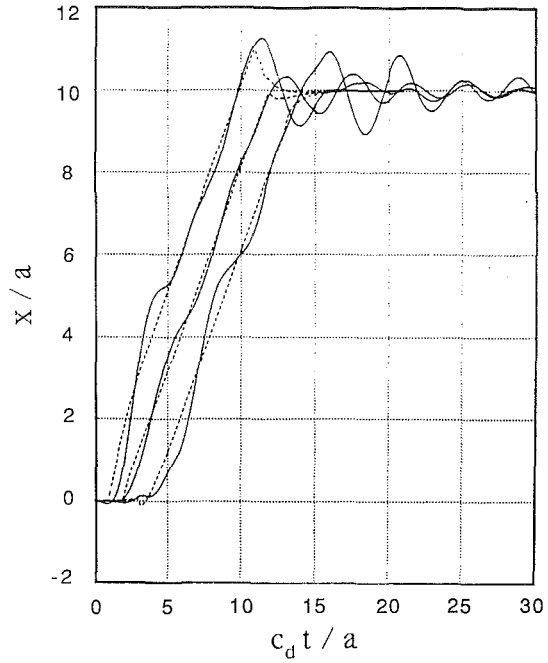


Fig. 8 Displacement response histories for spherical shell (solid curves: finite elements around shell; dashed curves: DAA boundary on shell surface)

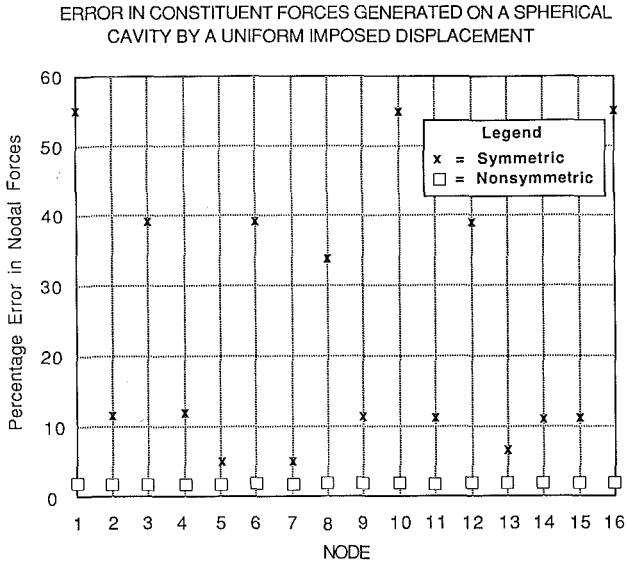


Fig. 9 Error in nodal-force values produced by symmetric and nonsymmetric medium stiffness matrices

canonical problems pertaining to infinite-cylindrical and spherical shells.

From this study, the following observations may be made:

1. Doubly asymptotic approximations are clearly superior to singly asymptotic approximations, the former incorporating both radiative energy dissipation and elastic restoring forces, the latter accounting for only one or the other.
2. While the medium damping matrix may be interpreted in terms of local dashpots positioned on the DAA surface, the medium stiffness matrix is not so easily regarded; attempts to simplify the fully coupled nature of $\mathbf{K}_m$ merely degrade the validity of low-frequency approximation.
3. Although it is tempting to use a symmetric medium stiffness matrix in DAA computations, the resulting loss of accuracy constitutes too high a price.
4. The computational stability requirement (4.17) is a generous one when the soil is substantially softer than the structural material; when this is not the case, however, more efficient computations might be realized with an unconditionally stable staggered solution procedure, which is yet to be developed.
5. The use of modern software-engineering techniques greatly facilitates the implementation of methods for the analysis of coupled systems.
6. The results for the linear canonical problems once again demonstrate the difficulty of propagating a discontinuous wave front through a finite-element grid and, in contrast, the good performance of a boundary-element grid located directly on the surface of the structure.

Acknowledgment

This work was sponsored by the Defense Nuclear Agency under Contract No. DNA 001-83-C-0239. The authors express their appreciation to Dr. Kent Goering as technical monitor, and to their colleagues Drs. Carlos Felippa, K. C. Park, and Chi-Lin Yen for their consultation and support.

References

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

Bellman, R., 1970, *Introduction to Matrix Analysis*, McGraw-Hill, New York.

Burton, A. J., 1976, "Numerical Solution of Acoustic Radiation Problems," National Physical Laboratory Report OC5/535.

Cohen, M., and Jennings, P. C., 1983, "Silent Boundary Methods for Transient Analysis," *Computational Methods for Transient Analysis*, Belytschko, T., and Hughes, T. J. R., ed., North-Holland, New York, pp. 301–360.

Cruse, T. A., 1969, "Numerical Solutions in Three-Dimensional Elastostatics," *International Journal of Solids and Structures*, Vol. 5, pp. 1259–1274.

Cundall, P. A., et al., 1978, "Solution of Infinite Dynamic Problems by Finite Modelling in the Time Domain," *Proceedings of the Second International Conference on Applied Numerical Modelling*, Madrid, Spain.

Datta, S. K., ed., 1982, *Earthquake Ground Motion and Its Effects on Structures*, AMD-Vol. 53, American Society of Mechanical Engineers, New York.

Felippa, C. A., and Park, K. C., 1980, "Staggered Transient Analysis Procedures for Coupled Mechanical Systems: Formulation," *Computer Methods in Applied Mechanics and Engineering*, Vol 24, pp. 61–111.

Felippa, C. A., 1981, "Architecture of a Distributed Analysis Network for Computational Mechanics," *Computers and Structures*, Vol. 13, pp. 405–413.

Felippa, C. A., 1982, "The Global Database Manager EZ-GAL," LMSC-766995, Lockheed Palo Alto Research Laboratory, Palo Alto, CA.

Felippa, C. A., 1983, "A Command Language for Applied Mechanics Processors," LMSC-D878511, Lockheed Palo Alto Research Laboratory, Palo Alto, CA.

Felippa, C. A., 1978, "SKYPUL Users Manual," LMSC-D623146, Lockheed Palo Alto Research Laboratory, Palo Alto, CA.

Ferguson, G. H., and Cyr, N., 1984, "DIAL User's Manual," LMSC-D850112, Rev. L2D8, Lockheed Missiles and Space Company, Sunnyvale, CA.

Garnet, H., and Crouzet-Pascal, J., 1966, "Transient Response of a Circular Cylinder of Arbitrary Thickness, in an Elastic Medium, to a Plane Dilatational Wave," ASME Journal of Applied Mechanics, Vol. 33, pp. 521–531.

Geers, T. L., 1978, "Doubly Asymptotic Approximations for Transient Motions of Submerged Structures," *Journal of the Acoustical Society of America*, Vol. 64, pp. 1500–1508.

Geers, T. L., and Yen, C.-L., 1981, "Transient Excitation of Cylindrical and Spherical Shells Embedded in Elastic Media: Residual Potential and Doubly Asymptotic Solutions," presented at the Tenth U.S. National Congress of Applied Mechanics, Austin, Texas, June 16–20, 1986. LMSC-D767735, Lockheed Palo Alto Research Laboratory, Palo Alto, CA.

Grafton, P. E., and Fox, J. W. L., 1965, "Interaction of Ground Motion in an Elastic Soil with Buried Structures," Paper No. 65-410, AIAA.

Henrici, P., 1962, *Discrete Variables Methods in Ordinary Differential Equations*, Wiley, New York.

Kalinowski, A. J., ed., 1981, *Computational Methods for Infinite Domain Media-Structure Interaction*, AMD-Vol. 46, American Society of Mechanical Engineers, New York, NY.

Kupradze, V. D., 1965, *Potential Methods in the Theory of Elasticity*, D. Davey, London.

Lachat, J. C., and Watson, J. O., 1976, "Effective Numerical Treatment of Boundary Integral Equations: A Formulation for Three-Dimensional

Elastostatics," *International Journal for Numerical Methods in Engineeirng*, Vol. 10, pp. 991–1005.

Mathews, I. C., and Geers, T. L., 1985, "A Doubly Asymptotic, Non-Reflecting Boundary for Ground-Shock Analysis," *Advanced Topics in Boundary Element Analysis*, Cruse, T., et al., ed., AMD-Vol. 72, American Society of Mechanical Engineers, New York, pp. 271–290.

Park, K. C., et al., 1977, "Stabilization of Staggered Solution Procedures for Fluid-Structure Interaction Analysis," *Computational Methods for Fluid-Structure Interaction Problems*, Belytschko, T., and Geers, T. L., ed., AMD-Vol. 26, American Society of Mechanical Engineers, New York, pp. 95–124.

Park, K. C., 1980, "Partitioned Transient Analysis Procedures for Coupled-Field Problems: Stability Analysis," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 370–376.

Rizzo, F. J., 1967, "An Integral Equation Approach to Boundary Value Problems of Classical Elastostatics," *Quarterly of Applied Mathematics*, Vol. 25, pp. 83–95.

Smith, W. D., 1974, "A Nonreflecting Plane Boundary for Wave Propagation Problems," *Journal of Computational Physics*, Vol. 15, pp. 492–503.

Takahasi, H., and Mori, M., 1973, "Quadrature Formulas Obtained by Variable Transformation," *Numerische Mathematik*, Vol. 21, pp. 206–219.

Timoshenko, S., and Goodier, J. N., 1951, *Theory of Elasticity*, McGraw-Hill, New York.

Underwood, P. G., and Geers, T. L., 1980, "Doubly Asymptotic, Boundary-Element Analysis of Nonlinear Soil-Structure Interaction," *Innovative Numerical Analysis for the Applied Engineering Sciences*, Shaw, R. P., et al., ed., University Press of Virginia, Charlottesville, VA, pp. 413–422.

Underwood, P., and Geers, T. L., 1981, "Doubly Asymptotic, Boundary-Element Analysis of Dynamic Soil-Structure Interaction," *Internatinal Journal of Solids and Structures*, pp. 687–697.

Zienkiewicz, O. C., 1977, *The Finite Element Method*, McGraw-Hill, London.

APPENDIX

Symmetric and Nonsymmetric Medium Stiffness Matrices

The accuracy of symmetric and nonsymmetric medium stiffness matrices is evaluated here by computing the nodal forces generated by a uniform radial displacement applied to a spherical cavity in an infinite elastic medium. The correct nodal forces follow from the known traction solution (Timoshenko and Goodier, 1951) and equation (3.16), the nodal forces produced by the nonsymmetric stiffness matrix follow from equation (3.17), and the nodal forces produced by the symmetric stiffness matrix follow from equation (3.17) with $\mathbf{K}_m$ replaced by $\hat{\mathbf{K}}_m$. Figure 9 shows, for the discretization of Fig. 5, computational error in nodal-force magnitudes computed with the symmetric and nonsymmetric matrices; $\mathbf{K}_m$ clearly outperforms $\hat{\mathbf{K}}_m$. It should be noted that convergence of the nodal forces generated by the symmetric medium matrix $\hat{\mathbf{K}}_m$ was obtained by successive mesh refinement.

On Tensile Shock Waves in Rubber-Like Materials[1]

M. Mihăilescu-Suliciu

I. Suliciu

Institute of Mathematics,
Academiei 14,
Bucharest, Romania

The problem of generation of one-dimensional tensile shock waves in rubber-like materials is studied numerically and compared to the exact elastic nonlinear solution and the steady wave solution. It is shown that a rate-type semilinear visco-elastic model can describe the steepening of the wave during its propagation and a "thickness" of the wave is naturally incorporated. An energetic criterion for the numerical stability is discussed. The numerical results point out the uncertainty (difficulty) one may encounter in measuring the dynamic Young's modulus and Maxwell-type viscosity coefficient.

1 Introduction

The main purpose of this work is to investigate the possibility of describing rubber-like materials by means of a rate-type viscoelastic constitutive equation. In this sense, we try on one hand to gain some insight in the way shock waves in nonlinear elasticity could be "captured" by using a "smoothing" Maxwell-type viscosity approach. On the other hand, we want to see how the steepening of the tensile waves in rate type viscoelasticity may take place when the equilibrium curve has an upwards oriented concavity. We also intend to test an energetic condition for the numerical stability of the integration scheme, obtained under different circumstances (Mihailescu-Suliciu and Suliciu, 1985).

We do not claim to give here a precise constitutive equation which could describe accurately the behavior of a certain material under dynamic test conditions. However, for the purpose of running the numerical experiments we did choose from the experimental literature (Treloar, 1949; Bell, 1973) the elastic (or equilibrium) stress-strain curve, the Young's modulus, etc., which are appropriate for a certain kind of rubber.

Governing Equations. The equations of one-dimensional motion of a body are

$$\rho \frac{\partial v}{\partial t} - \frac{\partial \sigma}{\partial X} = 0, \quad \frac{\partial \epsilon}{\partial t} - \frac{\partial v}{\partial X} = 0, \tag{1}$$

where $v = v(X,t)$ is the particle velocity, $\sigma = \sigma(X, t)$, $\epsilon = \epsilon(X, t)$ are the engineering stress and strain, respectively, and $\rho = \text{const.} > 0$ is the mass density in the reference configuration; t is the time coordinate and X is the space coordinate in the reference configuration.

We shall complete the system (1) either by a nonlinear elastic constitutive equation

$$\sigma = R(\epsilon) > 0,\ R'(\epsilon) > 0,\ R''(\epsilon) > 0,\ \epsilon \in [\epsilon_1, \epsilon_2], \tag{2}$$

or by a rate-type, viscoelastic constitutive equation

$$\dot{\sigma} - E\dot{\epsilon} = -k(\sigma - R(\epsilon)),\ \epsilon \in [\epsilon_1, \epsilon_2],\ -\infty < \sigma < \infty \tag{3}$$

where $\dot{\sigma} = \partial\sigma/\partial t$, $\dot{\epsilon} = \partial\epsilon/\partial t$, $k = \text{const.} > 0$ is called the Maxwell-type viscosity coefficient or simply viscosity coefficient; dimension $(k) =$ dimension $\{(\text{time})^{-1}\}$. $E = \text{const.} > 0$ is called the dynamic Young's modulus. The Newtonian viscosity coefficient μ is formally related to k by $\mu = E/k$ since if in (3) $\dot{\sigma}/k$ is neglected then $\sigma = R(\epsilon) + \mu\dot{\epsilon}$. But completing the system (1) with such a constitutive equation, one obtains a semilinear parabolic system of equations while the semilinear system (1) and (3) preserves the hyperbolic character of the system (1) and (2). Thus, by a Maxwell-type viscosity approach to a nonlinear elastic problem, the hyperbolic character of the governing system is preserved.

It is known that an initial or an initial and boundary value problem for system (1) and (2) may, in general, lead to solutions which involve shock waves after a finite interval of time even if the data are continuous (or even C^∞).

One can think of the constitutive equation (3) as a better model than the elastic one for the description of rubber behavior in dynamic experiments. Then we have to determine k (eventually as function of stress and strain) and E for each kind of rubber.

On the other hand, as it was suggested by Mihailescu-Suliciu and Suliciu (1985), if we are given an elastic nonlinear problem which involves shock waves we can artificially build the constitutive equation (3) taking a positive and large enough k and an $E > R'(\epsilon)$ on $[\epsilon_1, \epsilon_2]$. We add to the elastic initial conditions for strain and particle velocity, the initial condition for stress

$$\sigma(X,0) = \sigma_0(X) = R(\epsilon(X,0)).$$

In this way, instead of the elastic nonlinear problem we get a viscoelastic semilinear problem consisting of system (1) and (3) and the initial and boundary data.

In the viscoelastic problem, continuous initial and boundary data do not generate shock waves at any time. This means that

[1] This paper is dedicated to William F. Ames on the occasion of his 60th Birthday.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 20, 1985; final revision, April 15, 1986.

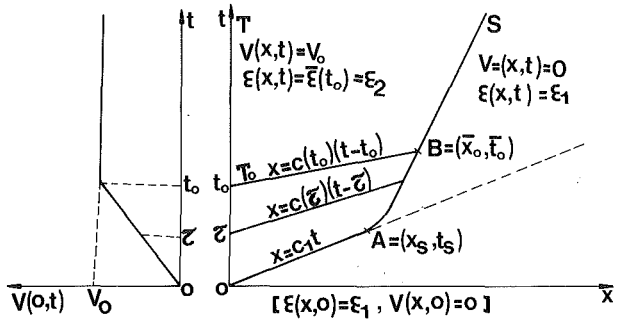


Fig. 1 The characteristics plane for an elastic nonlinear solution involving a shock wave

at a given spacial coordinate X_s where the elastic problem will lead, at a time t_s, to a jump of v and ϵ, the viscoelastic problem must lead to a continuous but fast increase in time of v, ϵ, and σ and the increase is faster for larger k (if E is fixed).

Another purpose of this paper is to get some information on the way shock waves can be "captured" and "shock structure" described by means of a Maxwell-type viscosity approach instead of a Newton-type viscosity approach.

On the Numerical Method and the Experimental Data. In order to obtain numerical solutions, the method of characteristics is often used for the initial and boundary value problem for the system (1) and (3). We only go up to the second approximation in this method since it is known (see for instance Rozhdestvensky and Yanenko, 1978) that, in general, the higher order approximations do not improve the order of accuracy over Mh^2, where h is the time integration step and $M>0$ is a constant.

For an isolated body problem in rate-type viscoelasticity (i.e., for a problem in which the body does not exchange energy with the surroundings) one can prove (Mihăilescu-Suliciu and Suliciu, 1985) using an energy estimate based on the second law of thermodynamics that one gets numerical stability if the time integration step h is smaller than a determined value h_m. For the initial and boundary value problem for which the strain $\epsilon(X,t)$ remains in the interval $[\epsilon_1, \epsilon_2]$ we have

$$h_m = \frac{E-R'(\epsilon_2)}{E-R'(\epsilon_1)} \cdot \frac{2}{k}. \tag{4}$$

Another purpose of this work is to test numerically if condition (4) is sufficient for the numerical stability in a non-isolated body problem. Such a problem is, for instance, that of the motion of a semi-infinite body with prescribed one-end particle velocity which will be considered here.

The elastic curve (2) which is the same as the equilibrium curve of equation (3) is taken of the form

$$\sigma = R(\epsilon) = A\exp(B\epsilon), \quad \epsilon\epsilon[\epsilon_1,\epsilon_2]. \tag{5}$$

When

$$A = 2.26\times 10^6 \text{ dyne/cm}^2, \quad B = 0.34, \quad \epsilon_1 = 3.5, \quad \epsilon_2 = 5.5 \tag{6}$$

it gives a good fit to an experimentally reported curve for a certain kind of rubber (Treloar 1949, Chapter 1, Section 1). For the mass density (cf Treloar, 1949, Chapter IX, Section 3, Fig. 69) we take

$$\rho = 1 \text{ gramme/cm}^3. \tag{7}$$

These experimental data seem also to agree with those discussed by Bell (1973, table 141, and Fig. 4.257 on p. 733).

For the same rubber for which (6) and (7) hold, the Young's modulus at zero stress and strain is experimentally found to be of order of magnitude 1×10^7 dyne/cm^2 (Treloar, Chapter I, Section 1). Since we have

$$R'(\epsilon) < E = 1\times 10^7 \text{ dyne/cm}^2 \tag{8}$$

we may take this value of E as the dynamic Young's modulus. The inequality (8) is essentially used in constructing the free energy function compatible with the second law of thermodynamics for the constitutive equation (3) (see Mihăilescu-Suliciu and Suliciu, 1985, and the literature quoted there). Based on that energy the estimate (4) is derived.

The Initial and Boundary Value Problems. We choose a continuous initial and boundary value problem similar to Kolsky's experiment, i.e.,

$$\epsilon(x,t)=\epsilon_1>0, \quad v(X,0)=0, \quad \sigma(X,0)=R(\epsilon_1), \quad X\geqq 0$$
$$v(0,t)=v_o(t)=\begin{cases} -at, & 0\leqq t\leqq t_o, \quad a>0 \\ -at_o=v_o, & t>t_o. \end{cases} \tag{9}$$

Due to the simple form of $v_o(t)$ we explicitly construct in Section 2 the exact solution of the problem (1) + (2) + (9) for $R(\epsilon)$ given by equation (5). In this solution a shock wave is developed in a finite time from continuous data.

In Section 3 we find, by the method of characteristics the numerical solution of the viscoelastic problem (1), (3), and (9) and compare it with the exact elastic solution for different viscosity coefficients.

2 A Nonlinear Elastic Solution Involving a Shock Wave

We consider the problem consisting of equations (1), (5), and (9) with $\rho=1$. This problem describes the motion of a semi-infinite elastic nonlinear body initially at rest, with the section $X=0$ subjected to a prescribed particle velocity (9). For the purpose of our numerical experiments we need the exact solution of the above problem. We construct this solution by elementary methods (see for instance Rozhdestvensky and Yanenko, 1978, Chapter I and Chapter IV) and we give it below.

The strain at $X=0$ is

$$\epsilon(0,\tau)=\bar{\epsilon}(\tau)=\begin{cases} \dfrac{2}{B}\ln[(aB\tau/2+C_1)/\sqrt{AB}], & \text{for } \tau\epsilon[0,t_o] \\ \bar{\epsilon}(t_o)=\epsilon_2>\epsilon_1, & \text{for } \tau>t_o \end{cases} \tag{10}$$

where $C_1=\sqrt{R'(\epsilon_1)}$. We obviously have $\bar{\epsilon}(\tau)\epsilon[\epsilon_1,\epsilon_2]$.

The curve $OABS$ (see Fig. 1) separates the rest region from the perturbed one. The straight line OA is a segment of the characteristic line $X=C_1t$. At the point

$$A=(X_s,t_s)=(2C_1^2/(aB),2C_1/(aB)) \tag{11}$$

the shock wave starts to develop. The parametric equation of the shock wave AB is $t=\tilde{t}(\tau)$, $X=\tilde{X}(\tau)$, $\tau\epsilon[0,t_o]$ with

$$\tilde{t}(0)=t_s, \tilde{t}(\tau)=\frac{\tau(aB\tau/3+C_1/2)}{C(\tau)-U(\tau)}, \quad \text{for } \tau\epsilon(0,t_o]$$
$$\tilde{X}(0)=X_s, \tilde{X}(\tau)=\frac{\tau C(\tau)(U(\tau)-aB\tau/6-C_1/2)}{C(\tau)-U(\tau)}, \tag{12}$$
$$\text{for } \tau\epsilon(0,t_o],$$

where

$$C(\tau)=\sqrt{R'(\bar{\epsilon}(\tau))}=aB\tau/2+C_1, \quad \tau\epsilon[0,t_o]$$
$$U(0)=C_1, \quad U(\tau)=[C^2(\tau)-C_1^2]/\left[2\ln\frac{C(\tau)}{C_1}\right], \quad \tau\epsilon(0,t_o]. \tag{13}$$

Here $C(\tau)$ and $U(\tau)$ are the acceleration and skock wave

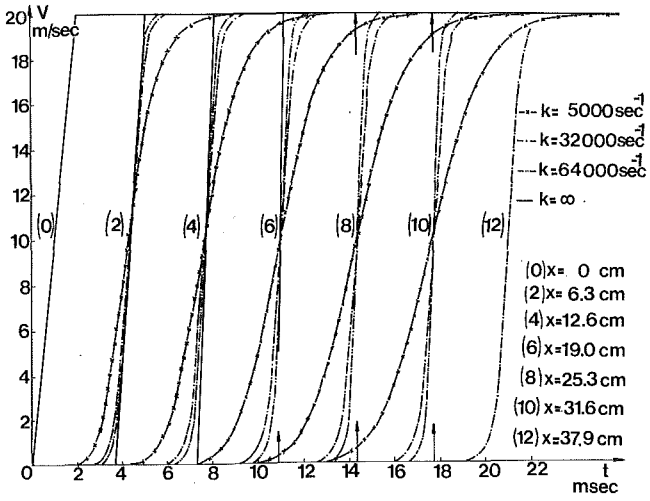


Fig. 2 Particle velocity-time profiles at different cross sections for several viscosity coefficients. The full line represents the exact elastic solution, and the cross sections, where there are jumps, are indicated by full arrows. Double dotted line (— ·· —) is for $k = 64000\ s^{-1}$, dotted line (— · —) is for $k = 32000\ s^{-1}$, and crossed line (— × —) is for $k = 5000\ s^{-1}$. The numbers (0), (2), . . . , (12) represent the arrival time in millisec. at the corresponding cross sections computed from the wave speed $c = \sqrt{E/\rho}$ by use of dynamic Young's modulus $E = 10^7$ dyne/cm^2 and $\rho = 1$ gr/cm^3.

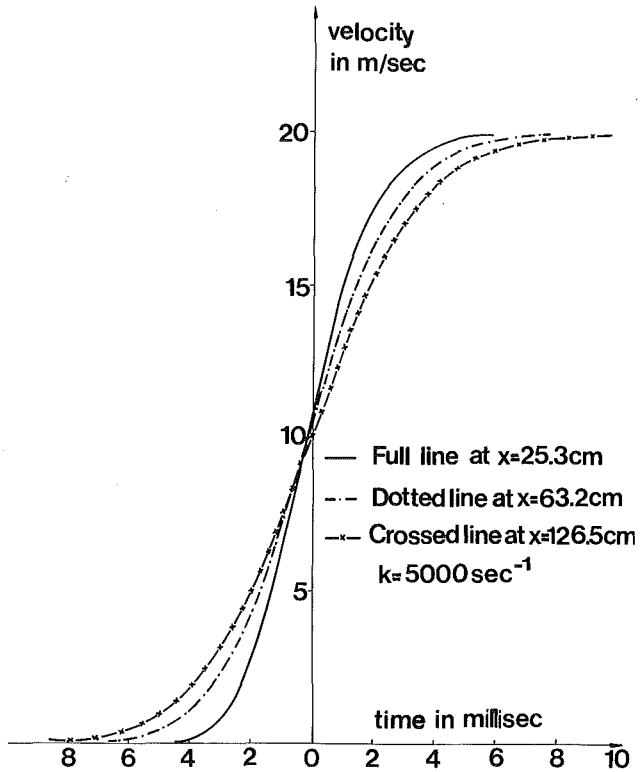


Fig. 3 Particle velocity-time profiles for the viscosity coefficient $k = 5000\ s^{-1}$ at different cross sections with the time origin on the elastic shock wave (or as they are seen by observers moving on the corresponding elastic shock waves)

speeds, respectively. Above the point $B = (\bar{X}_o, \bar{t}_o) = (\tilde{X}(t_o), \tilde{t}(\tilde{t}_o))$ the shock wave BS becomes a straight line

$$X = \bar{X}_o + U_o(t - \bar{t}_o), \quad t > \bar{t}_o;$$
$$U_o^2 = U^2(t_o) = [R(\epsilon_2) - R(\epsilon_1)]/(\epsilon_2 - \epsilon_1). \tag{14}$$

The solution in the perturbed region is as follows: In the region bounded by $OABT_oO$ (see Fig 1), $\epsilon(X, t) = \bar{\epsilon}(\tau)$, $v(X, t) = v_o(\tau)$ along $X = C(\tau)(t - \tau)$ for each fixed $\tau \epsilon [0, t_o]$. In the region bounded by TT_oBS, $\epsilon(X, t) = \epsilon_2$, $v(X, t) = v_o$.

3 Some Numerical Experiments in Rate-Type Viscoelasticity

We construct here by the method of characteristics, the numerical solutions of the viscoelastic problem (VEP) (1), (3), and (9) with the equilibrium curve $\sigma = R(\epsilon)$ given by equation (5). This VEP corresponds to the nonlinear elastic problem (NEP) solved in Section 2.

In the case when in the VEP the curve $\sigma = R(\epsilon)$ has downwards oriented concavity (i.e., $R''(\epsilon) < 0$),. the following facts are known from both laboratory and numerical experiments. The solutions $-v(X,t)$, $\epsilon(X,t)$, $\sigma(X,t)$ of the VEP are increasing functions of t for any fixed X and they tend asymptotically in time to a "plateau" value $-v_o$, ϵ_2, $R(\epsilon_2)$, respectively. Here the strain $\epsilon_2 > \epsilon_1$ is determined by v_o.

The increase in time for a fixed X is faster for larger k and for fixed k the increase in time is slower for larger X (see, for instance, Cristescu and Suliciu, 1982, Chapter III and IV, and also Daimaruya and Naitoh, 1983). However, in this case ($R''(\epsilon) < 0$), the corresponding NEP does not generate shock waves.

To understand some laboratory data as well as the behavior of the rate-type constitutive equation (3), we consider the case when the NEP associated with the equilibrium curve $\sigma = R(\epsilon)$ generates a shock wave starting at some section $X_s > 0$.

Kolsky's measurements show that the wave becomes steeper as it propagates. On the other hand, theoretical results show that the wave cannot become a mathematical shock wave if the material has a finite viscosity coefficient k and suggest that the wave becomes steeper at the same section X for a material with a larger viscosity coefficient. However, we would like to see the above facts in more detail and also to see the behavior of the wave with the propagating distance for different fixed viscosity coefficients k.

Numerical Experiments. The function $R(\epsilon)$, its numerical entries as well as the values of E and ρ are those of Section 1. The other numerical entries which will be used are

$$\epsilon_1 = 4, \ a = 1 \times 10^6 \text{ cm/s}^2, \ t_o = 2 \times 10^{-3} \text{ s}. \tag{15}$$

With these data, from equations (10) and (4) we get

$$\epsilon_2 = 5.055, \ h_m = 1.631/\text{k s}. \tag{16}$$

The strain ϵ_2 is asymptotically reached in time by the solution $\epsilon(X,t)$ of VEP, for fixed X.

We discusss here the results obtained for the following choices of k:

$$k = 5000, 32000, 64000\ \text{s}^{-1}. \tag{17}$$

Figure 2 shows the particle velocity in m/s versus time in milliseconds at sections $X = 0$; 6.32; 12.65; 18.97; 25.30; 31.62 cm obtained as the numerical solution of the VEP for the above values of k as well as the exact elastic solution for the above numerical data.

The exact elastic solution has the following properties: The solution stays continuous for $0 \leqq t < t_s = 10.18$ ms and all $X \geqq 0$. At a point (X_s, t_s), $X_s = 17.61$ cm, a shock wave starts to be generated and it is completely formed at the point $(\bar{X}_o, \bar{t}_o)$, $\bar{t}_o = 12.66$ ms. $\bar{X}_o = 22.06$ cm. For $t \geqq \bar{t}_o$ a single shock wave will propagate with a speed $U = 1897.73$ cm/s. Across this shock wave (ϵ, v) will jump from $(\epsilon_1 = 4, v = 0)$ to $(\epsilon_2 = 5.055, v_o = 2000$ cm/s). This solution is plotted in Fig. 2 by full and arrowed lines.

One observes that for all three values of k, at all sections $X > 0$, the velocity-time profile $v = v(t)$ starts tangent to the

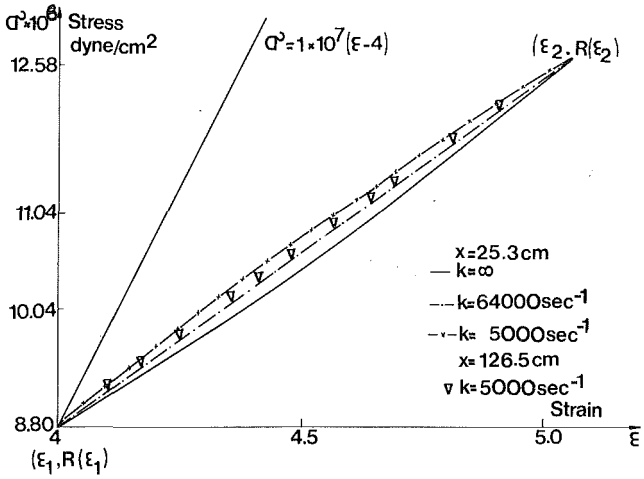


Fig. 4 Stress-strain relations at $X = 25.3$ cm and at $X = 126.5$ cm. The line $k = \infty$ represents the equilibrium curve $\sigma = R(\epsilon)$. The straight line with higher slope represents the instantaneous response curve at $\epsilon = 400$ percent (and $\sigma = 8.8 \times 10^6$ dyne/cm^2). The crossed straight line represents the steady wave stress-strain relation as well as the computed solutons for $k = 32000$/s and $k = 64000$/s at $X = 25.3$ cm.

line $v = 0$ and ends asymptotically tangent to the line $v = v_o = 2000$ cm/s. For $k = 5000$/s the wave spreads with the increase in distance from the end $X = 0$ (see Figs. 2 and 3). For $k = 32000$/s and $k = 64000$/s, Fig. 2 shows that the wave gets steeper with the propagating distance and the viscoelastic solution gets closer to the elastic one with the increase of k, i.e., the wave becomes steeper with the increase of both k and X.

Another important point is: suppose we deal with a real material described by the constitutive equation (3) with $k = 64000$/s, E and $R(\epsilon)$ as above. This material will have the bar velocity $c = \sqrt{E/\rho} = 3162.28$ cm/s. Suppose we measure the particle velocity-time profile of the wave at $X = 31.6$ cm (say); then the first signals arrives at this section at $t = 10$ ms but at the scale of our Fig. 2 we see something significant arriving only for $t > 16.6$ ms. This kind of behavior, which one must have in mind, places serious experimental difficulties for measuring both dynamic Young's modulus E and viscosity coefficient k. The same effect is present even if one uses the linear standard model of viscoelasticity.

Numerical Stability. According to the results of Mihăilescu-Suliciu and Suliciu (1985), a discrete isolated body problem formulated for the system (1) and (3) remains stable in energy if the maximum time integration step does not exceed the value h_m given by equation (4). This restriction on the time integration step may be physically interpreted in a similar way one does with the Courant number, i.e., for the numerical stability of the integration scheme it is necessary to index in time at a rate that allows the viscous effects time to develop. Since we do not know a similar proof for a nonisolated body problem such as (1), (3), and (9), we have tested here numerically the validity of the restriction $0 < h \leqq h_m$, with h_m given by equation $(16)_2$. The following experiments have been run. For each k in (17) we have chosen a time integration step $h_1 < h_m$, but close enough to h_m in order to have $h_2 = 2h_1 > h_m$.

For all k in (17), the numerical solution for $h = h_1$ was stable and agreed with the solution for $h = h_1/2$. For $k = 32000$/s or $k = 64000$/s the numerical solution did not behave properly for $h = 2h_1$ in the sense that the stress starts to decrease while the strain still increases for a few time steps, after which both of them decrease down to very large negative numbers. For $k = 5000$/s the same behavior is observed for $h = 4h_1$. For $h = 2h_1$ the numerical solution looks stable and it agrees with the solution for $h = h_1$ at the level of stresses which are not far from the equilibrium. However, the two numerical solutions (for $h = h_1$ and $h = 2h_1$) do not quite agree for stresses which are not close to the equilibrium.

Steady Waves. Our numerical experiments show how the particle velocity-time profile transforms in a steady wave profile for large k after propagating a certain distance (compare the plottings of Fig. 2 for $k = 32000$ and 64000/s at the sections $X = 25.3$, 31.6 cm and see also Fig. 4 for $k = 64000$/s and $X = 25.3$ cm).

Greenberg (1986) studied steady wave propagation through materials with governing equations of the form (1) and (3) but under the assumptions $E = E(\epsilon, \sigma)$, $\partial E/\partial\sigma < 0$, $\partial E/\partial\epsilon > 0$. These assumptions are not satisfied here since $E = \text{const}$. However, in a similar manner we find the following steady wave solutions

$$v(\zeta) = U_o[\epsilon(\zeta) - \epsilon_1],\ \sigma(\zeta) = R(\epsilon_1) + E_o[\epsilon(\zeta) - \epsilon_1],$$
$$\zeta = t - X/U_o,\ E_o = \rho U_o^2 = [R(\epsilon_2) - R(\epsilon_1)]/(\epsilon_2 - \epsilon_1),$$
$$(E - E_o)\epsilon'(\zeta) = k[E_o(\epsilon(\zeta) - \epsilon_1) - (R(\epsilon(\zeta)) - R(\epsilon_1))], \tag{18}$$

for our equations (1) and (3). (For a detailed discussion on steady waves, see Nunziato et al., 1974). The differential equation for ϵ is of the form $\epsilon'(\zeta) = kf(\epsilon(\zeta))/(E - E_o)$, $\zeta\epsilon(-\infty, \infty)$ or $\epsilon'(\eta) = f(\epsilon(\eta))$, $\eta\epsilon(-\infty, \infty)$ with $\eta = k\zeta/(E - E_o)$. Any solution $\epsilon(\eta)$ (or $\epsilon(\zeta)$) is a strictly increasing function on $(-\infty, \infty)$ with $\epsilon(\eta) \to \epsilon_1$ when $\eta \to -\infty$ and $\epsilon(\eta) \to \epsilon_2$ when $\eta \to \infty$.

Now let Δ, $0 < \Delta < (\epsilon_2 - \epsilon_1)/2$ be an arbitrarily small number ($\Delta = (\epsilon_2 - \epsilon_1)/10$ say). One can determine the maximal finite interval $(\eta'_\Delta, \eta''_\Delta)$ such that $\epsilon(\eta)\epsilon(\epsilon_1 + \Delta, \epsilon_2 - \Delta)$ when $\eta\epsilon(\eta'_\Delta, \eta''_\Delta)$. The quantity $\ell_\Delta = \eta''_\Delta - \eta'_\Delta$ can be called the conventional thickness of the wave in the plane $\eta - \epsilon$; it is independent on k and E and it depends only of $R(\epsilon)$ on $[\epsilon_1, \epsilon_2]$. (For a similar notion in gas dynamics see Witham, 1974, Sections 2.4, 6.15; for an exact and finite thickness of wave see Suliciu, 1974, and also Cristescu and Suliciu, 1982, Chapter V, Section 2). The conventional thickness of the wave L_Δ in the physical plane $\zeta - \epsilon$ is then

$$L_\Delta = (E - E_o)\ell_\Delta/k. \tag{19}$$

Since the conventional thickness of the steady wave for $k = 64000$/s does not exceed 0.8 ms (see Fig. 2) if we take $k = 1 \times 10^5$/s, then by equation (19) the thickness of the wave will be about 0.05 ms which at the scale of Fig. 2 makes this wave appear as a shock wave. We also note that for the equilibrium curve (5) and (6), for any E verifying equation (8) and any $k > 0$, the steady waves (i.e., the solutions of equation (18)) always exist. However, for small k the solution of our problem (1), (3), and (9), first spreads with the distance (see Fig. 3) and then it becomes a steady wave. One can see that this wave is almost steady from Fig. 4, where for $k = 5000$/s at the distance of 126.5 cm the triangles representing the stress-strain relation along the wave are very slightly above the straight line $\sigma = R(\epsilon_1) + E_o(\epsilon - \epsilon_1)$ representing the steady wave stress-strain relation.

4 Concluding Remarks

1. The shock waves, in nonlinear elasticity, generated by continuous data can be described by using a viscoelastic model with a large enough Maxwell-type viscosity coefficient. However, one has to decrease accordingly the time integration step, at least in the neighborhood of the shock waves.

2. In the semilinear rate-type viscoelasticity, shock waves are not generated by continuous data, but a wave can become steeper during its propagation. The wave does not transform itself into a mathematical shock wave if the viscosity coefficient is large but finite, instead it becomes a steep but smooth

steady wave. Therefore, a thickness of the wave is obtained, which is different in nature from the corresponding notion in gas dynamics.

3. There is a close relation between the size of the time integration step and magnitude of the viscosity coefficient. If this relation is not violated, one gets numerical stability; on the other hand, in a given problem it allows a proper choice of the time integration step in order to minimize the computational time.

4. The numerical results presented here show that one can not easily identify, in a laboratory experiment, the dynamic Young's modulus from the wave speeds since a high viscosity coefficient flattens very much the front of the wave during its propagation. Therefore, the dynamic Young's modulus can be much larger than it may seem. From this point of view, the results are better if one uses measurements closer to the impacted end. This remark is true even for the standard linear model of viscoelasticity.

Acknowledgment

The authors wish to express their gratitude to an unknown reviewer for his useful suggestions on the previous version of this work.

References

Bell, J. F., 1973, "The Experimental Foundation of Solid Mechanics," *Encyclopedia of Physics*, Vol. VI a/1, Springer, p. 733.

Cristescu, N., Suliciu, I., 1982, *Viscoplasticity*, Martinus Nijhoff, The Hague/The Netherlands, Ch. III, IV, and V, Sec. 2.

Daimaruya, M., Naitoh, M., 1983, "On the Existence of Strain Plateau in the Strain-Rate Dependent Theory of Malvern for Plastic Wave Propagation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 678–679.

Greenberg, J. M., 1968, "Existence of Steady Waves for a Class of Non-Linear Dissipative Materials," *Quarterly Journal of Applied Mathematics*, Vol. 26, pp. 27–34.

Kolsky, H., 1969, "Production of Tensile Shock Waves in Stretched Natural Rubber," *Nature*, Vol. 224, p. 1301.

Mihailescu-Suliciu, M., Suliciu, I., 1985, "On the Method of Characteristics in Rate-Type Viscoelasticity," *Z. Angew, Math. Mech.*, Vol. 65, pp. 479–486.

Nunziato, J. W., Walsh, E. K., Schuler, K. W., Barker, L. M., 1974, *Wave Propagation in Nonlienar Viscoelastic Materials*, Encyclopedia of Physics, Vol. VI a/4, Springer, Ch. V.

Rozhdestvensky, B. L., Yanenko, N. N., 1978, *Quasilinear Systems of Equations*, in Russian, Nauka, Moscow, Ch. I, Ch. IV, p. 434.

Suliciu, I., 1974, "Steady Waves in Materials with Rate-Type Behavior," *Mechanics Research Communications*, Vol. 1, pp. 167–172.

Treloar, L., 1949, *The Physics of Rubber Elasticity*, Oxford, Ch. I, Sec. 1, Ch. 9, Sec. 3.

Witham, G. B., 1974, *Linear and Nonlinear Waves*, Wiley, New York, Sections 2.4, 6.15.

Elastic Wave Scattering from an Interface Crack: Antiplane Strain

A. Boström
Division of Mechanics,
Chalmers University of Technology,
S-412 96 Göteborg, Sweden

The two-dimensional scalar problem of scattering of elastic waves under antiplane strain from an interface crack between two elastic half-spaces is considered. The method used is a direct integral equation method with the crack-opening displacement as the unknown. Chebyshev polynomials are used as expansion functions and the matrix in the resulting equations is simplified by contour integration techniques. The scattered far field is expressed explicitly in simple functions and the expansion coefficients. The consequences of energy conservation are explored and are used as a check in the numerical implementation. For incoming plane waves numerical results are given for the total scattered energy and the far field amplitude.

1 Introduction

The detection of cracks is an important problem in the nondestructive evaluation of materials. Cracks have thus been extensively studied in the literature, though mainly in highly idealized cases. The penny-shaped crack is exhaustively treated by Martin and Wickman (1983), who also give many further references, and by Krenk and Schmidt (1982). High-frequency aspects are considered by Keogh (1985 a, b) and Achenbach et al. (1978). Nonplanar cracks and partly debonded inclusions are studied by Boström and Olsson (1986) and Olsson (1986).

In the present paper we consider an interface crack between two homogeneous elastic half-spaces for the case of antiplane strain. The corresponding problem with an interface crack in a layered half-space is solved by Neerhoff (1970) and Yang and Bogy (1985), and Kundu (1986) solves for the transient response of an interface crack in a layered plate. Our method of solution is an integral equation method with the crack opening displacement as the unknown. To derive the integral equation we follow a procedure similar to that of Krenk and Schmidt (1982), and we thus avoid the introduction of any Green's function. One could instead proceed as Neerhoff (1979) and derive the integral equation from an integral representation containing a suitable Green's function. The two procedures should be equivalent but the one avoiding the Green's function is probably easier to generalize to more complicated cases.

2 Formulation of the Problem

Consider a scattering geometry as depicted in Fig. 1 with two homogeneous half-spaces $y > 0$ and $y < 0$ in welded contact except for an interface crack for $|x| < a$. Time-harmonic conditions are assumed throughout and the factor $\exp(-i\omega t)$ is suppressed. In the upper half-space the shear modulus is μ_1, the shear wavespeed is c_1, and the wavenumber is k_1 and the corresponding quantities in the lower half-space are μ_2, c_2, and k_2. For antiplane strain the only displacement component is in the z direction and this component satisfies the Helmholtz equation so that the displacments u^1 for $y > 0$ and u^2 for $y < 0$ satisfy

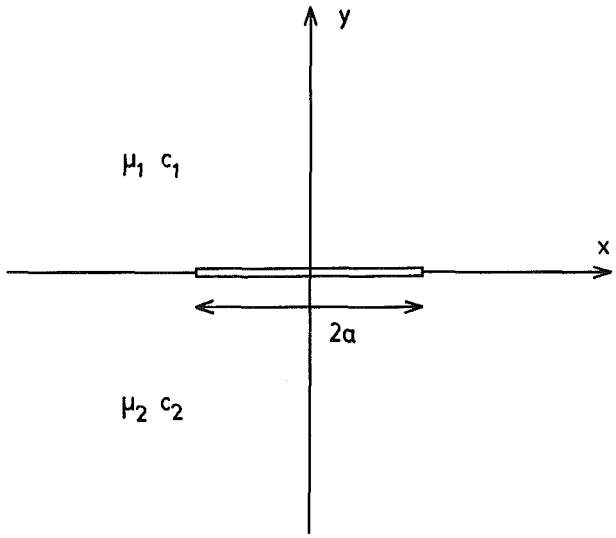


Fig. 1 The geometry

$$\nabla^2 u^1 + k_1^2 u^1 = 0 \quad y > 0 \tag{1}$$

$$\nabla^2 u^2 + k_2^2 u^2 = 0 \quad y < 0 \tag{2}$$

Along the welded part of the interface the displacement and traction are continuous:

$$u^1 = u^2 \qquad y = 0, \quad |x| > a \tag{3}$$

$$\frac{\partial u^1}{\partial y} = \beta \frac{\partial u^2}{\partial y} \qquad y = 0, \quad |x| > a \tag{4}$$

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 2, 1986; final revision, February 1, 1987.

where $\beta = \mu_2/\mu_1$ is the shear modulus quotient. Along the cracked part of the interface the traction vanishes and thus

$$\frac{\partial u^1}{\partial y} = \frac{\partial u^2}{\partial y} = 0 \quad y = 0, \quad |x| < a \tag{5}$$

To completely specify the problem we must also give the incoming field and the radiation conditions at infinity.

For simplicity we choose the incoming field as a plane wave of unit amplitude incident from below; more complicated incoming fields can then be obtained by superposition. To take advantage of the known reflection and transmission properties of the plane wave at the interface we make the following ansatz:

$$u^1 = Te^{i(q_0 x + h_{01} y)} + U^1 \qquad y > 0 \tag{6}$$

$$u^2 = e^{i(q_0 x + h_{02} y)} + Re^{i(q_o x - h_{02} y)} + U^2 \qquad y < 0 \tag{7}$$

where

$$h_{01} = (k_1^2 - q_0^2)^{1/2}, \quad Imh_{01} \geq 0$$

$$h_{02} = (k_2^2 - q_0^2)^{1/2}, \quad Imh_{02} \geq 0$$

and the reflection and transmission coefficients are

$$R = \frac{\beta h_{02} - h_{01}}{h_{01} + \beta h_{02}} \tag{8}$$

$$T = \frac{2\beta h_{02}}{h_{01} + \beta h_{02}} \tag{9}$$

The fields U^1 and U^2 are the extra fields due to the presence of the interface crack. These crack-scattered fields still satisfy the Helmholtz equations (1) and (2) and the welded boundary conditions (3) and (4), but instead of equation (5) they fulfill an inhomogeneous boundary condition on the crack surfaces.

3 The Integral Equation

The crack-scattered fields are now written in the form of Fourier representations:

$$U^1 = \int_{-\infty}^{\infty} f_1(q) e^{i(qx + h_1 y)} dq \tag{10}$$

$$U^2 = \int_{-\infty}^{\infty} f_2(q) e^{i(qx - h_2 y)} dq \tag{11}$$

where

$$h_1 = (k_1^2 - q^2)^{1/2}, \quad Imh_1 \geq 0$$

$$h_2 = (k_2^2 - q^2)^{1/2}, \quad Imh_2 \geq 0$$

In writing the representations (10) and (11) we have taken the radiation conditions into account. One of the boundary conditions says that

$$\frac{\partial U^1}{\partial y} = \beta \frac{\partial U^2}{\partial y} \qquad y = 0, \quad \text{all } x$$

and we especially note that this condition holds for *all* x. This gives

$$f_1(q) = -\frac{\beta h_2}{h_1} f_2(q) \tag{12}$$

The other boundary condition along the welded part of the interface is that

$$U^1 = U^2 \quad y = 0, \quad |x| > a$$

and this gives

$$-\int_{-\infty}^{\infty} \frac{\beta h_2 + h_1}{h_1} f_2(q) e^{iqx} dq = \begin{cases} 0 & |x| > a \\ \Delta U(x) & |x| < a \end{cases} \tag{13}$$

where ΔU is the so far unknown crack-opening displacement.

Inverting the Fourier transform in equation (13) we can express the crack-scattered field amplitude f_2 in the crack-opening displacement ΔU:

$$f_2(q) = \frac{-h_1}{2\pi(h_1 + \beta h_2)} \int_{-a}^{a} \Delta U(x) e^{-iqx} dx \tag{14}$$

The last boundary condition, equation (5), gives with equations (6), (12), and (14)

$$h_{01} Te^{iq_o x} + \int_{-\infty}^{\infty} dq \, \frac{\beta h_1 h_2}{2\pi(h_1 + \beta h_2)} e^{iqx}$$

$$\int_{-a}^{a} \Delta U(x') e^{-iqx'} dx' = 0 \tag{15}$$

This is the sought integral equation for the crack-opening displacement ΔU. Once ΔU is determined, equations (14) and (12) give the field amplitudes and equations (10) and (11) give the crack-scattered fields.

To solve the integral equation we expand ΔU in a complete set of Chebyshev polynomials:

$$\phi_n(x) = \begin{cases} \cos(n \arcsin(x/a)) & n = 1,3,5, \ldots \\ i \sin(n \arcsin(x/a)) & n = 2,4,6, \ldots \end{cases} \tag{16}$$

This set is convenient because it satisfies the correct edge condition (cf Neerhoff, 1979), and have the following property:

$$\int_{-a}^{a} \phi_n(x) e^{-i\gamma x} dx = \frac{n\pi}{\gamma} J_n(\gamma a) \tag{17}$$

We thus expand

$$\Delta U(x) = \sum_n \alpha_n \phi_n(x) \tag{18}$$

Inserting this in equation (15) we have

$$h_{01} Te^{iq_0 x} + \int_{-\infty}^{\infty} dq \, \frac{\beta h_1 h_2}{2q(h_1 + \beta h_2)} e^{iqx} \sum_n n\alpha_n J_n(qa) = 0$$

Multiplying by $\phi_n(x)$ and integrating over $[-a, a]$ we get

$$\sum_{n'} Q_{nn'} \alpha_{n'} = -2nTh_{01}/q_0 J_n(q_0 a) \tag{19}$$

where the symmetric matrix $Q_{nn'}$ is

$$Q_{nn'} = nn' \int_{-\infty}^{\infty} \frac{\beta h_1 h_2}{q^2(h_1 + \beta h_2)} J_n(qa) J_{n'}(qa) dq \tag{20}$$

The integration range in equation (20) can be reduced to a finite range. For a very similar integrand this has been discussed by van den Berg (1981) so we do not repeat the details here. The result is that $Q_{nn'}$ can be written as:

$$Q_{nn'} = 2nn' \int_0^{k_2} \frac{\beta\sqrt{k_1^2 - q^2}\sqrt{k_2^2 - q^2}}{\sqrt{k_1^2 - q^2} + \beta\sqrt{k_2^2 - q^2}} \left[J_>(qa) H_<^{(1)}(qa) + \frac{i}{n\pi} \delta_{nn'} \right] \frac{dq}{q^2}$$

$$+ 2nn' \int_{k_2}^{k_1} \frac{\beta^2(k_2^2 - q^2)\sqrt{k_1^2 - q^2}}{(1-\beta^2)q^2 + \beta^2 k_2^2 - k_1^2} \left[J_>(qa) H_<^{(1)}(qa) + \frac{i}{n\pi} \delta_{nn'} \right] \frac{dq}{q^2}$$

$$+ \frac{in\beta}{1+\beta} \delta_{nn'} \tag{21}$$

Here $>$ ($<$) denotes the larger (smaller) of n and n'. We have

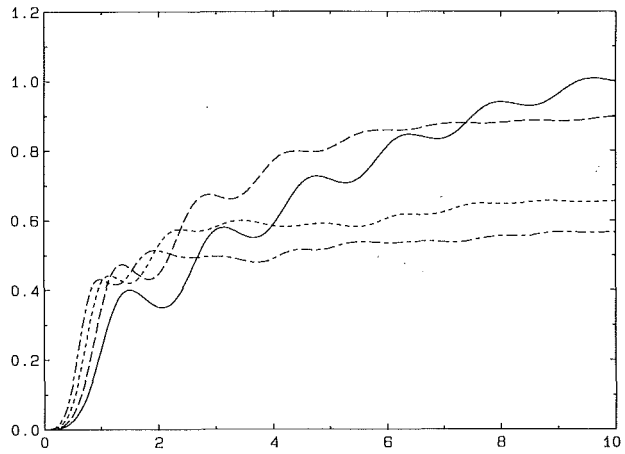


Fig. 2 The total crack-scattered energy as a function of frequency k_2a for an incoming plane wave making the angle $\phi_0 = 30$ deg with the interface and four material combinations: $c_2/c_1 = 1$, $\mu_2/\mu_1 = 1$ (————); $c_2/c_1 = 1.2$, $\mu_2/\mu_1 = 2$ (---); $c_2/c_1 = 1.5$, $\mu_2/\mu_1 = 5$ (·····); $c_2/c_1 = 1.8$, $\mu_2/\mu_1 = 8$ (-·-·-)

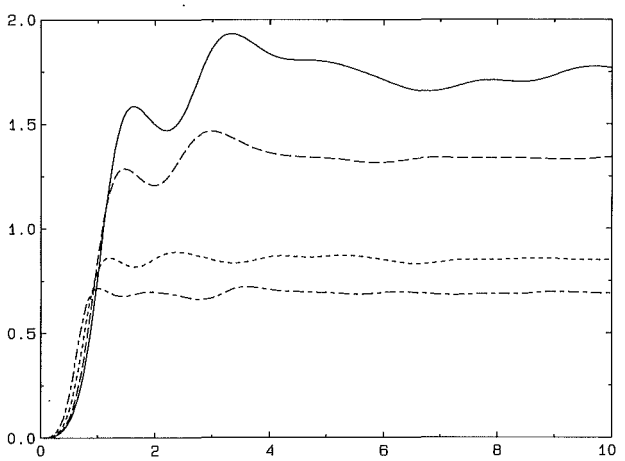


Fig. 3 Same as Fig. 2 but direction of incidence $\phi_0 = 60$ deg

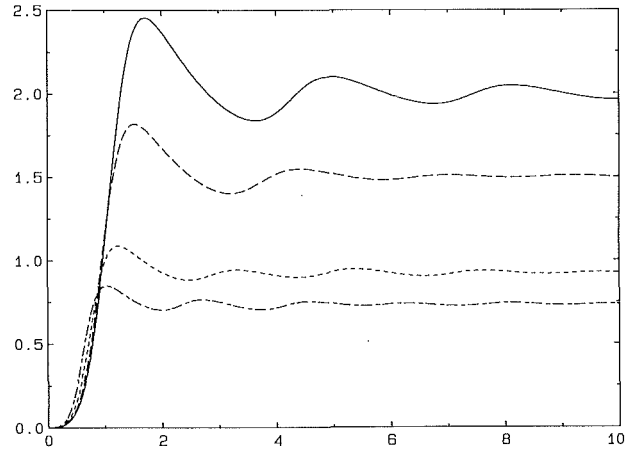


Fig. 4 Same as Fig. 2 but direction of incidence $\phi_0 = 90$ deg

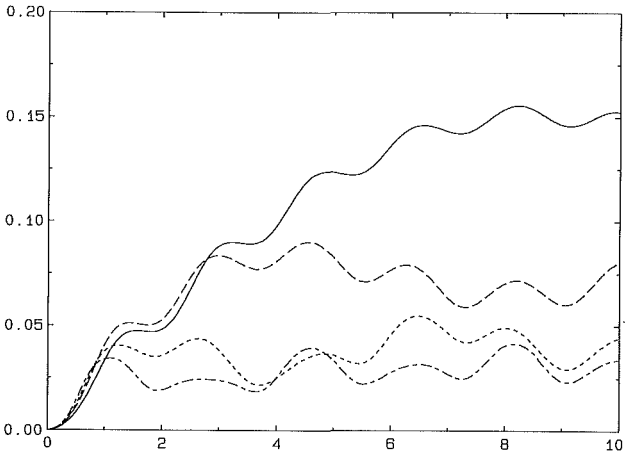


Fig. 5 The absolute value of the back-scattered far field amplitude as a function of frequency k_2a for an incoming plane wave making the angle $\phi_0 = 30$ deg with the interface and four material combinations: $c_2/c_1 = 1$, $\mu_2/\mu_1 = 1$ (————); $c_2/c_1 = 1.2$, $\mu_2/\mu_1 = 2$ (---); $c_2/c_1 = 1.5$, $\mu_2/\mu_1 = 5$ (·····); $c_2/c_1 = 1.8$, $\mu_2/\mu_1 = 8$ (-·-·-)

assumed that $k_1 < k_2$, but a very similar form holds when $k_2 > k_1$. For $k_1 = k_2$ and $\beta = 1$ our problem reduces to that considered by van den Berg (1981) and our $Q_{nn'}$ also reduces to his corresponding matrix.

Once equation (19) is solved for the unknown expansion coefficients α_n, the crack-opening displacement is determined and it in turn gives the field amplitude from equation (14)

$$f_2(q) = \frac{-h_1}{2q(h_1+\beta h_2)} \sum_n n\,\alpha_n J_n(qa) \tag{22}$$

and the crack-scattered fields from equations (10)–(12)

$$U^1 = \frac{\beta}{2} \sum_n n\,\alpha_n \int_{-\infty}^{\infty} \frac{h_2}{h_1+\beta h_2} J_n(qa) e^{i(qx+h_1 y)} \frac{dq}{q} \tag{23}$$

$$U^2 = -\frac{1}{2} \sum_n n\,\alpha_n \int_{-\infty}^{\infty} \frac{h_1}{h_1+\beta h_2} J_n(qa) e^{i(qx-h_2 y)} \frac{dq}{q} \tag{24}$$

These integrals have in general to be computed numerically, but we shall see in the next section that in the far field region this is not necessary.

4 The Far Fields

In the far field the integrals in equations (23) and (24) can be computed with the stationary-phase method. Introducing polar coordinates $x = \rho \cos\phi$, $y = \rho \sin\phi$ with ρ assumed large we easily see that the stationary point occurs at $q = k \cos\phi$. Using equations (8), (9), and (19), the far fields become

$$U^1 = -\beta^2 \sqrt{8\pi/k_1\rho}\; e^{i(k_1\rho - \pi/4)} \tan\phi \frac{\sqrt{k_2^2 - k_1^2\cos^2\phi}}{k_1 \sin\phi + \beta\sqrt{k_2^2 - k_1^2\cos^2\phi}}$$

$$\tan\phi_0 \frac{\sqrt{k_1^2 - k_2^2\cos^2\phi_0}}{\sqrt{k_1^2 - k_2^2\cos^2\phi_0} + \beta k_2 \sin\phi_0}$$

$$\sum_{n,n'} nn' J_n(k_1 a\cos\phi) J_{n'}(k_2 a\cos\phi_0) Q^{-1}_{nn'} \tag{25}$$

$$U^2 = -\beta\sqrt{8\pi/k_2\rho}\; e^{i(k_2\rho - \pi/4)} \tan\phi \frac{\sqrt{k_1^2 - k_2^2\cos^2\phi}}{\sqrt{k_1^2 - k_2^2\cos^2\phi} - \beta k_2 \sin\phi}$$

$$\tan\phi_0 \frac{\sqrt{k_1^2 - k_2^2\cos^2\phi_0}}{\sqrt{k_1^2 - k_2^2\cos^2\phi_0} + \beta k_2 \sin\phi_0}$$

$$\sum_{n,n'} nn' J_n(k_2 a\cos\phi) J_{n'}(k_2 a\cos\phi_0) Q^{-1}_{nn'} \tag{26}$$

where we have introduced the direction of propagation of the incoming wave ϕ_0 by $q_0 = k_2 \cos\phi_0$. We note that the reciprocity relations that U^1 is symmetrical under the ex-

change $k_2 \cos \phi_0 \leftrightarrow k_1 \cos \phi$ and U^2 under the exchange $\phi_0 \leftrightarrow -\phi$ are satisfied.

In the far field region it is easy to check the consequences of energy conservation and this we now do. The mean value in time of the total energy flux in the radial direction is

$$\langle P \rangle = \frac{1}{2} \omega \, Im \int_0^{2\pi} \mu \, u^* \frac{\partial u}{\partial \rho} \rho d\phi \tag{27}$$

where the star denotes the complex conjugate. As no energy is created or destroyed inside the circle with radius ρ we must evidently have

$$\langle P \rangle = 0 \tag{28}$$

Introduce now the far field amplitudes F_1 and F_2:

$$U^1 = \sqrt{8\pi/k_1\rho} \, e^{i(k_1\rho - \pi/4)} F_1(\phi) \tag{29}$$

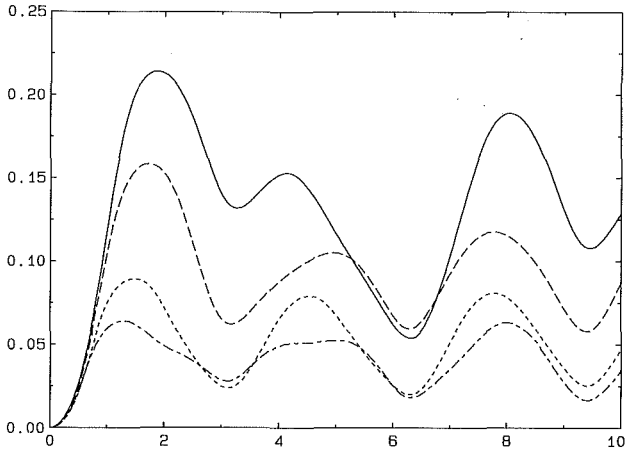


Fig. 6 Same as Fig. 5 but direction of incidence ϕ_0 = 60 deg

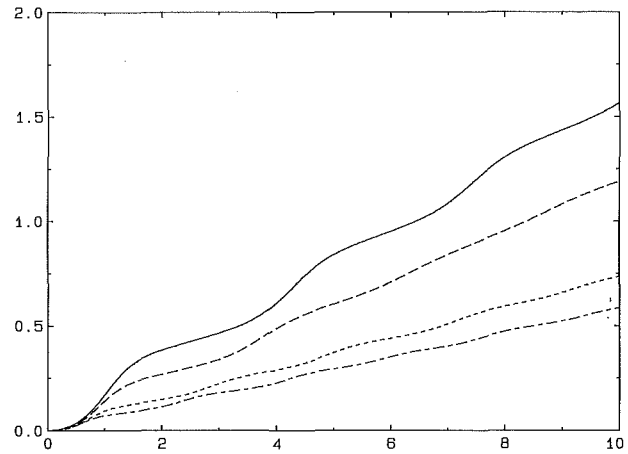


Fig. 7 Same as Fig. 5 but direction of incidence ϕ_0 = 90 deg

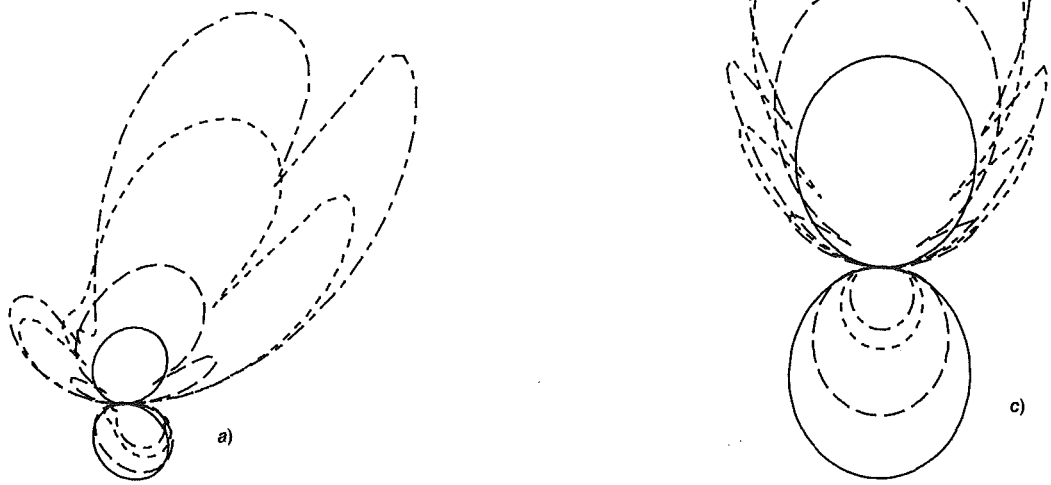


Fig. 8 The absolute value of the far field amplitude as a function of angle for the frequency $k_2 a$ = 2 and the same material combinations as in Figs. 2-7 and direction of incidence: (*a*) ϕ_0 = 30 deg; (*b*) ϕ_0 = 60 deg; (*c*) ϕ_0 = 90 deg

$$U^2 = \sqrt{8\pi/k_2\rho}\, e^{i(k_2\rho - \pi/4)} F_2(\phi) \tag{30}$$

where the expressions for F_1 and F_2 can thus be read off from equations (25) and (26). Inserting equations (6), (7), (29), and (30) into equation (27) and performing some of the resulting integrals by the stationary-phase method, equation (28) finally yields

$$\int_0^{\pi} |F_1(\phi)|^2 d\phi + \beta \int_{\pi}^{2\pi} |F_2(\phi)|^2 d\phi + \mathrm{Re}[T^* F_1(\phi_t) + \beta R^* F_2(\phi_r)] = 0 \tag{31}$$

Here ϕ_t and ϕ_r are the directions of the transmitted and reflected plane waves and are thus determined by Snell's law, i.e.,

$$\sin\phi_t = h_{01}^*/k_1, \qquad \cos\phi_t = q_0/k_1$$

$$\sin\phi_r = -h_{02}/k_2, \quad \cos\phi_r = q_0/k_2$$

The energy conservation as expressed by equation (31) should be compared with the corresponding expression given by Neerhoff (1979) for the case with an interface crack in a layered half-space. The energy conservation is a valuable check in numerical work and we thus use it during the numerical implementation.

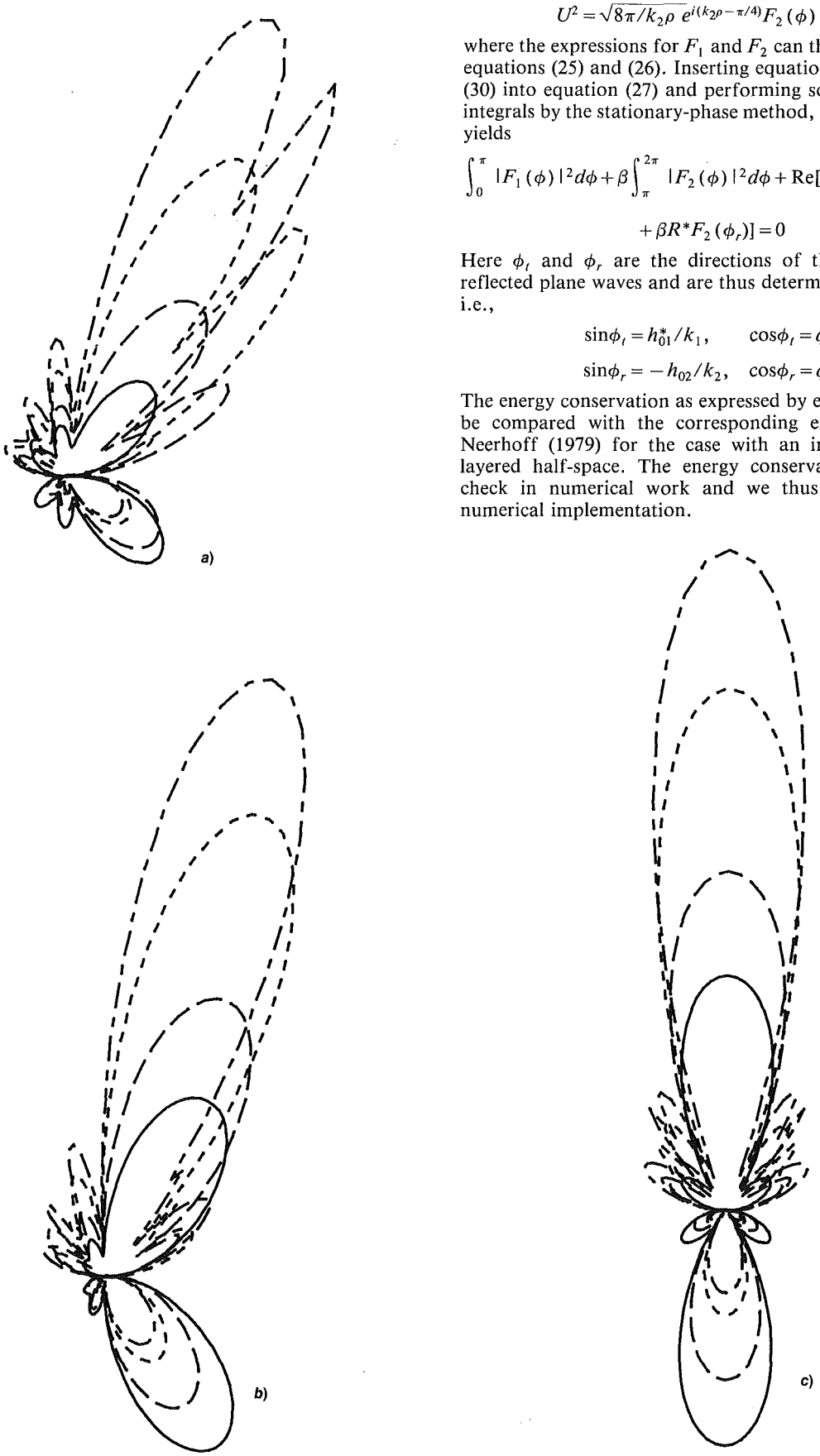


Fig. 9 The absolute value of the far field amplitude as a function of angle for the frequency $k_2a = 5$ and the same material combinations as in Figs. 2–8 and direction of incidence: (a) $\phi_0 = 30$ deg; (b) $\phi_0 = 60$ deg; (c) $\phi_0 = 90$ deg

5 Numerical Examples

We now turn to some numerical applications of the foregoing. We only compute far-field quantities, but we give both the total scattered energy and the back-scattered far field amplitude as functions of frequency and the far-field amplitude as a function of angle. The far-field amplitude was defined in the previous section, cf equations (25), (26), (29), and (30), and the total crack-scattered energy normalized with the energy density $1/2\ \omega\ \mu\ k_2$ in the incoming plane wave and the crack width $2a$ is

$$\sigma=\frac{4\pi}{k_2 a}\left(\beta^{-1}\int_0^{\pi}|F_1(\phi)|^2 d\phi+\int_{\pi}^{2\pi}|F_2(\phi)|^2 d\phi\right) \tag{32}$$

In the case of no contrast between the half-spaces this reduces to the usual definition of the total scattering cross section. By means of the condition of energy conservation, equation (31), the total scattered energy can be written in a form that contains no integrals.

All the computations are easy to perform and one only has to ascertain that the truncation, i.e., the number of terms in the summations in equations (25) and (26), is sufficiently high. The truncations needed depend on the frequency and by trial and error it is seen that it is in fact enough to take a number of terms that are slightly larger than max $(k_1 a, k_2 a)$.

For the numerical computations, four different material combinations have been used: $c_2/c_1 = 1$ and $\beta = \mu_2/\mu_1 = 1$ (a homogeneous space with a crack), $c_2/c_1 = 1.2$ and $\beta = 2$, $c_2/c_1 = 1.5$ and $\beta = 5$, and $c_2/c_1 = 1.8$ and $\beta = 8$. The lower half-space thus has the larger stiffness and wavespeed. The incoming plane wave in the lower half-space propagates in a direction which makes the angle $\phi_0 = 30$ deg, 60 deg, or 90 deg with the interface.

Of the numerical results we first show the total crack-scattered energy as a function of frequency $k_2 a$ for the incoming directions $\phi_0 = 30$ deg, 60 deg, and 90 deg in Figs. 2, 3, and 4, respectively. All four material combinations are shown in each figure: $c_2/c_1 = 1$ and $\beta = 1$ is shown with a full line, $c_2/c_1 = 1.2$ and $\beta = 2$ with a dashed line, $c_2/c_1 = 1.5$ and $\beta = 5$ with a dotted line, and $c_2/c_1 = 1.8$ and $\beta = 8$ with a dashed-dotted line. For the homogeneous cracked space the total scattered energy should approach $2 \cos \phi_0$ (= 1, 1.73, and 2 in Figs. 2–4, respectively) at higher frequencies and this is well satisfied. With increasing contrasts between the half-spaces the crack-scattered energy becomes smaller as the interface reflects some energy even in the absence of the crack.

In Figs. 5–7 the back-scattered far field amplitude is shown as a function of frequency for the same directions of incidence and material combinations as in Figs. 2–4. As for the total crack-scattered energy the values decrease with increasing contrast. For $\phi_0 = 30$ deg and 60 deg the curves have a periodic behavior with the peaks appearing when the wavefront of the incoming wave contains an integral number of half-wavelengths on the part that can be projected on the crack.

The angular dependence of the far field amplitude is shown in Figs. 8 and 9 for the same directions of incidence and material combinations as before and for the two frequencies $k_2 a = 2$ and $k_2 a = 5$, respectively. Figures 8(*a*), 8(*b*), and 8(*c*) which differ by the direction of the incoming wave are drawn to the same scales and likewise for Figs. 9(*a*), 9(*b*), and 9(*c*). As the contrasts between the half-spaces increase the far field amplitude in the upper half-space grows and develops additional side peaks. The growth is, however, not associated with any growth in crack-radiated energy, cf Figs. 2–4, as the material in the upper half-space becomes softer with increasing contrasts. The additional side peaks can be attributed to the fact that $k_1 a$ increases with increasing contrasts. In Figs. 9(*a*) and 9(*b*) we notice that the main lobe in the upper half-space is closer to the normal for larger contrasts quite in accordance with Snell's law. The main lobe in the lower half-space is, on the other hand, in the direction of specular reflection independently of the contrasts.

6 Concluding Remarks

We have seen how by a modification of the integral equation approach of Krenk and Schmidt (1982) it is possible to treat an interface crack. Only the two-dimensional case with antiplane strain has been considered in the present paper. More general cases can, however, also be solved with the present approach. The three-dimensional acoustic (scalar) problem with a soundhard (or soft) disk at an interface can thus be solved by the present approach and work in this direction is in progress. It will also be possible to treat the even more interesting case with a penny-shaped crack (which could be fluid-filled) at the interface between two elastic materials.

Acknowledgment

The present work is sponsored by the National Swedish Board for Technical Development (STU) and this is gratefully acknowledged.

References

Achenbach, J. D., Gautesen, A. K., and McMaken, H., 1978, "Diffraction of Pointsource Signals by a Circular Crack," *Bull. Seism. Soc. Am.*, Vol. 68, pp. 889–905.

Boström, A., and Olsson, P., 1987, "Scattering of Elastic Waves by Non-Planar Cracks," *Wave Motion*, Vol. 9, pp. 61–76.

Keogh, P. S., 1985a, "High-Frequency Scattering by a Griffith Crack I: A Crack Green's Function," *Q. J. Mech. Appl. Math.*, Vol. 38, pp. 185–204.

Keogh, P. S., 1985b, "High-Frequency Scattering by a Griffith Crack II: Incident Plane and Cylindrical Waves," *Q. J. Mech. Appl. Math.*, Vol. 38, pp. 205–232.

Krenk, S., and Schmidt, H., 1982, "Elastic Wave Scattering by a Circular Crack," *Phil. Trans. R. Soc. Lond.*, Vol. A 308, pp. 167–198.

Kundu, T., 1986, "Transient Response of an Interface-Crack in a Layered Plate," ASME Journal of Applied Mechanics, Vol. 53, pp. 579–586.

Martin, P. A., and Wickham, G. R., 1983, "Diffraction of Elastic Waves by a Penny-Shaped Crack: Analytical and Numerical Results," *Proc. R. Soc. Lond.*, Vol. A 390, pp. 91–129.

Neerhoff, F. L., 1979, "Diffraction of Love Waves by a Stress-Free Crack of Finite Width in the Plane Interface of a Layered Composite," *Appl. Sci. Res.*, Vol. 35, pp. 237–249.

Olsson, P., 1987, "Elastodynamic Scattering by Fluid-Filled Non-Planar Cracks," *J. Nondestructive Evaluation*, Vol. 5, pp. 161–168.

van den Berg, P. M., 1981, "Transition Matrix in Acoustic Scattering by a Strip," *J. Acoust. Soc. Am.*, Vol. 70, pp. 615–619.

Yang, H. J., and Bogy, D. B., 1985, "Elastic Wave Scattering From an Interface Crack in a Layered Half-Space," ASME Journal of Applied Mechanics, Vol. 52, pp. 42–50.

Rayleigh-Lamb Waves in an Elastic Plate With Voids

D. S. Chandrasekharaiah
Department of Mathematics,
Bangalore University,
Central College Campus,
Bangalore, India 560 001

Rayleigh-Lamb waves in a homogeneous and isotropic linear elastic plate containing a distribution of vacuous pores (voids) are studied. Assuming that the plate is of uniform thickness and that its faces are stress-free, it is found that the waves move, in general, in two uncoupled families, of which one is symmetrical with respect to the midplane of the plate and the other antisymmetrical; each of these families is affected by the presence of voids. If the plate is thin and the frequency is small, the voids influence only the symmetric waves and, because of this influence, the waves propagate slower than their classical counterparts. If the plate thickness and the frequency are large, each of the two families degenerates into two uncoupled waves; one of these is a classical Rayleigh wave and the other is a new wave not encountered in the classical theory.

1 Introduction

The theory of elastic materials with voids is one of the most recent generalizations of the classical theory of elasticity. This theory is concerned with elastic materials consisting of a distribution of small pores (voids) which contain nothing of mechanical or energetic significance. The general version of this theory was obtained by Nunziato and Cowin (1979), and the linearized version by Cowin and Nunziato (1983). A novel feature of this theory, over other theories on porous materials, is that it permits a porous body to enlarge or reduce the overall volume the body occupies in the absence of body forces. It is believed that the new theory is of practical utility in investigating various types of geological, biological, and synthetic porous materials for which the classical theory is inadequate. Some problems revealing interesting characterizations of the theory have been considered by Cowin and Nunziato (1983), Cowin and Puri (1983), Passman (1984), Cowin (1984 a, b; 1985 a, b), Puri and Cowin (1985), and Chandrasekharaiah (1986, 1987). Some basic theorems and properties of solutions have been obtained by Iesan (1985); the interrelationships between this theory and various other continuum theories have been analyzed by Cowin (1984 b; 1985 b).

The object of this paper is to discuss the propagation of free plane waves (of the Rayleigh-Lamb type) in a homogeneous and isotropic elastic plate with voids, by employing the field equations obtained by Cowin and Nunziato (1983). Assuming that the plate is of uniform thickness and that the faces are stress-free, we find that there occur, as in the corresponding classical problem, two uncoupled families of waves (in general) of which one consists of symmetrical motions about the midplane of the plate and the other antisymmetrical motions; each of these families is influenced by the presence of voids. We obtain the frequency equations for the two families and find that the voids have no influence on the cutoff frequencies of pure shear motions. As limiting cases, we consider small frequency waves in a thin plate and high frequency waves in a plate of large thickness. In the former case, we find that the voids have influence only on the symmetric family and that, because of this influence, the symmetric family propagates *slower* than its classical counterpart; for a typical (hypothetical) material model (Puri and Cowin, 1985) the decrease in the speed is found to be about 6.64 percent. In the latter case, each of the two (symmetric and antisymmetric) families degenerates into two uncoupled waves, of which one is a classical Rayleigh type wave not influenced by the presence of voids and the other is a new wave (caused by the porosity of the material) not encountered in the classical theory. A comparison between the present analysis and its counterpart in thermoelasticity is made at appropriate places.

2 Basic Equations and Boundary Conditions

In the context of the theory presented by Cowin and Nunziato (1983), the field equations for a homogeneous and isotropic material, in the absence of body forces, are given as follows:

$$\mu \nabla^2 u_i + (\lambda+\mu) u_{k,ki} + \beta \phi_{,i} = \rho \frac{\partial^2 u_i}{\partial t^2} \tag{2.1}$$

$$\alpha \nabla^2 \phi - \xi \phi - \omega \frac{\partial \phi}{\partial t} - \beta u_{k,k} = \rho k \frac{\partial^2 \phi}{\partial t^2} \tag{2.2}$$

In these equations, u_i is the displacement vector, ϕ is the so-called volume fraction field (defined in Cowin and Nunziato, 1983), λ, μ are the usual elastic constants, ρ is the mass density, α, β, ξ, ω, and k are new material constants characterizing the presence of voids, and t is time. The notation of Cartesian tensors is also adopted. In the absence of voids, we have $\phi = 0$, and equation (2.1) reduces to the classical Navier's equation.

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 23, 1986; final revision, September 15, 1986.

The stress tensor τ_{ij} associated with u_i and ϕ is given by (Cowin and Nunziato, 1983)

$$\tau_{ij} = \lambda\, \delta_{ij}\, u_{k,k} + \mu(u_{i,j} + u_{j,i}) + \beta\delta_{ij}\phi \tag{2.3}$$

If we set

$$u_i = p_{,i} + \epsilon_{ijk}\, q_{k,j} \tag{2.4}$$

then equation (2.1) is satisfied, provided p and q_i are governed by the following equations:

$$\left(\nabla^2 - \frac{1}{a^2}\frac{\partial^2}{\partial t^2}\right)p = -\frac{\beta}{\rho a^2}\,\phi \tag{2.5}$$

$$\left(\nabla^2 - \frac{1}{b^2}\frac{\partial^2}{\partial t^2}\right)q_i = 0 \tag{2.6}$$

Eliminating ϕ from equations (2.5) and (2.2) we obtain

$$\left[\left\{\nabla^2 - \frac{1}{\alpha^*}\left(1 + \omega^*\frac{\partial}{\partial t} + k^*\frac{\partial^2}{\partial t^2}\right)\right\}\left\{\nabla^2 - \frac{1}{a^2}\frac{\partial^2}{\partial t^2}\right\} + \beta^*\nabla^2\right]p = 0 \tag{2.7}$$

In the above equations, we have set

$$a^2 = \frac{\lambda + 2\mu}{\rho},\quad b^2 = \frac{\mu}{\rho},\quad \alpha^* = \frac{\alpha}{\xi}$$

$$\omega^* = \frac{\omega}{\xi},\quad k^* = \frac{\rho k}{\xi},\quad \beta^* = \frac{\beta^2}{\rho\alpha a^2} \tag{2.8}$$

From equations (2.5)–(2.7), it is evident that of the three unknowns p, q_i, and ϕ, only p and q_i are independent and that pure shear waves are not affected by the presence of voids. If in a given problem, p and q_i are determined by solving equations (2.7) and (2.6), then ϕ follows from (2.5).

Substituting for u_i from (2.4) in equation (2.3), and using equations (2.5) and (2.8), we obtain the following expression for τ_{ij} in terms of p and q_i:

$$\tau_{ij} = \mu\left[2p_{,ij} - \left\{2\nabla^2 - \frac{1}{b^2}\frac{\partial^2}{\partial t^2}\right\}\delta_{ij}p + \epsilon_{irs}\, q_{s,rj} + \epsilon_{jrs}\, q_{s,ri}\right] \tag{2.9}$$

If the boundary of the body is free of external loads, the following conditions hold on the boundary (Cowin and Nunziato, 1983):

$$\tau_{ij}\, n_j = 0 \tag{2.10}$$

$$\phi_{,j}\, n_j = 0 \tag{2.11}$$

Here, n_i is the unit outward normal to the boundary. By substituting for τ_{ij} from equation (2.9) we may express the boundary condition (2.10) in terms of p and q_i.

Under the assumptions made, equations (2.5)–(2.7) and equations (2.10), (2.11) serve as field equations and boundary conditions for the theory considered.

3 Plane Waves in a Plate

Suppose that the body considered is a plate occupying the Cartesian space $-\infty < x < \infty$, $-H \le y \le H$, $-\infty < z < \infty$ and that free plane waves propagate in the plate in the positive x direction causing plane deformation parallel to the xy plane. Then the displacement vector lies completely parallel to the xy plane and all the field variables depend only on x, y, and t. Under these conditions, only the z component, q_3, of q_i is relevant; we denote it by q. We denote the x and y components of u_i by u and v, respectively.

If the waves propagate with frequency $\theta/2\pi$, we seek, as usual, solutions for p and q in the form

$$(p, q) = (p_0, q_0)\exp\left[i(\gamma x - \theta t)\right] \tag{3.1}$$

where p_0 and q_0 are functions of y and γ is a complex number. If $\gamma = \gamma_1 + i\gamma_2$, then for the waves to be physically realistic, we should have $\gamma_1 > 0$ and $\gamma_2 \ge 0$; $2\pi/\gamma_1$ represents the wave length.

Putting equation (3.1) into equations (2.6) and (2.7), we arrive at the following expressions for p_0 and q_0:

$$p_0 = A_1\cosh m_1 y + A_2\cosh m_2 y + B_1\sinh m_1 y + B_2\sinh m_2 y \tag{3.2}$$

$$q_0 = A_3\sinh m_0 y + B_3\cosh m_0 y \tag{3.3}$$

Here A_i and B_i are arbitrary constants,

$$m_0^2 = \gamma^2 - \frac{\theta^2}{b^2} \tag{3.4}$$

and m_1, m_2 are (complex) roots of the equation

$$(\gamma^2 - m^2)^2 - \left[\frac{\theta^2}{a^2} - \frac{1}{\alpha^*}(1 - i\omega^*\theta - k^*\theta^2) + \beta^*\right](\gamma^2 - m^2) - \frac{\theta^2}{a^2\alpha^*}(1 - i\omega^*\theta - k^*\theta^2) = 0 \tag{3.5}$$

With the aid of equations (3.1) and (3.2), equation (2.5) yields

$$\phi = -\frac{\rho a^2}{\beta}\left[r_1(A_1\cosh m_1 y + B_1\sinh m_1 y) + r_2(A_2\cosh m_2 y + B_2\sinh m_2 y)\right]\exp\{i(\gamma x - \theta t)\} \tag{3.6}$$

where

$$r_{1,2} = m_{1,2}^2 - \gamma^2 + \frac{\theta^2}{a^2} \tag{3.7}$$

If the faces $y = \pm H$ of the plate are stress-free, the boundary conditions (2.10) and (2.11) become $\tau_{12} = \tau_{22} = \partial\phi/\partial y = 0$ for $y = \pm H$. With the aid of equations (2.9), (3.1)–(3.3), and (3.6), these conditions yield the following system of equations:

$$\begin{aligned}
2i\gamma(m_1 s_1 A_1 + m_2 s_2 A_2) + \gamma_0 s_0 A_3 &= 0\\
2i\gamma(m_1 c_1 B_1 + m_2 c_2 B_2) + \gamma_0 c_0 B_3 &= 0\\
\gamma_0(c_1 A_1 + c_2 A_2) - 2i\gamma m_0 c_0 A_3 &= 0\\
\gamma_0(s_1 B_1 + s_2 B_2) - 2i\gamma m_0 s_0 B_3 &= 0\\
r_1 m_1 s_1 A_1 + r_2 m_2 s_2 A_2 &= 0\\
r_1 m_1 c_1 B_1 + r_2 m_2 c_2 B_2 &= 0
\end{aligned} \tag{3.8}$$

Here we have set

$$s_{1,2} = \sinh m_{1,2}H,\quad c_{1,2} = \cosh m_{1,2}H,\quad \gamma_0 = 2\gamma^2 - \frac{\theta^2}{b^2}$$

$$s_0 = \sinh m_0 H,\quad c_0 = \cosh m_0 H \tag{3.9}$$

From equations (3.8), we see that all the A's are linked together, all the B's are linked together, and A's and B's are unlinked. It may be verified that for the waves of the desired type to exist, at least one of the two constants A_1 and B_1 should be nonzero; if $B_1 = 0$, the solution (3.1) corresponds to waves symmetrical with respect to the midplane ($y = 0$) of the plate and if $A_1 = 0$, the waves are antisymmetrical. In general, the solution (3.1) represents a superposition of two uncoupled families of waves, of which one is symmetrical and the other antisymmetrical with respect to the plane $y = 0$.

4 Frequency Equation

We now proceed to obtain the characteristic equations determining the phase speeds of symmetric and antisymmetric families of waves.

If we put

$$\Gamma = \gamma H,\quad M_0 = m_0 H,\quad M_{1,2} = m_{1,2}H$$

$$\Omega=\frac{\theta}{b}H, \qquad V=\frac{\Omega b}{\Gamma},$$

$$R_{1,2}^2=1-\frac{a^2}{\theta^2}\,r_{1,2} \tag{4.1}$$

then equations (3.4) and (3.7) yield

$$M_0^2=\Gamma^2-\Omega^2=\Gamma^2\left(1-\frac{V^2}{b^2}\right)$$

$$M_{1,2}^2=\Gamma^2-\frac{\Omega^2b^2}{a^2}R_{1,2}^2=\Gamma^2\left(1-\frac{V^2}{a^2}R_{1,2}^2\right) \tag{4.2}$$

It is evident that V represents the complex phase speed, only the real part of which is physically relevant.

With the aid of equations (4.1) and (4.2), we find from equation (3.5) that R_1 and R_2 satisfy the equation

$$\Omega^*R^4+(1-N^*-\Omega^*)R^2-1=0 \tag{4.3}$$

where

$$N^*=\frac{\alpha^*\beta^*H^2}{H^2-i\omega^*\Omega bH-k^*\Omega^2b^2} \tag{4.4}$$

and

$$\Omega^*=\frac{\alpha^*\Omega^2b^2}{a^2[H^2-i\omega^*\Omega bH-k^*\Omega^2b^2]} \tag{4.5}$$

Eliminating the constants A_i and B_i from the system of equations (3.8) and simplifying the resulting expressions with the aid of equations (3.9), (4.1), and (4.2), we arrive at the following two frequency equations:

$$(2\Gamma^2-\Omega^2)^2[M_2(1-R_2^2)T_1^{\pm1}-M_1(1-R_1^2)T_2^{\pm1}]$$
$$=4\Gamma^2M_0M_1M_2(R_1^2-R_2^2) \tag{4.6}$$

Here we have put

$$T_{1,2}=\frac{c_0s_{1,2}}{s_0c_{1,2}} \tag{4.7}$$

Of the two equations given in (4.6), the one which contains T_1^{-1} and T_2^{-1}, obtained by eliminating A's from equations (3.8), corresponds to the symmetric family of waves. The other equation which contains T_1 and T_2, obtained by eliminating B's from equations (3.8), corresponds to the antisymmetric family.

In view of equations (4.2), equations (4.6) may be regarded as equations connecting V and Γ as well. Accordingly, each of these two equations, being transcendental, yields infinitely many discrete roots for V in terms of Γ, each root corresponding to a mode of vibration. The motions – both symmetric and antisymmetric – are obviously dispersive and the analysis of their behavior in the general case is complicated. However, it is possible to obtain readily the cutoff frequencies by setting $\Gamma = 0$ in the frequency equations (4.6). The cutoff frequencies so obtained are governed by the following equations:

(*i*) For symmetric family

$$\Omega=n\pi,\ n=1,2,\ldots.\infty \tag{4.8}$$

$$R_2(1-R_2^2)\tan(\Omega bR_2/a)=R_1(1-R_1^2)\tan(\Omega bR_1/a) \tag{4.9}$$

(*ii*) For antisymmetric family

$$\Omega=(2n-1)\frac{\pi}{2},\ n=1,2,\ldots.\infty \tag{4.10}$$

$$R_2(1-R_2^2)\tan(\Omega bR_1/a)=R_1(1-R_1^2)\tan(\Omega bR_2/a) \tag{4.11}$$

Evidently, the cutoff frequencies given by equations (4.8) and (4.10) are not influenced by the presence of voids and are associated with pure shear motions. The cutoff frequencies determined by equations (4.9) and (4.11) are associated with the compressional motions influenced by the presence of voids. Because of the presence of the material constants α^*, β^*, and ω^* in the governing equation (4.3) of R_1 and R_2, the cutoff frequencies determined by the transcendental equations (4.9) and (4.11) differ from those occurring in the classical elasticity theory.

In the absence of voids, we may take $R_1 = 1$ and $R_2 = 0$ (see, equations (4.3)–(4.5)). Then the frequency equations (4.6) reduce, with the aid of equations (4.2), (4.7), and (3.9), to

$$\left[\frac{\tanh\left\{\Gamma\left(1-\frac{V^2}{a^2}\right)^{1/2}\right\}}{\tanh\left\{\Gamma\left(1-\frac{V^2}{b^2}\right)^{1/2}\right\}}\right]^{\pm1}$$
$$=\frac{4\left(1-\frac{V^2}{a^2}\right)^{1/2}\left(1-\frac{V^2}{b^2}\right)^{1/2}}{\left(2-\frac{V^2}{b^2}\right)^2} \tag{4.12}$$

These are well-known frequency equations for Rayleigh-Lamb waves in classical elasticity. (Apart from the notation, these equations are identical with equations (12) and (48) of Lamb, 1917.)

It may be noted that the field equations (2.1) and (2.2) and the constitutive equation (2.3) resemble the corresponding equations in linear coupled thermoelasticity; the boundary condition (2.11) is analogous to the condition of thermal insulation. Cowin (1985b) has discussed some aspects of the similarities between the theory of elastic materials with voids and the thermoelasticity theory. In the context of thermoelasticity, the problem of Rayleigh-Lamb waves has been analyzed, among others, by Nowacki (1975). We find that our frequency equations (4.6) are analogous to equations (15) and (16) on p. 208 of Nowacki (1975). However equation (3.5) and its counterpart in thermoelasticity, viz. equation (8) on p. 207 of Nowacki (1975), have different structures.

5 Limiting Cases

In what follows, we analyze equations (4.6) in two limiting cases.

Case (*i*) Waves in a Thin Plate. We first suppose that $|\Gamma| << 1$, $\Omega \to 0$. In this case, which obviously corresponds to small values of H and θ, we find from equations (4.2) and (4.3) that

$$R_1=(1-N)^{-1/2}, \quad R_2\to\infty.$$

Here

$$N=\alpha^*\beta^*=\frac{\beta^2}{\rho a^2\xi} \tag{5.1}$$

Equations (4.6) now yield the following roots:

(*i*) For symmetric family

$$V=2b\left[1-\frac{b^2}{a^2(1-N)}\right]^{1/2} \tag{5.2}$$

(*ii*) For antisymmetric family

$$V=2b\Gamma\left[\frac{1}{3}\left(1-\frac{b^2}{a^2}\right)\right]^{1/2} \tag{5.3}$$

Evidently, the antisymmetric family is *not* affected by the presence of voids; its speed, given by equation (5.3), is identical with that obtained in the classical theory (see equation (50) of Lamb, 1917).

On the other hand, the symmetric family is *affected* by the presence of voids; while its speed in the classical theory is

$$V_s = 2b\left(1-\frac{b^2}{a^2}\right)^{1/2} \tag{5.4}$$

(see equation (15) of Lamb, 1917), in the presence of voids the speed is modified to

$$V_s^* = 2b\left(1-\frac{b^2}{a^{*2}}\right)^{1/2} \tag{5.5}$$

where $a^* = a(1-N)^{1/2}$; see equation (5.2) above. It is known (Puri and Cowin, 1985) that the constant N satisfies the inequality $0 \leq N < 1$, with $N = 0$ holding in the absence of voids. Consequently, we have $V_s^* < V_s$ in the presence of voids. Thus, symmetric waves of small frequency in a thin elastic plate with voids are *slower* than their classical counterparts.

The values of the material constants characterizing the presence of voids are not known as yet for any material. In their analysis of plane waves in an unbounded elastic material with voids, Puri and Cowin (1985) have considered a typical (hypothetical) material model for which $a = 3873$ m/s, $b = 1937$ m/s, and $N = 0.2778$. For this model, we find that $V_s^* = 3132.0918$ m/s. For the corresponding classical model ($N=0$), we get $V_s = 3354.6933$ m/s. Thus, in this model, the reduction in the speed of small frequency symmetric waves in a thin plate, due to the presence of voids, is about 6.64 percent.

Puri and Cowin (1985) have shown that the constant $a^* = a(1-N)^{1/2}$ represents the speed of predominantly elastic low frequency waves in an unbounded elastic material with voids. As such, the expression for V_s^* obtained above is structurally consistent with the expression for V_s, valid in the classical theory.

In the linear thermoelasticity theory, it is known that small frequency dilatational waves propagate with speed $a_T = a(1+\epsilon)^{1/2}$, where $\epsilon > 0$ is the thermoelastic coupling constant (Nowacki, 1975, p. 108). From the frequency equation (15) on p. 208 of Nowacki (1975), it may be shown that the counterpart in thermoelasticity of our equation (5.5) above reads thus:

$$V_s^{**} = 2b\left(1-\frac{b^2}{a_T^2}\right)^{1/2} \tag{5.6}$$

This expression is also structurally similar to equation (5.4) valid in the classical elasticity theory, and we have $V_s^{**} > V_s$. Thus, while the presence of voids *decreases* the speed of small frequency symmetric waves in a thin elastic plate, the presence of the thermal field *increases* the speed; in both the cases the qualitative behavior of the waves remains unchanged.

Case (*ii*) Waves in a Plate of Large Thickness. We now suppose that $|\Gamma| > b/a\,\Omega\,|R_{1,2}|$. In this case, which obviously corresponds to large values of H, we find, with the aid of equations (4.2), that the two equations in (4.6) reduce to one and the same equation given below:

$$\Gamma^2\left[M_1^2 + M_1M_2 + M_2^2 + \frac{\Omega^2 b^2}{a^2} - \Gamma^2\right]\left(2-\frac{V^2}{b^2}\right)^2 = 4M_0M_1M_2(M_1+M_2) \tag{5.7}$$

Accordingly, in this case, both symmetric and antisymmetric motions behave alike. We find that equation (5.7) is identical, apart from the notation, with the frequency equation for surface waves in a halfspace with voids, obtained by Chandrasekharaiah (1986).

For $\Omega \to \infty$, equations (4.3) and (4.2) yield

$$M_1 = \Gamma\,(1-V^2/a^2)^{1/2}, \quad M_2 = \Gamma\,(1-V^2/V_\phi^2)^{1/2}, \quad M_0 = \Gamma\,(1-V^2/b^2)^{1/2} \tag{5.8}$$

where

$$V_\phi^2 = \frac{\alpha^*}{k^*} = \frac{\alpha}{\rho k} \tag{5.9}$$

With the aid of expressions (5.8), equation (5.7) leads to the following equation:

$$\left(1-\frac{V^2}{V_\phi^2}\right)^{1/2}\left[\left(1-\frac{V^2}{2b^2}\right)^2 - \left(1-\frac{V^2}{a^2}\right)^{1/2}\left(1-\frac{V^2}{b^2}\right)^{1/2}\right] = 0 \tag{5.10}$$

Evidently, each of the two families (symmetric and antisymmetric) now degenerates into two uncoupled waves. One of these is the classical Rayleigh wave not influenced by the presence of voids and the other is purely a volume fraction wave, caused by the presence of voids, propagating with speed V_ϕ. Thus, the effect of the presence of voids in this case is just to exhibit new waves not encountered in the classical theory; we note that their speed V_ϕ is identical with the speed of high frequency volume fraction waves occuring in an unbounded elastic material with voids (Puri and Cowin, 1985).

In the corresponding situation in thermoelasticity, it may be shown, from equations (15) and (14) on p. 208 of Nowacki (1975), that the presence of the thermal field does not exhibit any new wave; there occur only classical Rayleigh waves uninfluenced by the thermal field.

Acknowledgment

The author is thankful to the University Grants Commission, New Delhi, for providing financial assistance through research grant No. F.8-3/84, SR III.

References

Chandrasekharaiah, D. S., 1986, "Surface Waves in an Elastic Half-Space With Voids," *Acta Mech.*, Vol. 62, pp. 77–85.

Chandrasekharaiah, D. S., 1987, "Effects of Surface Stresses and Voids on Rayleigh Waves in an Elastic Solid," *International Journal of Engineering Science*, Vol. 25, pp. 205–211.

Cowin, S. C., 1984a, "A Note on the Problem of Pure Bending for a Linear Elastic Material With Voids," *Journal of Elast.*, Vol. 14, pp. 227-233.

Cowin, S. C., 1984b, "The Stresses Around a Hole in a Linear Elastic Material With Voids," *Quarterly Journal of Mechanics Applied Mathematics*, Vol. 37, pp. 441–465.

Cowin, S. C., 1985a, "The Viscoelastic Behavior of Linear Elastic Materials with Voids," *Journal of Elast.*, Vol. 15, pp. 185–192.

Cowin, S. C., 1985b, "Modelling Shrinkage Mechanisms in Porous Elastic Solids," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 351–355.

Cowin, S. C, and Nunziato, J. W., 1983, "Linear Elastic Materials With Voids," *Journal of Elast.*, Vol. 13, pp. 125–147.

Cowin, S. C., and Puri, P., 1983, "The Classical Pressure Vessel Problems for Linear Elastic Materials With Voids, *Journal of Elast.*, Vol. 13, pp. 157–163.

Iesan, D., 1985, "Some Theorems in the Theory of Elastic Materials With Voids," *Journal of Elast.*, Vol. 15, pp. 215–224.

Lamb, H., 1917, "On Waves in an Elastic Plate," *Proc. Roy. Soc. London*, Vol. 93A, pp. 114–128.

Nowacki, W., 1975, *Dynamic Problems of Thermoelasticity*, Noordhoff, Leyden.

Nunziato, J. W., and Cowin, S. C., 1979, A Non-Linear Theory of Elastic Materials With Voids, *Arch. Rat. Mech. Anal.*, Vol. 72, pp. 175–201.

Passman, S. L., 1984, "Stress Relaxation, Creep, Failure and Hysteresis in a Linear Elastic Material With Voids, *Journal of Elast.*, Vol. 13, pp. 157–163.

Puri, P., and Cowin, S. C., 1985, "Plane Waves in Linear Elastic Materials With Voids," *Journal of Elast.*, Vol. 15, pp. 167–183.

T. Belytschko
Professor of Civil and Mechanical/Nuclear Engineering.

Xiao-Jun Wang[1]

Z. P. Bazant
Professor of Civil Engineering.

Y. Hyun
Research Assistant.

Department of Civil Engineering, Northwestern University, Evanston, IL 60201

Transient Solutions for One-Dimensional Problems With Strain Softening

Closed-form solutions are presented for the transient response of rods in which strain softening occurs and the stress-strain laws exhibit nonvanishing stresses after the strain-softening regime. It is found that the appearance of any strain softening results in an infinite strain rate if the material is inviscid. For a stress-strain law with a monotonically decreasing stress the strains are infinite also. If the stress increases after the strain-softening portion, the strains remain finite and the strain-softening point moves through the rod.

Introduction

A negative slope is found in the constitutive equations for phenomena such as erosion in penetration, shear banding and other damage mechanisms. Yet, the understanding of the behavior of continua which are governed by such constitutive equations is very limited. In fact, Hadamard (1903) discarded the possibility of such continua by stating that the wavespeed is imaginary, so that the continuum cannot exist. Numerical solutions for such materials are also quite strange. For example, Belytschko et al. (1984, 1985) have recently shown that in spherical geometries, strain-softening models can lead to numerical solutions characterized by many large peaks in strain, and that the locations of these peaks depend very much on the mesh size. However, constitutive models with strain softening are so prevalent and important in practice that their behavior must be understood.

The only closed-form solutions for problems in which the stress tends monotonically to zero are those of Bazant and Belytschko (1985), who presented a transient solution for a one-dimensional rod problem. These solutions exhibited a localization of the strain softening to a domain of measure zero, a discontinuity in the displacement and a singularity in the strains at the point of strain softening. The argument of Hadamard was shown to be irrelevant since the strain softening does not occur in a finite domain. However, the energy dissipation in the strain softening domain was shown to vanish, which raised questions as the applicability of this constitutive model to damage.

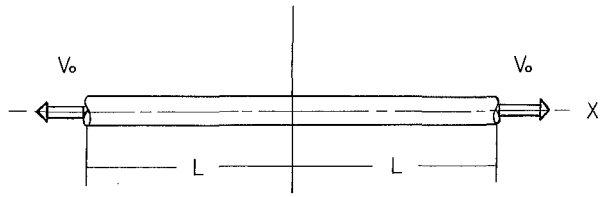


Fig. 1 Problem description: one-dimensional rod of length 2L with velocities prescribed at both ends

We consider here two different stress-strain laws in which the stress does not become a zero in the strain-softening branch of the stress-strain law. In the first material law, the slope of the stress-strain law after the onset of strain softening remains nonpositive; we call this law strain softening/perfectly plastic. In the second material law, the slope of the stress-strain law is negative for an interval and then reverses; we call this law strain softening-rehardening. It is found that if the stress decreases monotonically to any nonzero positive value after strain softening is initiated, a singularity appears in the strain, and the displacement is discontinuous. However, with the rehardening law, the strain remains finite and the strain softening point traverses through the material.

Very few closed-form or numerical-transient solutions with strain softening in which the slope of the stress-strain curve remains nonpositive have appeared in the literature. Some works relevant to this one are Bazant (1976), Aifantis and Serrin (1983), Wu and Freund (1984), Sandler and Wright (1984), Belytschko et al. (1984, 1985), Willam et al. (1984), and Schreyer and Chen (1984). For materials with rehardening, excellent theoretical studies have been reported by James (1980).

Problem Formulation

Consider a bar of length $2L$ with a unit cross section and mass ρ per unit length as shown in Fig. 1. The axis of the bar coincides with the coordinate x; the origin of the coordinate system is the midpoint of the rod so the interval of x is $[-L, +L]$. The equation of motion is

[1] Visiting Scholar, on leave from the University of Science and Technology of China.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 1, 1985; final revision, July 9, 1986.

Paper No. 87-WA/APM-3.

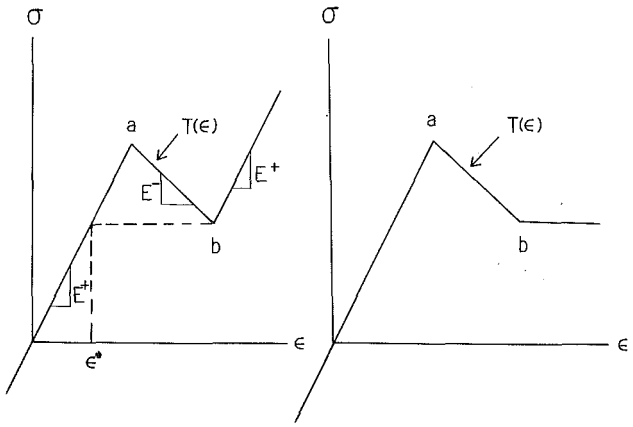


Fig. 2 Stress-strain laws for problems 1 and 2

$$\frac{\partial \sigma}{\partial x} = \rho \ddot{u} \tag{1}$$

where σ is the stress, $u(x,t)$ the displacement and superposed dots are time derivatives. The stress-strain law is taken to be

$$\dot{\sigma} = E(\sigma, \epsilon, \dot{\epsilon}) \frac{\partial v}{\partial x} \tag{2}$$

$$v = \dot{u} \tag{3}$$

In the elastic part of the response, $E > 0$ and equations (1)–(3) can be combined to yield

$$c^2 \frac{\partial^2 u}{\partial x^2} = \ddot{u} \qquad c^2 = \frac{E}{\rho} \tag{4}$$

Initially, the bar is undeformed and at rest so

$$u(x,0) = v(x,0) = 0 \quad -L \le x \le L \tag{5a}$$

The boundary conditions are

$$v(L,t) = v_o H(t) \tag{5b}$$

$$v(-L,t) = -v_o H(t) \tag{5c}$$

where $H(\ )$ is the Heaviside step function, v_o is a prescribed constant velocity, and t is the time.

Solutions

The solution to the above system is elastic until the stress associated with the onset of strain softening is reached (we will not be concerned with any purely elastic solutions). Strain softening always occurs first at the midpoint, where the stresses of the two elastic waves are superimposed and would reach a stress of twice the intensity of the initial waves if the material remained elastic.

The procedure of constructing a solution once strain softening is attained depends on the following hypothesis: strain softening is limited to a single point x_s (a set of measure zero) and at that point the strain instantaneously increases at least to where the stress attains a minimum value along the stress-strain curve, so after strain-softening $\epsilon \ge \epsilon_b$ (see Fig. 2).

Remark 1. This hypothesis was demonstrated in Bazant and Belytschko (1985). While this step may need more rigorous proof, it enables all of the governing equations to be satisfied; furthermore, it is borne out by numerical solutions.

Problem 1: Strain Softening-Rehardening. In the first problem, the stress-strain law is shown in Fig. 2(*a*). The stress-strain law can be characterized as follows:

initial conditions: $\sigma = \epsilon = 0; \quad S = \sigma_a$

algorithm: if $\epsilon > \epsilon_a$ and $S > \sigma_b$ and $\dot{\epsilon} > 0$

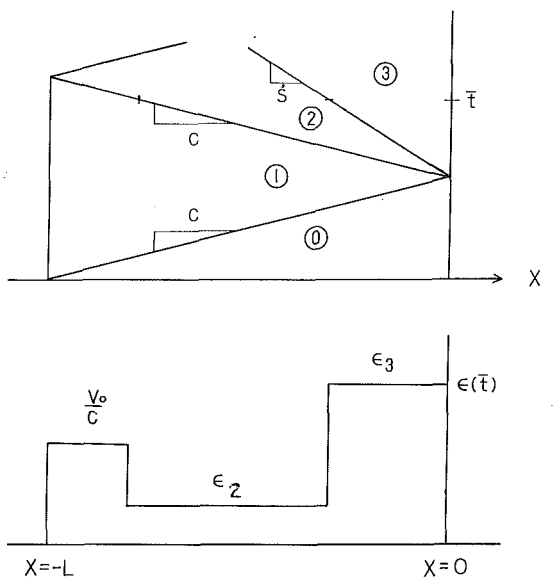


Fig. 3 Wavefronts in problem 1 and the strain distribution $\epsilon(\bar{t})$

then $\dot{\sigma} = E^- \dot{\epsilon}, \ \dot{S} = \dot{\sigma}$ (6*a*)

otherwise $\dot{\sigma} = E^+ \dot{\epsilon}, \ \dot{S} = 0$ (6*b*)

In the above, S is a state variable for the material. An alternative algorithm can be written in difference form:

difference algorithm:

$$\sigma^{\text{new}} = \sigma^{\text{old}} + E^+ \Delta\epsilon \tag{6c}$$

if $\sigma^{\text{new}} > T(\epsilon^{\text{new}})$ then, replace above by $\sigma^{\text{new}} = T(\epsilon^{\text{new}})$ (6*d*)

$$T(\epsilon) = \begin{cases} E^+ \epsilon & \text{for } \epsilon \le \epsilon_a \\ \sigma_a + E^- (\epsilon - \epsilon_a) & \text{for } \epsilon_a < \epsilon < \epsilon_b \\ \sigma_b + E^+ (\epsilon - \epsilon_b) & \text{for } \epsilon_b < \epsilon \end{cases} \tag{6e}$$

It is assumed that $Ev_o/c > \sigma_a/2$, so that when the two waves meet at $x = 0$, strain softening is initiated, so that equation (6*a*) applies. Since it is hypothesized that at the strain softening point the strain jumps instantaneously, the stress then instantaneously takes on a value which we will call σ_s, $\sigma_s \ge \sigma_b$.

We then have 2 boundary-value problems (BVP):

BVP(A) $-L \le x \le 0$:

Equations (4) and (5*a*), $\sigma(0,t) = \sigma_s \qquad \frac{L}{c} \le t \le t_3$ (7*a*)

BVP(B) $0 x \le L$:

Equations (4) and (5*b*), $\sigma(0,t) = \sigma_s \qquad \frac{L}{c} \le t \le t_3$ (7*b*)

where t_3 is a time to be determined as part of the solution. Since these two BVPs are symmetric with respect to the origin, we consider only BVP(A).

It will be shown in the following that a solution to this problem can be found if σ_s assumes any value in the range

$$\sigma_b \le l_s \le \sigma_a \tag{8}$$

see Fig. 2. We will parametrize this family of solutions by the strain in the initial reflected wave, the strain in domain 2 in Fig. 3, which is denoted by ϵ_2.

The structure of the solution is shown in Fig. 3. In domain 1, behind the initial elastic wavefront, the velocity, stress, and strain are given by

$$\dot{u}_1 = -v_o \quad \epsilon_1 = \frac{v_o}{c} \quad \sigma_1 = \frac{E v_o}{c} \tag{9}$$

(subscripts on the left-hand variables designate the domain to which the variables pertain).

It will be shown that the reflected elastic wave moves faster than the wave associated with strain softening, so that we have elastic behavior in domain 2; hence

$$\epsilon_2 = \frac{\sigma_2}{E} \tag{10}$$

Also, as can be seen from Fig. 2(*a*), since $\sigma_s \geq \sigma_b$, the stress-strain law gives

$$\sigma_s = \sigma_b + E(\epsilon_2 - \epsilon^*) \tag{11}$$

and since $\sigma_s \geq \sigma_b$,

$$\epsilon_2 \geq \epsilon^* \tag{12}$$

where $\epsilon^* = \sigma_b/E$. The remainder of the solutions will be constructed by using the jump conditions

$$[\dot{u}] = \dot{s}[\epsilon] \quad [\sigma] = \rho \dot{s}^2 [\epsilon] \tag{13}$$

where [ ] designates a jump and $\dot{s}$ is the velocity of the discontinuity.

From the velocity jump condition between domains 1 and 2, we obtain

$$\dot{u}_2 - \dot{u}_1 = c(\epsilon_2 - \epsilon_1) \tag{14a}$$

which, upon the use of equations (9), gives

$$\dot{u}_2 = c\left(\epsilon_2 - \frac{2v_o}{c}\right) \tag{14b}$$

The displacement field in domains 1 and 2 is then given by

$$u(x,t) = -v_o < \xi - \frac{2x}{c} > + (c\epsilon_2 - v_o) < \xi > \tag{15}$$

where $\xi = t - (L-x)/c$ and $< f > = fH(\ )$. Hence $u(0,t) = (c\epsilon_2 - 2v_o) < t - L/c >$. Therefore, if the displacement field is to remain continuous at $x = 0$, another wave must emanate from that point; the only exception is the unusual situation where $c\epsilon_2 = 2v_o$, which will be examined later. The speed of this wave will be denoted by $\dot{s}$ and it represents the interface between domains 2 and 3 in Fig. 3.

The velocity-strain jump condition gives

$$\dot{s}(\epsilon_3 - \epsilon_2) = \dot{u}_3 - \dot{u}_2 = 2v_o - c\epsilon_2 \tag{16a}$$

where the last equality is obtained by noting $\dot{u}_3 = 0$ because of symmetry and using equation (14*b*). The stress jump condition gives

$$\sigma_3 - \sigma_2 = \rho \dot{s}^2 (\epsilon_3 - \epsilon_2) \tag{16b}$$

and the stress-strain law in the strain softening domain gives

$$\sigma_3 - \sigma_b = E(\epsilon_3 - \epsilon_b) \tag{16c}$$

Equations (16) are solved as follows: we can put equation (16*c*) in the form

$$\sigma_3 - \sigma_2 = \rho c^2 (\epsilon_3 - \epsilon_b - \epsilon_2 + \epsilon^*) \tag{17}$$

and using equations (16*b*) and (17) yields

$$\dot{s}^2 (\epsilon_3 - \epsilon_2) = c^2 [\epsilon_3 - \epsilon_b - \epsilon_2 + \epsilon^*] \tag{18}$$

Using equation (16*a*) to eliminate $\dot{s}$ from equation (18) yields a quadratic equation for ϵ_3

$$\epsilon_3^2 - \epsilon_3(2\epsilon_2 + \epsilon_b - \epsilon^*) + \epsilon_2(\epsilon_2 + \epsilon_b - \epsilon^*) - \left(\frac{2v_o}{c} - \epsilon_2\right)^2 = 0 \tag{19}$$

which gives a one-parameter of solutions for ϵ_3

$$\epsilon_3 = \frac{2\epsilon_2 + \epsilon_b - \epsilon^*}{2} + \left[\frac{1}{4}(\epsilon_b - \epsilon^*)^2 + \left(\frac{2v_o}{c} - \epsilon_2\right)^2\right]^{1/2} \tag{20}$$

in terms of the parameter ϵ_2. Only the solution with the positive sign on the radical has been selected in the above since it is necessary that $\epsilon_3 > \epsilon_b$; this inequality is violated with the negative sign.

Combining equations (16*b*) and (16*c*) to eliminate the stresses and using equation (16*a*) to then eliminate ϵ_3, we obtain the following equation for $\dot{s}$

$$\left(\frac{\dot{s}}{c}\right)^2 + 2A\left(\frac{\dot{s}}{c}\right) - 1 = 0 \tag{21a}$$

where

$$A = \frac{c(\epsilon_b - \epsilon^*)}{2(2v_o - c\epsilon_2)} \tag{21b}$$

Hence

$$\dot{s} = c(-A + \sqrt{1 + A^2}) \tag{22}$$

and it follows immediately that if $A > 0$, then by the triangle inequality $\dot{s} < c$. The condition that $A \geq 0$ is satisfied if

$$2v_o - c\epsilon_2 > 0 \tag{23}$$

which must be satisfied if strain softening is to be initiated.

Thus we have a one-parameter family of solutions for this problem in which the parameter ϵ_2 is restricted by

$$\epsilon^* \leq \epsilon_2 \leq \frac{\sigma_a}{E} \tag{24}$$

An interesting case, which we will see is usually obtained in numerical solutions of these equations, corresponds to $\epsilon_2 = \epsilon^*$. Equations (20) then becomes

$$\epsilon_3 = \frac{\epsilon_b + \epsilon^*}{2} + \left[\frac{(\epsilon_b - \epsilon^*)^2}{4} + \left(\frac{2v_o}{c} - \epsilon^*\right)^2\right]^{1/2} \tag{25}$$

The strain ϵ_3 can then be shown to be bounded by

$$\frac{1}{2}(2\epsilon_2 + \epsilon_b + \epsilon^*) \leq \epsilon_3 \leq \frac{2v_o}{c} + (\epsilon_b - \epsilon^*) \tag{26}$$

Note that if $2v_o \to c\epsilon^*$, equation (21*b*) shows that $A \to \infty$ and from equation (22), $\dot{s} \to 0$.

The solution for $-L \leq x < -s$ is then

$$u(x,t) = -v_o < \xi - \frac{2x}{c} > + (c\epsilon^* - v_o) < \xi > \tag{27a}$$

$$\epsilon = \frac{v_o}{c} H\left(\xi - \frac{2x}{c}\right) + \left(\epsilon^* - \frac{v_o}{c}\right) H(\xi) \tag{27b}$$

For $-s \leq x \leq 0$

$$u = \epsilon_3 x H < \xi > \tag{27c}$$

$$\epsilon = \epsilon_3 H < \xi > \tag{27d}$$

The character of the solution is shown in Fig. 3. A noteworthy feature which distinguishes it from an elastic-plastic solution is the unloading wave emanating from the center.

Remark 2. Although the point of strain softening moves in the solution, this does not contradict the statement in Bazant and Belytschko (1985) (for the case in which $\sigma_b = 0$ and $E \leq 0$ in the softening domain) that the strain softening/elastic interface must be stationary. In the case considered here with $\sigma_b \neq 0$ and E becoming positive again after softening, the strain softening occurs instantaneously and the point subsequently becomes elastic. Thus, the interface $s(t)$ can be considered to be between two elastic domains.

Remark 3. Note that if equation (16*a*) is satisfied, $\dot{s} > 0$ as required, since $\epsilon_3 > \epsilon_2$.

Remark 4. The solution poses some peculiar mathematical difficulties, for at the points $x = \pm s$ the stress takes on the values in the range $\sigma_a \leq \sigma \leq \sigma_s$ twice in one point in time; thus whether it is differentiable, and whether $\sigma_{,x}$ in the governing equation (1) is defined, is not clear.

Remark 5. The propagation of the jump discontinuity and

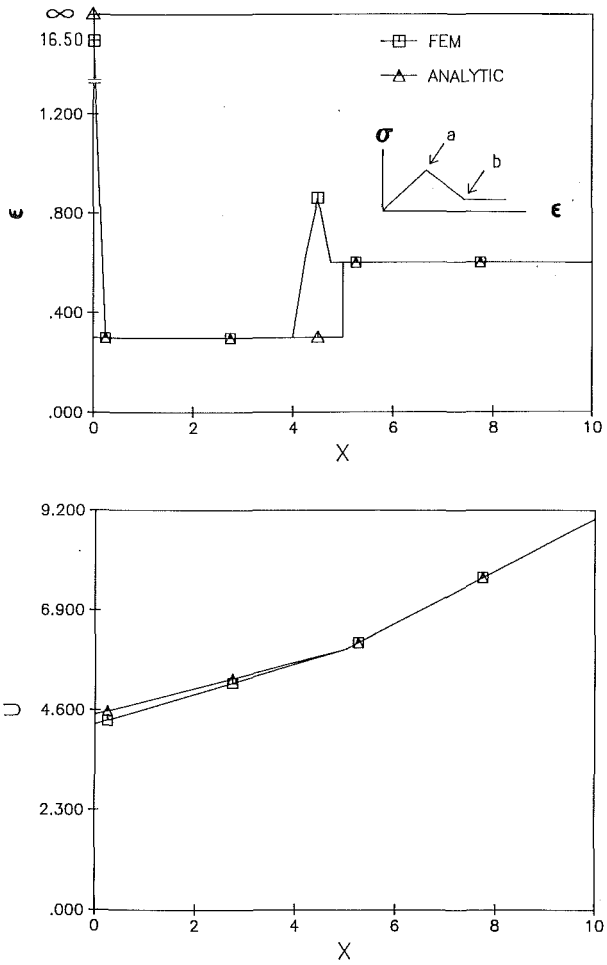


Fig. 4 Strain and displacement distributions at time $t = 15$ for problem 2 for the case when $\sigma_b = 0.3$

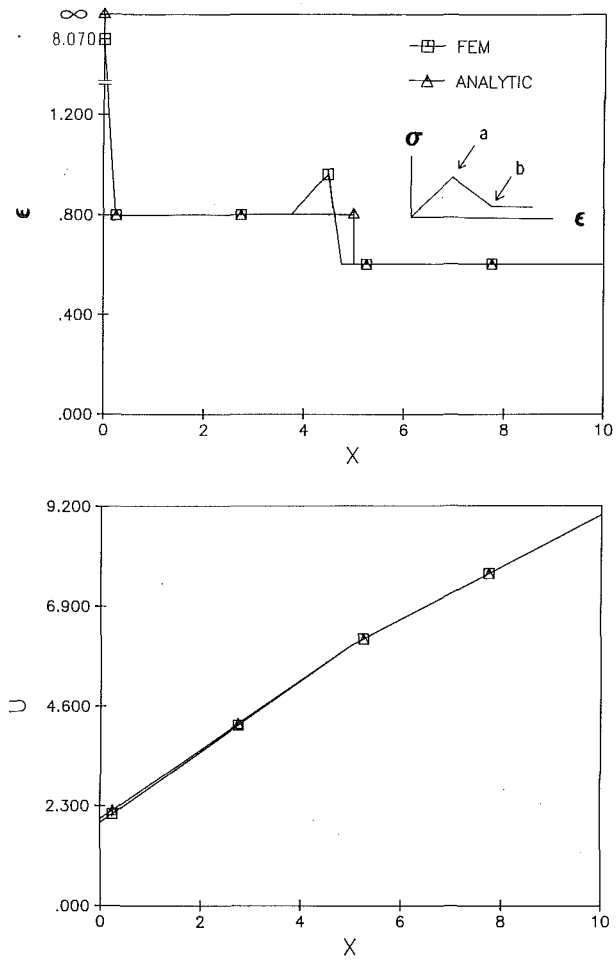


Fig. 5 Strain and strain displacement distributions at time $t = 15$ for problem 2 for the case when $\sigma_b = 0.8$

subsequent loading may require that $\sigma_3 = T(\epsilon_3) > \sigma_a$. Under certain conditions this requirement is not met.

Solution 2. The same BVP is solved with the stress-strain law shown in Fig. 2(*b*). The constitutive algorithm is as follows:

$$\text{initial conditions:} \quad \sigma = \epsilon = 0 \quad S = \sigma_a, \tag{28a}$$

$$\text{if } \sigma = S = \sigma_b \text{ and } \dot{\epsilon} > 0, \quad \dot{\sigma} = 0 \tag{28b}$$

$$\text{if } \sigma_b \le S \le \sigma_a, \ \sigma = \dot{S} \text{ and } \dot{\epsilon} > 0, \quad \dot{\sigma} = E^- \dot{\epsilon} \tag{29a}$$

$$\text{otherwise} \quad \dot{\sigma} = E^+ \dot{\epsilon} \tag{29b}$$

The difference algorithm of equation (6*d*-*e*) with $T(\epsilon)$ redefined as in Fig. 2(*b*) may also be used.

In constructing the solution for this material law, we note that the stress at the strain-softening point becomes σ_b after the jump in the strain, so the elastic solution in domains 1 and 2 becomes

$$\epsilon_2 = \epsilon^*, \quad \sigma_2 = \sigma_b \tag{30a}$$

$$u(x,t) = -v_o \left\langle \xi - \frac{2x}{c} \right\rangle + (c\epsilon^* - v_o) \langle \xi \rangle \tag{30b}$$

The size of the strain-softening domain is characterized by s with $s(t=0) = 0$, and from (30*b*), the elastic solution at $s = 0$ is given by

$$u(0,t) = (c\epsilon^* - 2v_o) \left\langle t - \frac{L}{c} \right\rangle \tag{31}$$

If strain softening has occurred, $c\epsilon^* - 2v_o < 0$, so since the stresses in both the elastic and strain-softening domain are σ_b, there is no mechanism for developing a wave to eliminate the displacement discontinuity. The only way to satisfy the boundary value problems (7) is to allow a discontinuity in the displacement at $x = 0$ and an associated infinite strain. Hence equation (30*b*) holds in the left-hand plane with $s = 0$ and the magnitude of the displacement discontinuity is $2\,c\epsilon^* - 4v_o$ and the strain field is given by

$$\epsilon(x,t) = \frac{v_o}{c} H\left(t - \frac{L+x}{c}\right) - \left(\frac{v_o}{c} - \epsilon_o\right) H\left(t - \frac{L-x}{c}\right)$$

$$+ (2c\epsilon^* - 4v_o)\delta(x) H\left(t - \frac{L}{c}\right) \quad \text{for } x \le 0. \tag{32}$$

where $\delta(\ )$ is the Dirac delta function.

The energy dissipation due to nonlinear material behavior in the region $-s \le x \le s$, where s tends to zero, results strictly from the Dirac delta term in equation (32) and is given by

$$W = 2\sigma_b (c\epsilon^* - 2v_o) \tag{33}$$

This agrees with the result of Bazant and Belytschko (1985) when $\sigma_b \to 0$, i.e., the dissipation vanishes when the stress goes to zero in the strain-softening domain. When $\sigma_b \neq 0$, a finite dissipation of energy can be achieved, but it is solely due to the plastic response and equivalent to that of an ideal plastic material with yield stress σ_b.

Figures 4 and 5 show the strain and displacement fields for problem 2 in the left-hand plane. Both the analytic solution and the finite element solution are shown. In these examples, $c = 1$, $Ev_o/c\sigma_a = 0.6$, and $\sigma_b = (0.3,\ 0.8)$ in Figs. 4 and 5, respectively. For the finite element solution, 40 elements were

Table 1 Parameters for problem 1 – Figs. 6–8

$\epsilon_a = 1.0$	$\sigma_a = 1.0$
$\epsilon_b = 1.2$	$\sigma_b = 0.2$
$v_o = 0.8$	$L = 50$

no. of elements for $(0 \le x \le L) = 80$
$c = 1.0$
Courant number ~ 0.7
$\dot{s}$ (for closed form solution) $= 0.704$

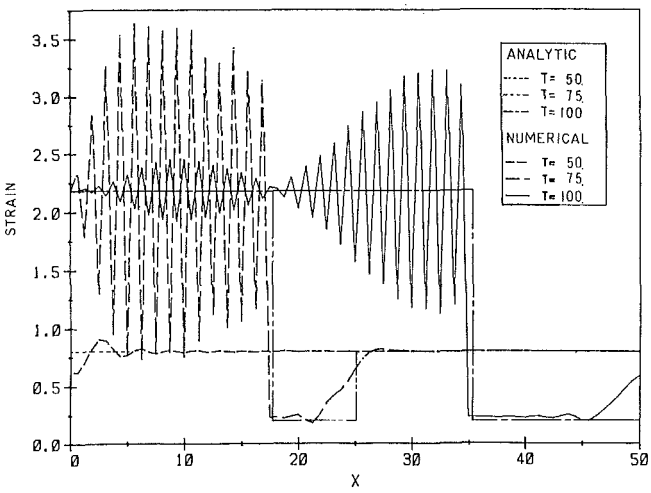


Fig. 6 Strain for example of problem 1

used on the left-hand plane and the Courant number was 0.6. Whereas the analytic solution predicts an infinite strain at $x = 0$, the strain in the finite element solution is finite but much larger than the surrounding strains. The analytic and finite element solutions otherwise agree quite well except at the wavefront generated by strain-softening (at $x = 5$ in the figures).

Figure 6 shows the strains in the right-hand plane for a finite element solution for problem 1; the corresponding analytic solution with $\epsilon_2 = \epsilon^*$ is also shown. The problem parameters are listed in Table 1. Several features are noteworthy: (*i*) the finite element solution exhibits the unloading wave (at which $\epsilon = 0.2$) which precedes the strain-softening wave $\dot{s}$; (*ii*) the finite element solution correctly captures the wave speed $\dot{s}$; (*iii*) the strains behind the wave $s(t)$ are extremely noisy, which probably reflects the difficulty the numerical solution has in reproducing the complex stress path associated with the wavefront (see Remark 4).

The noise significantly exceeds that found in finite element solutions of elastic wave propagation problems (see Holmes and Belytschko, 1976). Figure 7 shows the same solution with a five-point "averaging" digital filter described in Holmes and Belytschko (1976) applied to the strains and stresses. The filtering technique more clearly brings out the similarities of the finite element and analytic solutions. Figure 8 shows the displacements at 3 times, which again illustrates the presence of the unloading wave and the excellent agreement of the closed form and numerical solutions.

Capturing the unloading wave in a numerical solution does require some care. We used a time-step control so that during a time step no element can pass more than 10 percent beyond the point (ϵ_b, σ_b) in the stress-strain law. Attempts to obtain the same fidelity by reducing the Courant number (time step) to about 0.1 were unsuccessful because at such low Courant numbers the wavefronts are excessively dispersed.

A convergence study was made in the L_2-norm for this solution using meshes of 40 to 320 elements. The rates of convergence were quite sensitive to the time step and amount of

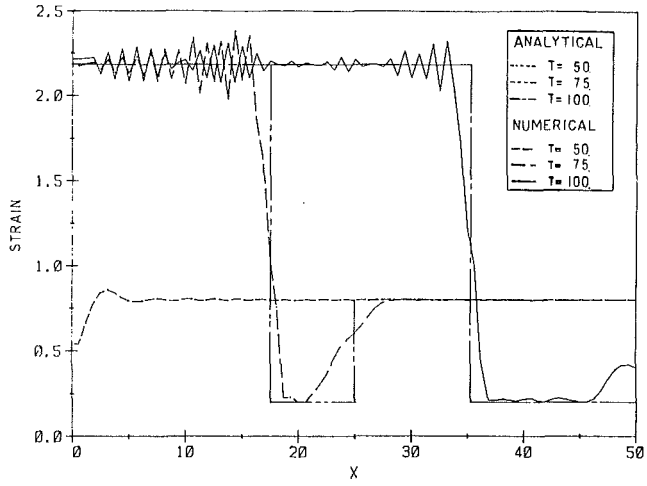


Fig. 7 Strain for example of problem 1 with five-point spatial averaging filter

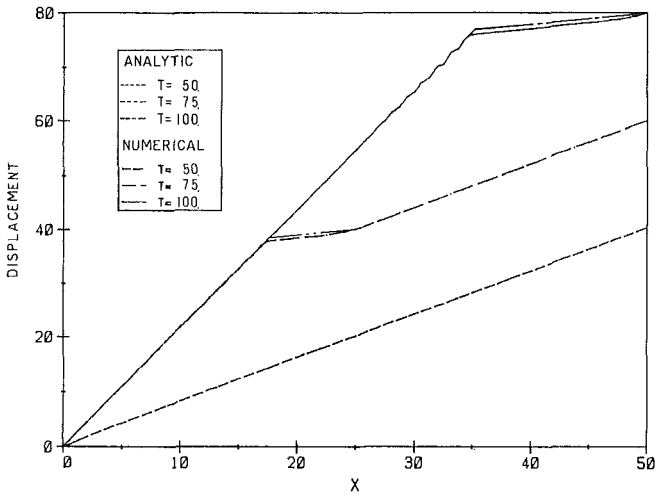


Fig. 8 Displacements for example of problem 1

artificial viscosity. The rates of convergence varied from $h^{0.6}$ to $h^{1.5}$, which is substantially below the h^2 rate in linear, static problems for this element.

Discussion and Conclusions

Closed-form transient solutions have been developed for rods with strain softening, a negative slope in the stress-strain curve. Two types of stress-strain curves were considered, one where the stress increases again and one where it remains constant after the strain softening. Finite element solutions were also obtained for representative problems. The following conclusions are drawn:

1 If the stress remains constant after the strain softening, a discontinuity appears in the displacement.

2 If the stress increases after strain softening, no discontinuity appears in the displacement, but the strain-softening point moves through the rod with jump discontinuities in the stresses and strains at a speed that is slower than the elastic wavespeed.

3 Finite element solutions reproduce the salient features of these solutions but exhibit excessive noise and slow rates of convergence.

4 When the stress remains constant, the displacement discontinuity is associated with a finite dissipation of energy; when the stress monotonically decreases to zero, the failure of the material associated with the discontinuity in displacements

requires no energy to be dissipated because it occurs on a set of measure zero.

Acknowledgment

The authors wish to express their gratitude for the support to the US Army Research Office under Contract DAAG29-84-K-0057 and the US Air Force Office of Scientific Research under Grant 83-0009.

References

Aifantis, E. C., and Serrin, J. B., 1983, "Equilibrium Solutions in the Micromechanical Theory of Fluid Microstructure," *Journal of Colloid and Interface Science*, Vol. 96, pp. 530–547.

Bazant, Z. P., 1976, "Instability, Ductility and Size Effects in Strain-Softening Concrete," *ASCE Journal of the Engineering Mechanics Division*, Vol. 102, EM2, pp. 331–344.

Bazant, Z. P., and Belytschko, T., 1985, "Wave Propagation in a Strain-Softening Bar; Exact Solution," *Journal of Engineering Mechanics*, Vol. 111, No. 3, pp. 381–389.

Belytschko, T., and Bazant, Z. P., 1984, "Strain-Softening Material and Finite Element Solutions," *Proc. ASME Symp. on Constitutive Equations; Macro and Computational Aspects*, William, K., ed., held at ASME Winter Annual Meeting, New Orleans, LA, pp. 253–272.

Hadamard, H., 1903, *Lecons sur la Propagation des Ondes*, Herman et Cie, Paris, France, Chapter VI.

Holmes, N., and Belytschko, T., 1976, "Postprocessing of Finite Element Transient Response Calculations by Digital Filters," *Computers and Structures*, Vol. 6, pp. 211–216.

James, R. D., 1980, "The Propagation of Phase Boundaries in Elastic Bars," *Arch. Rat. Mech. Anal.*, Vol. 73, pp. 125–158.

Sandler, I., and Wright, J., 1983, "Summary of Strain-Softening," *Theoretical Foundations for Large-Scale Computations of Nonlinear Material Behavior*, DARPA-NSF Workshop, Nemat-Nasser, S., ed., Northwestern University, Evanston, Ill., pp. 225–241.

Schreyer, H. L., and Chen, Z., 1984, "The Effect of Localization on the Softening Behavior of Structural Members," *Proc. ASME Symp. on Constitutive Equations; Macro and Computational Aspects*, presented at the Winter Annual Meeting of the ASME, New Orleans, pp. 193–203.

Willam, K. J., Bicanic, N., and Sture, S., 1984, "Constitutive and Computational Aspects of Strain-Softening and Localization in Solids," *Proc. ASME Symp. on Constitutive Equations; Macro and Computational Aspects*, presented at the Winter Annual Meeting of the ASME, New Orleans, LA, pp. 233–252.

Wu, F. H., and Freund, L. B., 1984, "Deformation Trapping Due to Thermoplastic Instability in One-Dimensional Wave Propagation," *Journal of Mechanics and Physics of Solids*, Vol. 32, No. 2, pp. 119–132.

Reexamination of Jumps Across Quasi-Statically Propagating Surfaces Under Generalized Plane Stress in Anisotropically Hardening Elastic-Plastic Solids

R. Narasimhan
Graduate Research Assistant.

A. J. Rosakis
Assistant Professor of Aeronautics and Applied Mechanics, Assoc. Mem. ASME

Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125

Strong discontinuities across quasi-statically propagating surfaces in anisotropic elastic-plastic solids under generalized plane stress are reexamined allowing for some generality in constitutive response and taking into account the phenomenon of necking. Jumps in stresses are ruled out on the basis of material stability postulates and a previous approach (by Pan, 1982) is discussed. It is noted that for elastic-perfectly plastic solids, sliding velocity discontinuities occur under restrictive and exceptional conditions (when both the surface and its normal are stress characteristics) for generalized plane stress as compared to plane strain. Necks may form along (stress) characteristic directions with the relative velocity vector orthogonal to the other family of characteristics.

1 Introduction

A variety of problems of physical interest involving the deformation of elastic-plastic solids may require the admission of discontinuities in the gradients of stresses and velocities (weak discontinuities) or in these quantities themselves (strong discontinuities). Such discontinuities may occur within regions that are currently deforming plastically or at elastic-plastic boundaries. These possibilities have received wide attention for rigid-perfectly plastic solids in plane strain (Hill, 1950) and in generalized plane stress (Hill, 1952) in the presence of either the isotropic Huber-von Mises or Tresca yield conditions in the plastic range. It is well known that for such solids, strong discontinuities in stress and velocity cannot be simultaneously present, and that velocity jumps occur across characteristic surfaces. It has been noted by Hill (1952) that when a rigid-plastic generalized plane-stress theory is employed in the study of the extension of thin plates, two types of strong discontinuities must be considered. These arise because of the neglect of elastic deformation and the averaged nature of generalized plane stress. A consideration of the second of these factors has led to the mathematical idealization of the experimentally observed phenomenon of localized necking in thin sheets (Nadai, 1950).

In a recent paper, Drugan and Rice (1984) investigated strong discontinuities across quasi-statically propagating surfaces in elastic-plastic solids under general three-dimensional conditions when all displacement components are assumed to be continuous. One important conclusion of their work is that all stress components are always continuous, a result that follows from certain material stability postulates.

Pan (1982) has also discussed quasi-statically moving strong discontinuities for elastic-perfectly plastic Huber-von Mises materials under generalized plane stress. He assumes that a strong discontinuity can be replaced by a transition layer of elastic material in which all stress components are assumed to vary continuously. He subsequently argues that all stress components are continuous across propagating surfaces, by using the specific nature of the Huber-von Mises locus and arriving at a contradiction.

In the present work, we reexamine quasi-static discontinuities for the more general case of an anisotropic hardening solid, using an integral form of the maximum plastic work inequality and the usual assumptions in the theory of generalized plane stress (Section 2). It is demonstrated in Section 5 that the use of the maximum plastic work inequality leads to full stress continuity for a broad class of solids, which includes some hardening materials and anisotropic behavior. Pan's assumptions and the limitations of his approach are discussed in Section 5. A complete analysis of all possible velocity jumps, including sliding discontinuity and localized necks, is carried out in Section 6 with some generality in constitutive behavior.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 10, 1986; final revision, October 28, 1986.

Paper No. 87-WA/APM-2.

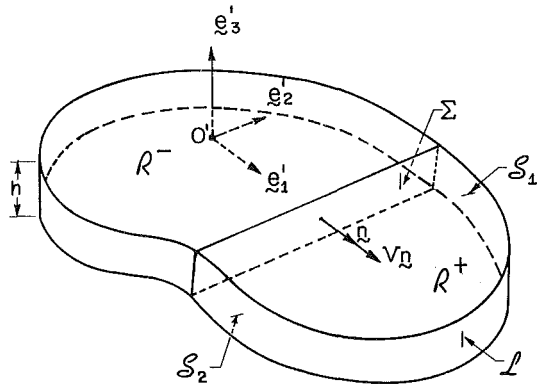


Fig. 1 Elastic-plastic body with discontinuity surface Σ

2 The Generalized Plane-Stress Problem

Consider an elastic-plastic body occupying an open cylindrical region R of height h (see Fig. 1). Let the boundary ∂R of the above region be composed of two traction-free planar surfaces S_1 and S_2 and a lateral surface L.

Consider further a fixed orthonormal coordinate system $\{0', \mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'\}$ such that $\mathbf{e}_3'$ is parallel to the generators of R.

Generalized plane stress conditions require that the height of the cylinder (also referred to later as the thickness of the cylindrical plate) be small as compared with any other dimension of the cylinder, and that the prescribed tractions $\mathbf{t}$ be such that:

$$\mathbf{t} = \mathbf{0} \text{ or } \sigma_{3i} = 0 \text{ on } S_1 \text{ and } S_2$$

and

$$t_3 = 0,\ t_\alpha = t_\alpha^*(x_1, x_2) \text{ on } L. \tag{2.1}$$

Here σ_{ij} are the components of the symmetric Cauchy stress tensor, Greek subscripts have the range 1, 2 while Latin subscripts take the values 1, 2, and 3. (This convention will be adopted through the following development.)

In what follows, field quantities such as σ, ϵ, $\mathbf{u}$, and v will represent thickness averages of the stress and strain tensors and the displacement and velocity vectors, respectively. It is also assumed that,

$$\sigma_{3i} \equiv 0 \text{ on } R. \tag{2.2}$$

The above assumptions result in solutions of the generalized plane stress problem which, in general, will not satisfy the exact three-dimensional field equations as discussed in detail by Timoshenko and Goodier (1970) and Hill (1950). This is because some of the compatibility equations are not generally satisfied, and errors are involved in using the averaged quantities in the constitutive law and the yield condition. However, if the plate thickness is sufficiently small, the generalized plane stress solution is expected to provide an accurate approximation.

Let Σ be a planar surface, parallel to the $x_2' - x_3'$ plane, dividing the region R in two open subregions R^+ and R^- such that

$$R = R^+ \cup R^- \cup \Sigma.$$

We will define the normal $\mathbf{n}(\mathring{\mathbf{x}})$ to Σ at a point $\mathring{\mathbf{x}} \in \Sigma$ as the outward normal of the closed subregion $\bar{R}^-$ ($\bar{R}^\pm \equiv R^\pm \cup \Sigma$) at the same point $\mathring{\mathbf{x}}$

In what follows, the surface Σ will be viewed as a potential surface of strong discontinuities (discontinuities in stresses and strains) and will be allowed to translate quasi-statically with a normal velocity $V\mathbf{n}$.

Since the approximate theory of generalized plane stress treats the thickness of the plate as vanishingly small, Hill (1952) points out that every quantity whose gradient is of order $(1/h)$ in a zone of breadth comparable to h *should be modelled as a discontinuity*. Thus, the experimentally observed formation of necks (Nadai, 1950) in thin plates subjected to tension (rapid variation of thickness of the plate in narrow zones) would be modelled as discontinuities in the out-of-plane displacement component u_3.

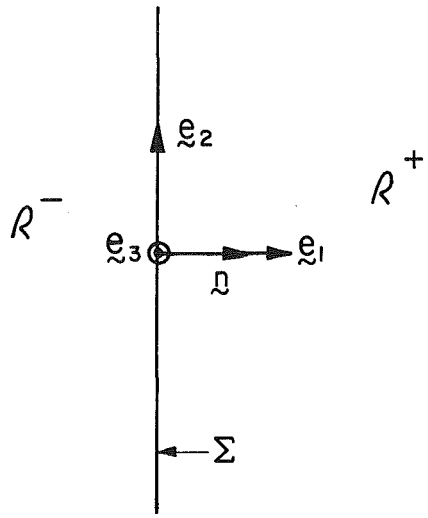


Fig. 2 Local coordinate system translating with the surface

The jump in a field quantity $\mathbf{g}(\mathbf{x})$, across the surface Σ, will be denoted by:

$$\left.\begin{aligned} &[\mathbf{g}] \equiv \mathbf{g}^+(\mathring{\mathbf{x}}) - \mathbf{g}^-(\mathring{\mathbf{x}}) \text{ where} \\ &g^\pm(\mathring{\mathbf{x}}) = \lim_{\epsilon \to 0} g(\mathring{\mathbf{x}} \pm \epsilon \mathbf{n}(\mathring{\mathbf{x}}))\ \ \mathring{\mathbf{x}} \in \Sigma \text{ and } \epsilon > 0 \end{aligned}\right\} \tag{2.3}$$

3 Smoothness Considerations

All field quantities will be referred to, with respect to an orthonormal frame $\{0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ translating with the surface Σ and such that $0 \in \Sigma$, $\mathbf{e}_3 = \mathbf{e}_3'$ and $\mathbf{e}_1 = \mathbf{n}$; see Fig. 2.

Inplane displacement components u_α are required to have the following smoothness properties:

$$\begin{aligned} &u_\alpha \in C(R) \text{ and} \\ &u_\alpha \in C^1(R - \Sigma), \end{aligned} \tag{3.1}$$

with the understanding that $\partial u_\alpha / \partial x_\beta$ need not be continuous across Σ. Then, according to the Hadamard compatibility relations (Hill, 1961) for jumps in the derivatives of a continuous function,

$$\left[\frac{\partial u_\alpha}{\partial x_\beta}\right] = \lambda_\alpha n_\beta \text{ on } \Sigma. \tag{3.2}$$

where λ_α are arbitrary functions of position on Σ. The out-of-plane displacement component u_3 will in general be allowed to suffer a jump across Σ, as discussed in Section 2. Thus:

$$u_3 \in C^1(R^+) \cap C^1(R^-), \tag{3.3}$$

with the understanding that on Σ, u_3 and its gradient need not be defined. On the other hand, $[u_3]$, the jump in the limiting value of u_3 from R^- to R^+, will be assumed to be a continuous and continuously differentiable function of position on Σ.

It is now possible to extend the Hadamard compatibility relations (3.2) for the treatment of jumps in the derivatives of discontinuous functions. This extension was first discussed by Thomas (1957). The following simpler version was later provided by Hill (1961),

$$\left[\frac{\partial u_3}{\partial x_i}\right] = \lambda_3 n_i + \frac{\partial \phi}{\partial x_i} \text{ on } \Sigma. \tag{3.4}$$

where λ_3 is an arbitrary function of position on Σ and ϕ is an arbitrary continuous function, together with its gradient on Σ and in one neighborhood, say R^-, with the additional restriction that

$$\phi = [u_3] \text{ on } \Sigma.$$

One choice of ϕ in $\bar{R}^+$ would be to consider ϕ continued analytically along the normals. Any other choice would merely change λ_3, which is given by

$$\lambda_3 = [\nabla u_3 \cdot \mathbf{n}] - \nabla\phi \cdot n.$$

Relations (3.2) and (3.4) allow definition of jumps in the strains across Σ, consistent with the assumptions of the approximate theory of generalized plane stress.

Within the contexts of a small strain formulation,

$$\epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}), \tag{3.5}$$

and the jumps in the inplane strain component $\epsilon_{\alpha\beta}$ can be expressed by equation (3.2) as:

$$[\epsilon_{\alpha\beta}] = \frac{1}{2}(\lambda_\alpha n_\beta + \lambda_\beta n_\alpha) \text{ on } \Sigma. \tag{3.6}$$

On the other hand, the jump in the out-of-plane strain component ϵ_{33} can be expressed by equations (3.2) and (3.4) as:

$$[\epsilon_{33}] = \lambda_3 n_3 + \frac{\partial\phi}{\partial x_3} \text{ on } \Sigma. \tag{3.7}$$

where $\phi \in C^1(\bar{R}^+)$ and $\phi = [u_3]$ on Σ.

4 Material Idealization

Within the context of the small-strain flow theory of plasticity, the total strain rate tensor can be decomposed into elastic and plastic parts:

$$\dot{\epsilon} = \dot{\epsilon}^e + \dot{\epsilon}^P \text{ on } R, \tag{4.1}$$

where the dot denotes differentiation with respect to time. The elastic strain rate tensor $\dot{\epsilon}^e$ is related to the stress rate tensor $\dot{\sigma}$ through a constant, positive definite four-tensor $\mathbf{H}$ (the inverse of the elasticity tensor $\mathbf{C}$). $\mathbf{H}$ is assumed to possess the usual major and minor symmetries. For an anisotropic elastic-plastic solid, $\dot{\epsilon}^e$ is given by:

$$\dot{\epsilon}^e = \mathbf{H}\dot{\sigma} \text{ on } R. \tag{4.2}$$

Attention will be focused on the class of materials obeying Drucker's stability postulate. A particular form of this postulate known as the maximum plastic work inequality can be expressed as:

$$(\sigma - \sigma^*) \cdot \dot{\epsilon}^P \geq 0. \tag{4.3}$$

$\forall f(\sigma, \bar{\epsilon}^P) = 0$, and $f(\sigma^*, \bar{\epsilon}^P) \leq 0$, where $f(\sigma, \bar{\epsilon}^P)$ is the yield function. An important implication of the above postulate is the normality of the plastic strain rate $\dot{\epsilon}^P$ to the yield surface leading to a flow rule of the form,

$$\dot{\epsilon}^P = \dot{\lambda}\mathbf{P}, \tag{4.4}$$

where $\dot{\lambda} \geq 0$ and $\mathbf{P} = \nabla_\sigma f$. $\dot{\lambda}$ and $\mathbf{P}$ are scalar valued and symmetric tensor valued functions of σ, respectively. In the following section, an integral form of (4.3) will be used in conjunction with equations (4.1) and (4.2), as well as the compatibility conditions for the jumps in total strains (3.6), (3.7) to define the jumps in the stresses and the plastic strains produced during the passage of a discontinuity Σ through a material point.

5 Stress Continuity Across the Propagating Surface

In this section it will be demonstrated that all stress components are continuous across the surface Σ, propagating quasi-statically through the thin plate. It will be shown that this is true even if the out-of-plane displacement u_3 suffers a discontinuity across Σ. The following proof is based on the maximum plastic work inequality and the positive definiteness of $\mathbf{H}$. It is an adaptation for plane stress of the proof given by Drugan and Rice (1984) for the general three-dimensional case. In the present analysis, only the in-plane displacement components u_α are assumed continuous, and the proof is adapted to suit the assumptions of the theory of generalized plane stress. Also, unlike the discussion by Pan (1982) and consistent with the assumptions of generalized plane stress (Hill, 1952), our discussion treats necks as jumps and *not* as narrow transition layers.

If inertia terms are neglected, the balance of linear momentum requires that across the quasi-statically moving surface Σ the traction be continuous. Thus

$$[t_\alpha] = [\sigma_{\alpha\beta} n_\beta] = 0 \text{ on } \Sigma.$$

With respect to the local orthonormal coordinate frame $\{0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ moving with Σ, $n_i = \delta_{1i}$ and the above conditions become:

$$[\sigma_{1\alpha}] = 0 \text{ on } \Sigma. \tag{5.1}$$

Equations (2.2) and (5.1) imply that the only stress component that can suffer a nontrivial jump is σ_{22}. The plastic work W^P accumulated discontinuously at a material point due to the passage of the surface Σ is given by:

$$W^P = \int_{\epsilon^{P+}}^{\epsilon^{P-}} \sigma \cdot d\epsilon^P. \tag{5.2}$$

It should be observed here that some error is involved in using the averaged stress and strain quantities of generalized plane stress in the above integral. The above integral is evaluated according to the assumptions of Section 2. On applying equation (2.2), we find that the plastic work accumulation in equation (5.2) reduces to

$$W^P = \int_{\epsilon^{P+}_{\alpha\beta}}^{\epsilon^{P-}_{\alpha\beta}} \sigma_{\alpha\beta} d\epsilon^P_{\alpha\beta}. \tag{5.3}$$

Using equation (5.1), the above becomes:

$$W^P = -\sigma_{11}[\epsilon^P_{11}] - 2\sigma_{12}[\epsilon^P_{12}] + \int_{\epsilon^{P+}_{22}}^{\epsilon^{P-}_{22}} \sigma_{22} d\epsilon^P_{22}. \tag{5.4}$$

Also, by using the fact that $n_\beta = \delta_{1\beta}$, equation (3.6) implies that:

$$[\epsilon_{22}] = 0 \text{ on } \Sigma \text{ or } [\epsilon^e_{22}] = -[\epsilon^P_{22}] \text{ on } \Sigma. \tag{5.5}$$

By setting $d\epsilon^P = d\epsilon - d\epsilon^e$, using the continuity of ϵ_{22} across Σ equation (5.5) and integrating by parts, equation (5.4) becomes

$$W^P = -\sigma_{11}[\epsilon^P_{11}] - 2\sigma_{12}[\epsilon^P_{12}] - \int_{\epsilon^{e+}_{22}}^{\epsilon^{e-}_{22}} \sigma_{22} d\epsilon^e_{22}. \tag{5.6}$$

The integral in equation (5.6) can now be evaluated by using equations (2.2), (5.1), and the constitutive law, to give:

$$\int_{\epsilon^{e+}_{22}}^{\epsilon^{e-}_{22}} \sigma_{22} d\epsilon^e_{22} = -\frac{1}{2} H_{2222}(\sigma^+_{22} + \sigma^-_{22})[\sigma_{22}]. \tag{5.7}$$

In addition, from equations (2.2), (5.1), and (5.5),

$$[\epsilon^P_{22}] = -[\epsilon^e_{22}] = -H_{2222}[\sigma_{22}] \text{ on } \Sigma. \tag{5.8}$$

Thus, equations (5.6) and (5.7) give:

$$W^P = -\sigma_{11}[\epsilon^P_{11}] - 2\sigma_{12}[\epsilon^P_{12}] - \frac{1}{2}(\sigma^+_{22} + \sigma^-_{22})[\epsilon^P_{22}]$$

or,

$$W^P = -\frac{1}{2}(\sigma^+_{ij} + \sigma^-_{ij})[\epsilon^P_{ij}]. \tag{5.9}$$

It should be observed that the restrictions imposed on the path in stress space in the evaluation of the integral in equation

(5.6) are the plane stress conditions and the continuity of tractions across Σ. This effectively implies a straight line path in stress space from σ_{22}^{+} to σ_{22}^{-}.

The integral form of the plastic work inequality (4.3) can now be used by setting $\boldsymbol{\sigma}^{*} = \boldsymbol{\sigma}^{+}$ where $f(\boldsymbol{\sigma}^{+}, \bar{\epsilon}^{P}) \leq 0$. Thus, $\boldsymbol{\sigma}^{+}$ is constrained to remain always at or inside the yield surface during passage of Σ. Thus, by our using equations (2.2), (5.1), and (5.9):

$$\int_{\epsilon_{ij}^{P+}}^{\epsilon_{ij}^{P-}} (\sigma_{ij} - \sigma_{ij}^{+}) d\epsilon_{ij}^{P} = -\frac{1}{2}(\sigma_{22}^{+} + \sigma_{22}^{-})[\epsilon_{22}^{P}] + \sigma_{22}^{+}[\epsilon_{22}^{P}] \geq 0$$

which, by equation (5.8), gives

$$\frac{1}{2}[\sigma_{22}][\epsilon_{22}^{e}] \leq 0 \text{ or } \frac{1}{2}([\sigma_{22}])^{2} H_{2222} \leq 0. \tag{5.10}$$

(5.10) now requires that $[\sigma_{22}] = 0$, since $H_{2222} > 0$.

Remarks. The following remarks are relevant:

1. Under generalized plane stress conditions, all stress components are continuous across the slowly propagating surface Σ, even if the out-of-plane displacement u_3 suffers a discontinuity.
2. The present discussion applies to general anisotropic elastic-plastic hardening solids obeying a flow rule of the associated type. The proof of full stress continuity is based on an integral form of the maximum plastic work inequality and the positive definiteness of the elastic potential.
3. An earlier discussion by Pan (1982) is limited to elastic-ideally plastic solids of a Huber-von Mises type under generalized plane stress conditions. His argument, which does not make use of the maximum plastic work inequality, follows from Hill's statement (Hill, 1950) that the stress state from σ_{22}^{+} to σ_{22}^{-} can be bridged only by a succession of elastic states. This assumes a smooth variation of stresses in a "transition layer." Such an assumption is questionable for generalized plane stress since, as pointed out in Section 2, any field quantity whose gradient is $O(1/h)$ in a zone of breadth comparable to h should be modelled as a discontinuity. Even if this assumption is accepted, Pan's argument clearly does not apply to arbitrary yield surfaces or general hardening solids. For instance, in elastic-perfectly plastic solids characterized by a Tresca yield condition when the neck (discontinuity in u_3) coincides with a principal stress direction and $\sigma_{11} = \pm \sigma_o$, the stress component σ_{22} can have *any* value between 0 and $\pm \sigma_o$ and still lie on the yield surface (Hill, 1950). Hence, $\dot{\lambda}$ in equation (4.4) is not necessarily zero in the transition from $\boldsymbol{\sigma}^{+}$ to $\boldsymbol{\sigma}^{-}$ (Hill, 1952), and the argument fails. Also, for any type of hardening solid, the consistency condition requires the stress state to lie on the yield surface during the process from $\boldsymbol{\sigma}^{+}$ to $\boldsymbol{\sigma}^{-}$ and no elastic unloading is possible.

6 Discontinuities in Strains and Velocities

In this section, the earlier result pertaining to continuity of stresses across Σ will be used to provide restrictions on the nature of admissible jumps in strains and material particle velocities across Σ for a general anisotropic hardening solid. Attention will then be turned to plastically incompressible, generally anisotropic, elastic-perfectly plastic solids with smooth but otherwise arbitrary yield surfaces. Specialized results will be given for Huber-von Mises solids at the end of the discussion.

General Considerations. The jumps in the in-plane velocity component v_α are given (Hill, 1961; Drugan and Rice, 1984) by:

$$[v_\alpha] = -V\left[\frac{\partial u_\alpha}{\partial x_1}\right] \text{ on } \Sigma, \tag{6.1}$$

where $V\mathbf{n}$ is the normal velocity of Σ. Making use of equations (3.2) and (3.6) the velocity jumps may be expressed as:

$$\begin{aligned} [v_1] &= -V[\epsilon_{11}] \\ [v_2] &= -2V[\epsilon_{12}] \end{aligned} \text{ on } \Sigma. \tag{6.2}$$

Full stress continuity and equation (4.2) require the elastic part of the strains to be continuous across Σ.

$$[\epsilon_{ij}^{e}] = 0 \text{ on } \Sigma. \tag{6.3}$$

The above, and equation (5.8), therefore imply

$$[\epsilon_{22}^{P}] = 0 \text{ on } \Sigma. \tag{6.4}$$

As a result, the expression for the positive plastic work accumulation in equation (5.9) becomes

$$W^{P} = -\sigma_{11}[\epsilon_{11}^{P}] - 2\sigma_{12}[\epsilon_{12}^{P}] \geq 0, \tag{6.5}$$

and the jumps in the velocity components v_1 and v_2 are given by:

$$\begin{aligned} [v_1] &= -V[\epsilon_{11}^{P}] \\ [v_2] &= -2V[\epsilon_{12}^{P}] \end{aligned} \text{ on } \Sigma. \tag{6.6}$$

The plastic work W^{P} can now be expressed in terms of velocity jumps as follows:

$$W^{P} = \frac{1}{V}(\sigma_{11}[v_1] + \sigma_{12}[v_2]) \geq 0. \tag{6.7}$$

No specific restrictions on the constitutive model other than the general assumptions made in Section 4 have been imposed in the derivation of equations (6.1)–(6.7).

For the specific class of plastically incompressible solids:

$$[\epsilon_{33}^{P}] = -[\epsilon_{11}^{P}] - [\epsilon_{22}^{P}] \text{ on } \Sigma. \tag{6.8}$$

which, by use of equation (6.4) simplifies to:

$$[\epsilon_{33}^{P}] = -[\epsilon_{11}^{P}] \text{ on } \Sigma. \tag{6.9}$$

Equation (6.9) serves to determine the jump in the out-of-plane plastic strain component ϵ_{33}^{P} in terms of the jump in the inplane plastic strain component ϵ_{11}^{P} for plastically incompressible solids.

If the displacement component u_3 happens to be continuous across Σ as in (Drugan and Rice, 1984), then ϵ_{33} and hence ϵ_{33}^{P} would also be continuous. Equations (6.8) and (6.6) will then imply that ϵ_{11}^{P} and v_i should also be continuous across Σ. Thus, it follows that for a plastically incompressible solid, when the surface Σ does not coincide with a neck (jump in u_3), only a sliding velocity discontinuity (jump in v_2) is permissible.

Elastic-Perfectly Plastic Solid. For such solids, the yield surface is represented by

$$f(\boldsymbol{\sigma}) = 0 \text{ on } R, \tag{6.10}$$

where $f(\boldsymbol{\sigma})$ depends symmetrically on $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}^{T}$. It will also be assumed here that the yield surface is smooth (has a continuous normal).

Under such circumstances the flow rule takes the following form:

$$\dot{\boldsymbol{\epsilon}}^{P} = \dot{\lambda}\mathbf{P} \text{ on } R, \tag{6.11}$$

where $\dot{\lambda} \geq 0$ is an undetermined scalar function of position, and

$$\mathbf{P}(\boldsymbol{\sigma}) \equiv \nabla_{\boldsymbol{\sigma}} f(\boldsymbol{\sigma}) \text{ on } R \tag{6.12}$$

is a symmetric tensor-valued function of $\boldsymbol{\sigma}$. Under conditions of generalized plane stress, equations (6.10) and (6.12) should be used in conjunction with the constraint (2.2). Inside regions that are currently deforming plastically, it can be shown from the two inplane equilibrium equations, the yield condition,

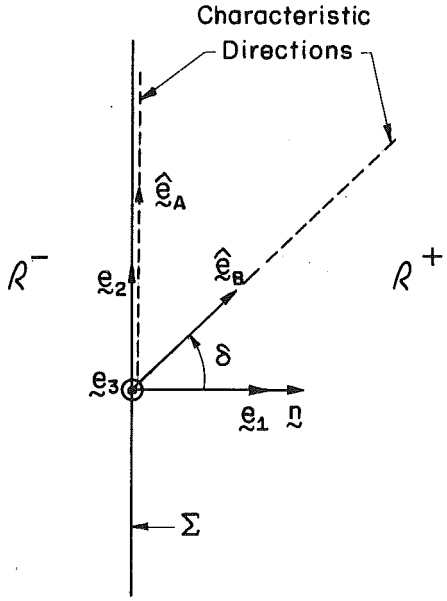


Fig. 3 Moving discontinuity surface and characteristic directions of the plastically deforming side

and the plane stress assumption (2.2) that along stress characteristic directions the direct components of $P_{\alpha\beta}$ should vanish (Hill, 1950).

It is also clear that, P_{ij} should be continuous across Σ from assumed smoothness of the yield surface and the requirement of full stress continuity. Then, from equation (6.11) the jumps in the plastic strain component ϵ_{ij}^P becomes

$$[\epsilon_{ij}^P] = -\eta P_{ij} \text{ on } \Sigma. \tag{6.13}$$

where $\eta = \int_{\lambda^+}^{\lambda^-} d\lambda \geq 0$ is an undetermined scalar function of position on Σ.

Since $[\epsilon_{22}^P] = 0$ across Σ, equation (6.13) implies that either $\eta = 0$ or $P_{22} = 0$ or both. If $\eta = 0$, equation (6.13) requires all strain components to be continuous. Thus, the *necessary* condition for nontrivial jumps in strains to exist across Σ is that P_{22} should vanish on Σ. In other words, Σ *should coincide with a stress characteristic direction of its plastic side.*

This condition is less restrictive than the necessary condition for nontrivial jumps in the plastic strain components derived by Drugan and Rice (1984) when *all* displacements were continuous across Σ. The corresponding necessary condition derived by them states that $P_{22} = P_{33} = P_{23} = 0$ on Σ.

From the above, the following important observation can be made. Consider at least one side of Σ (which coincides with a neck, say R^+) to be currently deforming plastically. If, in addition, Σ coincides with one of the stress characteristic directions, say direction A (see Fig. 3), then the velocity component along the other characteristic direction B, is continuous across Σ. Thus,

$$[V_B] = 0 \text{ on } \Sigma. \tag{6.14}$$

The above result follows by first observing that since Σ coincides with a stress characteristic direction, P_{22} vanishes on Σ. Also, if the other characteristic direction makes an angle δ ($\delta \neq \pm \pi/2$) with the x_1 axis, then by the fact that $P_{BB} = 0$ and the transformation relation, we have

$$\tan \delta = -\frac{P_{11}}{2P_{12}} \text{ for } P_{12} \neq 0. \tag{6.15}$$

In addition, combining equations (6.6) and (6.13) and noting that $P_{12} \neq 0$, we see that the following is true:

$$[v_1] = \frac{P_{11}}{2P_{12}} [v_2] \text{ on } \Sigma. \tag{6.16}$$

The velocity jump $[v_B]$ along the other characteristic direction will be given by

$$[v_B] = \cos \delta([v_1] + [v_2] \tan \delta), \ \delta \neq \pm \frac{\pi}{2},$$

which vanishes by use of equation (6.15) and (6.16). This general result was also noted by Pan (1982) for the special case of an isotropic Huber-von Mises solid and it also holds for stationary necks in a rigid-plastic solid (Hill, 1952).

If in addition $P_{12} = 0$, both the stress characteristics merge along Σ ($\delta = \pm \pi/2$) and as a result Σ becomes a "parabolic line." Equations (6.6) and (6.13) then imply that if $P_{12} = 0$,

$$[v_2] = 0 \text{ on } \Sigma. \tag{6.17}$$

Thus, when Σ coincides with a "parabolic line," the tangential velocity is continuous and *only* the normal velocity has a jump.

When Σ coincides with a neck and the two characteristic directions do not merge along Σ (see Fig. 3), then the accumulation of plastic work (6.7) due to the passage of Σ becomes:

$$W^P = \frac{1}{V}\left(\frac{\sigma_{11}P_{11} + 2\sigma_{12}P_{12}}{2P_{12}}\right)[v_2] \geq 0. \tag{6.18}$$

Also, the fact that $\sigma_{ij}\dot{\epsilon}_{ij}^P \geq 0$ implies that

$$\sigma_{ij}P_{ij} \geq 0. \tag{6.19}$$

By equation (2.2) and $P_{22} = 0$ along Σ, (6.19) becomes:

$$\sigma_{11}P_{11} + 2\sigma_{12}P_{12} \geq 0. \tag{6.20}$$

Inequalities (6.18) and (6.20) result in

$$\frac{1}{V}\frac{[v_2]}{2P_{12}} \geq 0, \ P_{12} \neq 0. \tag{6.21}$$

When the two characteristics merge along Σ ($\delta = \pm \pi/2$ and $P_{22} = P_{12} = 0$), it follows from equations (6.7) and (6.17) that

$$\frac{\sigma_{11}}{V}[v_1] \geq 0. \tag{6.22}$$

Isotropic Huber-von Mises Solids. The above results can now be specialized for an isotropic elastic-perfectly plastic solid that obeys the Huber-von Mises yield condition. For such a solid, the yield condition states

$$f(\sigma) = \frac{1}{2}\mathbf{S}\cdot\mathbf{S} - \tau_0^2 = 0 \text{ on } R, \tag{6.23}$$

where $\mathbf{S} \equiv \sigma - 1/3\ tr\ \sigma \mathbf{1}$ is the deviatoric stress tensor and τ_o is the yield stress in pure shear. For such a solid,

$$\mathbf{P}(\sigma) = \nabla_\sigma f(\sigma) = \mathbf{S} \text{ on } R. \tag{6.24}$$

All the results and corresponding remarks from equations (6.3)–(6.22) hold for this solid with $\mathbf{P}$ replaced by $\mathbf{S}$. In particular, equation (6.16) takes the form (Pan, 1982):

$$[v_1] = \frac{S_{11}}{2S_{12}}[v_2] \text{ if } S_{12} \neq 0, \tag{6.25}$$

and equation (6.18) reduces to

$$W^P = \frac{1}{V}\left(\frac{\tau_o^2}{\sigma_{12}}\right)[v_2] > 0. \tag{6.26}$$

Summary of Results. The results of Section 6 can now be summarized as follows:

(*a*) For a general anisotropic hardening solid that is also plastically incompressible, the following is true: When the propagating surface Σ does *not* coincide with a neck (full displacement continuity), only a jump in the tangential velocity component (sliding discontinuity) is admissible.

(*b*) If, however, the solid is perfectly plastic, Σ coincides with one characteristic direction ($P_{22} = 0$). In addi-

tion, full displacement continuity together with plastic incompressibility also give $P_{33} = P_{11} = 0$. This states that the direction normal to Σ is *also* a characteristic direction. Unlike plane strain, this occurs under plane stress conditions only under exceptional circumstances (Hill, 1950). In particular, for Huber-von Mises solids this is true when the surface coincides with a plane of maximum shear stress, and the latter is equal in magnitude to the yield stress in pure shear.

(*c*) For a general anisotropic elastic-perfectly plastic solid, when a surface *coincides* with a neck (discontinuity in u_3), both tangential and normal velocities have jumps. This requires that the neck should lie along one characteristic direction. Then the component of the velocity along the other characteristic direction (not generally perpendicular to Σ) is continuous (see equation (6.14)). Thus, necks cannot form if the plastically deforming side of the surface is in an elliptic state of stress.

(*d*) For an elastic perfectly plastic solid, if in addition to (*c*), $P_{12} = 0$, both the characteristics merge along the neck, and this results in a parabolic stress state. Then, the tangential velocity is continuous and *only the normal velocity has a jump*. For the special case of a Huber-von Mises solid, $P_{12} = S_{12} = 0$, and the characteristic surface coincides with a principal stress direction.

7 Remarks and Applications

The jump conditions discussed here have some relevance to the stress and strain fields near the tip of a quasi-statically growing crack in an elastic-plastic solid under generalized plane stress conditions. For instance, in the elastic-perfectly plastic Huber-von Mises material (Rice, 1982) a "constant stress" (asymptotic) plastic sector cannot occur direclty behind a "centered fan" plastic sector because the condition for positive plastic work accumulation (6.21) will be violated at the interface. This renders the asymptotic solution for the plane stress stationary crack by Hutchinson (1968) unacceptable when the crack begins to grow. From the preliminary asymptotic analysis by Rice (1982), it then follows that only an "elastic unloading" sector can occur behind the centered fan. Hutchinson's stationary crack solution also has a jump in the inplane stress component between two constant stress sectors. This is also inadmissible when the crack begins to propagate.

No solution for this problem, which satisfies all the conditions set forth in the present work, has yet been constructed. An open question that arises, for which detailed experimental and numerical studies may provide an answer, is whether necking occurs near the growing crack tip. Otherwise, except in special circumstances (e.g., a fan angle of 90 deg), no strong discontinuities near the growing crack tip can be admitted. In view of the fact that the (fully yielded) stationary crack tip solution (Hutchinson, 1968) has a strong discontinuity, one wonders whether the condition of full continuity in both stress and velocity near the propagating crack tip may be too restrictive to satisfy.

Acknowledgments

The authors would like to express their gratitude to Professor J. K. Knowles for his valuable advice and encouragement. The support of ONR Contract No. N00014-85-K-0596 is also gratefully acknowledged.

References

Drugan, W. J., and Rice, J. R., 1984, "Restrictions on Quasi-Statically Moving Surfaces of Strong Discontinuity in Elastic-Plastic Solids," *Mechanics of Material Behaviour* Dvorak, G. J., and Shield, R. T., eds., Elsevier, Amsterdam, pp. 59–73.

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

Hill, R., 1952, "On Discontinuous Plastic States with Special Reference to Localized Necking in Thin Sheets," *Journal of Mechanics and Physics of Solids*, Vol. 1, pp. 19–30.

Hill, R., 1961, "Discontinuity Relations in Mechanics of Solids," *Progress in Solid Mechanics*, Sneddon, I. N., and Hill, R., eds., Vol. 2, Ch. 6, pp. 247–276.

Hutchinson, J. W., 1968, "Plastic Stress and Strain Fields at a Crack Tip," *Journal of Mechanics and Physics of Solids*, Vol. 16, pp. 337–347.

Nadai, A., 1950, *Theory of Flow and Fracture in Solids*, McGraw-Hill, New York, Vol. 1, pp. 319–320.

Pan, H., 1982, "Some Discussion on Moving Strong Discontinuity under Plane Stress," *Mechanics of Materials*, Vol. 1, pp. 325–329.

Rice, J. R., 1982, "Elastic-Plastic Crack Growth," *Mechanics of Solids*, Hopkins, H. G., and Sewell, M. J., eds., Pergamon Press, Oxford, pp. 539–562.

Thomas, T. Y., 1957, "Extended Compatibility Relations for Study of Surfaces of Discontinuity in Continuum Mechanics," *Journal of Mathematics and Mechanics*, Vol. 6, pp. 311–322.

Timoshenko, S. P., and Goodier, G. N., 1970, *Theory of Elasticity*, McGraw-Hill, New York, p. 31, p. 274.

A Continuum Model for Void Nucleation by Inclusion Debonding

A. Needleman
Division of Engineering,
Brown University,
Providence, RI 02912

A cohesive zone model, taking full account of finite geometry changes, is used to provide a unified framework for describing the process of void nucleation from initial debonding through complete decohesion. A boundary value problem simulating a periodic array of rigid spherical inclusions in an isotropically hardening elastic-viscoplastic matrix is analyzed. Dimensional considerations introduce a characteristic length into the formulation and, depending on the ratio of this characteristic length to the inclusion radius, decohesion occurs either in a "ductile" or "brittle" manner. The effect of the triaxiality of the imposed stress state on nucleation is studied and the numerical results are related to the description of void nucleation within a phenomenological constitutive framework for progressively cavitating solids.

1 Introduction

The nucleation of voids from inclusions and second phase particles plays a key role in limiting the ductility and toughness of plastically deforming solids, including structural metals and composites. The voids initiate either by inclusion cracking or by decohesion of the interface, but here attention is confined to consideration of void nucleation by interfacial decohesion.

Theoretical descriptions of void nucleation from second phase particles have been developed based on both continuum and dislocation concepts, e.g., Brown and Stobbs (1971), Argon et al. (1975), Chang and Asaro (1978), Goods and Brown (1979), and Fisher and Gurland (1981). These models have focussed on critical conditions for separation and have not explicitly treated propagation of the debonded zone along the interface. Interface debonding problems have been treated within the context of continuum linear elasticity theory; for example, the problem of separation of a circular cylindrical inclusion from a matrix has been solved for an interface that supports neither shearing nor tensile normal tractions (Keer et al., 1973). The growth of a void at a rigid inclusion has been analyzed by Taya and Patterson (1982), for a nonlinear viscous solid subject to overall uniaxial straining and with the strength of the interface neglected.

The model introduced in this investigation is aimed at describing the evolution from initial debonding through complete separation and subsequent void growth within a unified framework. The formulation is a purely continuum one using a cohesive zone (Barenblatt, 1962; Dugdale, 1960) type model for the interface but with full account taken of finite geometry changes. Constitutive relations are specified independently for the matrix, the inclusion, and the interface. The constitutive equation for the interface is such that, with increasing interfacial separation, the traction across the interface reaches a maximum, decreases, and eventually vanishes so that complete decohesion occurs. Since the mechanical response of the interface is specified in terms of both a critical interfacial strength and the work of separation per unit area, dimensional considerations introduce a characteristic length.

Arbitrary inclusion geometries and quite general matrix and inclusion constitutive relations can be incorporated into the formulation. The specific boundary value problem analyzed here is one simulating a periodic array of rigid spherical inclusions in an isotropically hardening elastic-viscoplastic matrix. The aggregate is subject to both axial and radial stresses and a circular cylinder surrounding each inclusion is required to remain cylindrical throughout the deformation history in order to simulate the constraint of the surrounding material. By considering histories with different ratios of radial to axial stress, the effect of stress triaxiality on nucleation is studied. The numerical results are related to the description of void nucleation within the phenomenological constitutive framework of Gurson (1975, 1977).

2 Interface Model

Attention is directed toward an interface supporting a nominal traction field $\mathbf{T}$ (force/unit reference area) which, in general, has both normal and shearing components. Two material points, A and B, initially on opposite sides of the interface, are considered and the interfacial traction is taken to depend only on the displacement difference across the interface, $\Delta\mathbf{u}_{AB}$. At each point of the interface, we define

$$u_n = \mathbf{n}\cdot\Delta\mathbf{u}_{AB}, \quad u_t = \mathbf{t}\cdot\Delta\mathbf{u}_{AB}, \quad u_b = \mathbf{b}\cdot\Delta\mathbf{u}_{AB} \tag{2.1}$$

and

$$T_n = \mathbf{n}\cdot\mathbf{T}, \quad T_t = \mathbf{t}\cdot\mathbf{T}, \quad T_b = \mathbf{b}\cdot\mathbf{T} \tag{2.2}$$

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, NY, 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October 28, 1986; final revision, February 17, 1987.
Paper No. 87-WA/APM-9.

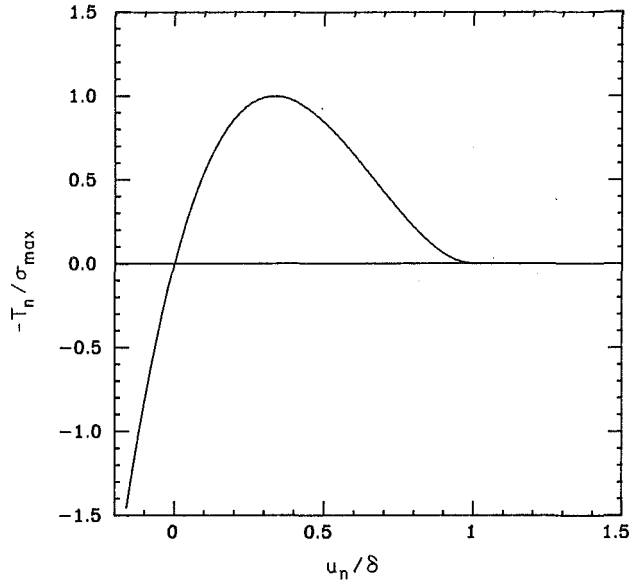


Fig. 1 Normal traction across the interface as a function of u_n with $u_t \equiv u_b \equiv 0$

In equations (2.1) and (2.2), **n**, **t**, **b** form a right-hand coordinate system chosen so that positive u_n corresponds to increasing interfacial separation and negative u_n corresponds to decreasing interfacial separation.

The mechanical response of the interface is described through a constitutive relation that gives the dependence of the tractions T_n, T_t, and T_b on u_n, u_t, and u_b. Here, this response is specified in terms of a potential $\phi(u_n, u_t, u_b)$, where

$$\phi(u_n,u_t,u_b) = -\int_0^u [T_n du_n + T_t du_t + T_b du_b] \tag{2.3}$$

As the interface separates, the magnitude of the tractions increases, achieves a maximum, and ultimately falls to zero when complete separation occurs. The magnitude of the tractions is taken to increase monotonically for negative u_n. Relative shearing across the interface leads to the development of shear tractions, but the dependence of the shear tractions on u_t and u_b is taken to be linear. The specific potential function used is

$$\begin{aligned}\phi(u_n,u_t,u_b) &= \frac{27}{4}\sigma_{max}\delta\left\{\frac{1}{2}\left(\frac{u_n}{\delta}\right)^2\left[1-\frac{4}{3}\left(\frac{u_n}{\delta}\right)+\frac{1}{2}\left(\frac{u_n}{\delta}\right)^2\right]\right.\\ &+\frac{1}{2}\alpha\left(\frac{u_t}{\delta}\right)^2\left[1-2\left(\frac{u_n}{\delta}\right)+\left(\frac{u_n}{\delta}\right)^2\right]\\ &\left.+\frac{1}{2}\alpha\left(\frac{u_b}{\delta}\right)^2\left[1-2\left(\frac{u_n}{\delta}\right)+\left(\frac{u_n}{\delta}\right)^2\right]\right\}\end{aligned} \tag{2.4}$$

for $u_n \le \delta$, where σ_{max} is the maximum traction carried by the interface undergoing a purely normal separation ($u_t \equiv u_b \equiv 0$), δ is a characteristic length and α specifies the ratio of shear to normal stiffness of the interface. When $u_n > \delta$, $\phi \equiv \phi_{sep}$, where ϕ_{sep} is the work of separation.

The interfacial tractions are obtained by differentiating equation (2.4) to give

$$\begin{aligned}T_n = \frac{-27}{4}\sigma_{max}&\left\{\left(\frac{u_n}{\delta}\right)\left[1-2\left(\frac{u_n}{\delta}\right)+\left(\frac{u_n}{\delta}\right)^2\right]\right.\\ &\left.+\alpha\left(\frac{u_t}{\delta}\right)^2\left[\left(\frac{u_n}{\delta}\right)-1\right]+\alpha\left(\frac{u_b}{\delta}\right)^2\left[\left(\frac{u_n}{\delta}\right)-1\right]\right\}\end{aligned} \tag{2.5}$$

$$T_t = \frac{-27}{4}\sigma_{max}\left\{\alpha\left(\frac{u_t}{\delta}\right)\left[1-2\left(\frac{u_n}{\delta}\right)+\left(\frac{u_n}{\delta}\right)^2\right]\right\} \tag{2.6}$$

$$T_b = \frac{-27}{4}\sigma_{max}\left\{\alpha\left(\frac{u_b}{\delta}\right)\left[1-2\left(\frac{u_n}{\delta}\right)+\left(\frac{u_n}{\delta}\right)^2\right]\right\} \tag{2.7}$$

for $u_n \le \delta$ and $T_n \equiv T_t \equiv T_b \equiv 0$ when $u_n > \delta$.

The motivation for choosing a potential of the form (2.4) is to obtain a response of the type shown in Fig. 1 where the normal traction, T_n, is plotted as a function of u_n with $u_t \equiv u_b \equiv 0$. The particular functional form (2.4) was chosen for analytical convenience; other forms can readily be used in the present framework. As can be seen in Fig. 1, the maximum interfacial stress is achieved at $u_n = \delta/3$ and complete separation occurs when $u_n = \delta$. The work of separation (in Fig. 1, the area under the curve between $u_n = 0$ and $u_n = \delta$) is

$$\phi_{sep} = 9\sigma_{max}\delta/16 \tag{2.8}$$

With $u_t \not\equiv u_b \not\equiv 0$, T_n, T_t, and T_b all vanish when $u_n = \delta$ so that, in general, δ serves as a characteristic length. Furthermore, due to the existence of a potential, equation (2.8) gives the work of separation regardless of the path. Equation (2.8) is regarded as defining the characteristic interface length δ by $\delta = 16\,\phi_{sep}/9\,\sigma_{max}$. Although δ has dimensions of length, it does not necessarily correspond to any physical distance.

The interface description adopted here is a phenomenological one characterized by the three parameters σ_{max}, δ and α. In the numerical examples parameter values representative of iron carbide particles in spheroidized carbon steels will be used. Based on the results of Argon et al. (1975), Goods and Brown (1979), and Fisher and Gurland (1981), the order of magnitude of δ can be estimated for this case; $\sigma_{max} \approx 10^3$ MPa and ϕ_{sep} is in the range of 1 to 10 Jm^{-2} so that $\delta \approx 10^{-9}$ to 10^{-8} m.

There does not appear to be any similar basis for specifying a value of the shear stiffness parameter α. In the calculations carried out here, the value of α is arbitrarily set to 10, the presumption being that the interface exhibits a stiffer response for relative sliding than for the normal displacement leading to separation. However, as will be illustrated subsequently, the numerical results in the specific cases analyzed are not very sensitive to the choice of α.

3 Finite Element Formulation

The finite element analysis is based on a convected coordinate Lagrangian formulation of the field equations with the initial unstressed state taken as reference. All field quantities are considered to be functions of convected coordinates, x^i, which serve as particle labels, and time t. This formulation has been employed extensively in previous finite element analyses, e.g., Needleman (1972) and Tvergaard (1976), and is reviewed by Needleman (1982).

Attention is confined to quasi-static deformations and, with body forces neglected, the principal of virtual work is written as

$$\int_V \tau^{ij}\delta E_{ij} dV + \int_{S_{int}} \delta\phi dS = \int_{S_{ext}} T^i \delta u_i dS \tag{3.1}$$

Here, τ^{ij} are the contravariant components of Kirchhoff stress ($\tau = J\sigma$, with σ the Cauchy stress) on the deformed convected coordinate net, V, S_{ext}, and S_{int} are the total volume (inclusion plus matrix), external surface and interfacial surface, respectively, of the body in the reference configuration, and

$$T^i = (\tau^{ij} + \tau^{kj}u^i_{,k})\nu_j \tag{3.2}$$

$$E_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i} + u^k_{,i}u_{k,j}) \tag{3.3}$$

where ν is the surface normal in the reference configuration, u_j are the components of the displacement vector on base vectors

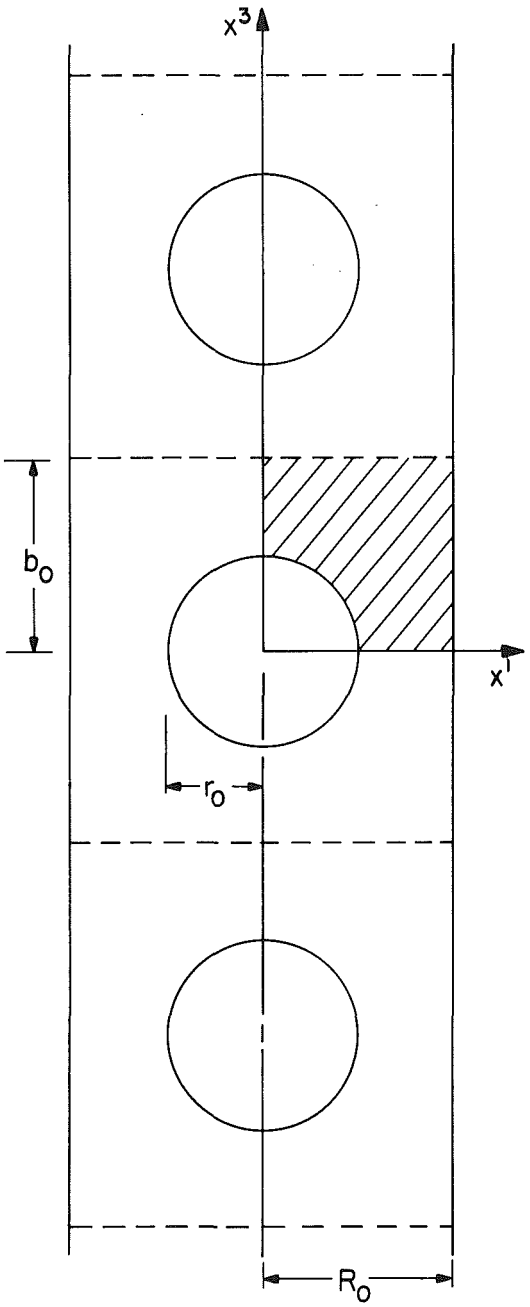


Fig. 2 Axisymmetric model of a material containing an array of spherical voids. Due to the assumed symmetry, only the shaded quadrant is analyzed numerically.

in the reference configuration and $()_{,i}$ denotes covariant differentiation in the reference frame.

For the specific boundary value problem considered here, we use a cylindrical coordinate system with radial coordinate x^1, circumferential angle x^2, and axial coordinate x^3. As sketched in Fig. 2, we consider spherical particles of radius r_0 located along the axis of a circular cylinder with an initial spacing of $2b_0$ between particle centers. The cylinder has initial radius R_0 and attention is confined to axisymmetric deformations so that all field quantities are independent of x^2. Furthermore, the circular cylindrical cell surrounding each particle is required to remain a circular cylinder throughout the deformation history and within each cell symmetry is assumed about the cell center line so that only the shaded region is analyzed numerically. As discussed by Tvergaard (1982), this axisymmetric configuration can be considered an approximation to a three dimensional array of hexagonal cylinders.

The boundary conditions for the axisymmetric region analyzed numerically are

$$\dot{u}^3 = 0, \quad \dot{T}^1 = 0, \quad \dot{T}^2 = 0, \quad \text{on } x^3 = 0 \tag{3.4}$$

$$\dot{u}^3 = \dot{U}_3 = \dot{\epsilon}_\infty b, \quad \dot{T}^1 = 0, \quad \dot{T}^2 = 0, \quad \text{on } x^3 = b_0 \tag{3.5}$$

$$\dot{u}^1 = \dot{U}_1, \quad \dot{T}^3 = 0, \quad \dot{T}^2 = 0, \quad \text{on } x^1 = R_0 \tag{3.6}$$

Here, $\dot{(\ )} = \partial()/\partial t$ and $\dot{\epsilon}_\infty$ is a prescribed constant while $\dot{U}_1$ is determined by the analysis. With these boundary conditions, the deformed circular cylindrical cell has radius $R = R_0 + U_1$, and height $2b = 2b_0 + 2U_3$.

The lateral displacement rate, $\dot{U}_1$ is determined from the condition that the average macroscopic true stresses acting on the cell follow the proportional history

$$\frac{\Sigma_1}{\Sigma_3} = \frac{\dot{\Sigma}_1}{\dot{\Sigma}_3} = \rho \tag{3.7}$$

with ρ a prescribed constant and

$$\Sigma_1 = \frac{R_0 b_0}{Rb}\left\{\frac{1}{b_0}\int_0^{b_0} [T^1]_{x^1 = R_0}\, dx^3\right\} \tag{3.8}$$

$$\Sigma_3 = \frac{R_0^2}{R^2}\left\{\frac{2}{R_0^2}\int_0^{R_0} [T^3]_{x^3 = b_0}\, x^1 dx^1\right\} \tag{3.9}$$

The matrix material is characterized as an elastic-viscoplastic isotropically hardening solid. The total rate of deformation, $\mathbf{D}$, is written as the sum of an elastic part, $\mathbf{D}^e$, and a plastic part $\mathbf{D}^P$, with

$$\mathbf{D}^e = \frac{1+\nu}{E}\,\hat{\tau} - \frac{\nu}{E}\,(\hat{\tau}:\mathbf{I})\mathbf{I} \tag{3.10}$$

$$\mathbf{D}^P = \frac{3\dot{\bar{\epsilon}}}{2\bar{\sigma}}\,\tau' \tag{3.11}$$

where $\hat{\tau}$ is the Jaumann rate of Kirchhoff stress, $\mathbf{I}$ is the identity tensor, $\hat{\tau}:\mathbf{I}$ is the trace of $\hat{\tau}$, $\dot{\bar{\epsilon}}$ is the effective plastic strain rate, E is Young's modulus, ν is Poisson's ratio and

$$\tau' = \tau - \frac{1}{3}\,(\tau:\mathbf{I})\mathbf{I}, \quad \bar{\sigma}^2 = \frac{3}{2}\,\tau':\tau' \tag{3.12}$$

$$\dot{\bar{\epsilon}} = \dot{\epsilon}_0[\bar{\sigma}/g(\bar{\epsilon})]^{1/m},$$

$$g(\bar{\epsilon}) = \sigma_0(\bar{\epsilon}/\epsilon_0 + 1)^N, \quad \epsilon_0 = \sigma_0/E \tag{3.13}$$

Here, $\bar{\epsilon} = \int \dot{\bar{\epsilon}} dt$ and the function $g(\bar{\epsilon})$ represents the effective stress versus effective strain response in a tensile test carried out at a strain-rate such that $\dot{\bar{\epsilon}} = \dot{\epsilon}_0$. Also, σ_0 is a reference strength and N and m are the strain hardening exponent and strain rate hardening exponent, respectively.

Expanding equation (3.1) about a state of approximate equilibrium gives

$$\Delta t \int_V [\dot{\tau}^{ij}\delta E_{ij} + \tau^{ij}\dot{u}^k_{,i}\delta u_{k,j}]dV$$

$$+ \Delta t \int_{S_{\text{int}}} [S^{nn}\dot{u}_n\delta u_n + S^{nt}(\dot{u}_n\delta u_t + \dot{u}_t\delta u_n)$$

$$+ S^{tt}\dot{u}_t\delta u_t]dS = \Delta t \int_{S_{\text{ext}}} \dot{T}^i\delta u_i dS - \left[\int_V \tau^{ij}\delta E_{ij}dV\right.$$

$$\left. + \int_{S_{\text{int}}} \delta\phi dS - \int_{S_{\text{ext}}} T^i\delta u_i dS\right] \tag{3.14}$$

where $S^{ij} = \partial^2\phi/\partial u_i\partial u_j$. In equation (3.14), the integral over S_{int} has been specialized to the case of axisymmetric deformations with the b direction identified with that of the circumferential angle so that $u_b \equiv 0$. The term in square brackets on the right-hand side of equation (3.14) is an equilibrium correction term that vanishes when the known state is an exact equilibrium state.

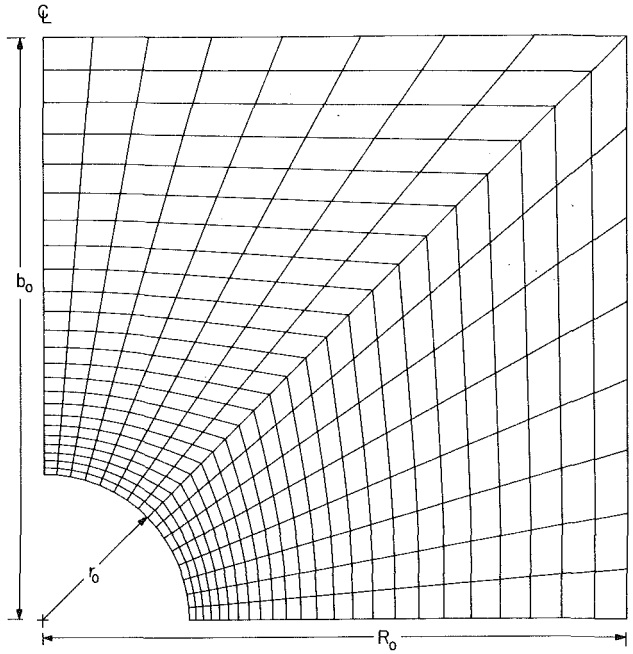


Fig. 3 The 16 × 24 finite element mesh used in the calculations. Each quadrilateral consists of four linear displacement "crossed" triangular elements.

The set of equations for the unknown displacement rates is obtained by combining equations (3.10) and (3.11) and then using the relation between the Jaumann and convected stress rates in equation (3.14). The finite element mesh used in the numerical calculations is shown in Fig. 3. The mesh has 16 quadrilaterals around the inclusion and 24 quadrilaterals in the radial direction. Each quadrilateral consists of four "crossed" triangles. The circular inclusion is approximated by a polygon consisting of linear segments that are the sides of the elements along the interface. The method for evaluating the integrals along S_{int} in equation (3.14) is similar to an approach used in surface diffusion calculations by Needleman and Rice (1980). The integration scheme uses four Gauss points within each linear segment and the interfacial tractions are evaluated at the Gauss integration points rather than at the finite element nodes. This permits partial debonding within a linear segment.

The deformation history is calculated in a linear incremental manner and, in order to increase the stable time step, the rate tangent modulus method of Peirce et al. (1984) is used. This is a forward gradient method based on an estimate of the plastic strain rate in the interval between t and $t + \Delta t$. The incremental boundary value problem is solved using a combined finite element Rayleigh-Ritz method (Tvergaard, 1976).

In most cases, the prescribed overall strain rate, $\dot{\epsilon}_\infty$ is taken constant and equal to the reference strain rate $\dot{\epsilon}_0$. However, equilibrium solutions do not necessarily exist to the boundary value problem so posed. For a certain range of interface characterizations, equilibrium solutions only exist if U_3 decreases during decohesion. In such cases, the boundary value problem is modified so that equilibrium solutions are sought for increasing interfacial separation and the prescribed velocity condition in equation (3.5) is replaced by $\dot{u}_3 = constant$ along $x^3 = b_0$; $\dot{\epsilon}_\infty$ then becomes an unknown determined by the solution procedure.

4 Numerical Results

In the numerical calculations carried out here, the inclusion volume fraction and geometry, as well as the matrix material properties, remain fixed; only the interface characteristics are varied. The inclusion geometry is specified by $b_0/R_0 = 1$ and $r_0/R_0 = 0.25$, giving an inclusion volume fraction of 1.04 percent. The matrix material properties are $E = 500\ \sigma_0$, $\nu = 0.3$, $N = 0.1$, and $m = 0.01$. In most calculations, the value $\sigma_{\max} = 3\ \sigma_0$ is employed which is a plausible value for iron carbide particles in a spheroidized steel, e.g., with a yield strength of 350 to 450 MPa and an interfacial cohesive strength in the range 1000 to 1400 MPa (Argon et al., 1975; Goods and Brown, 1979; Cialone and Asaro, 1979; and Brownrigg et al., 1983).

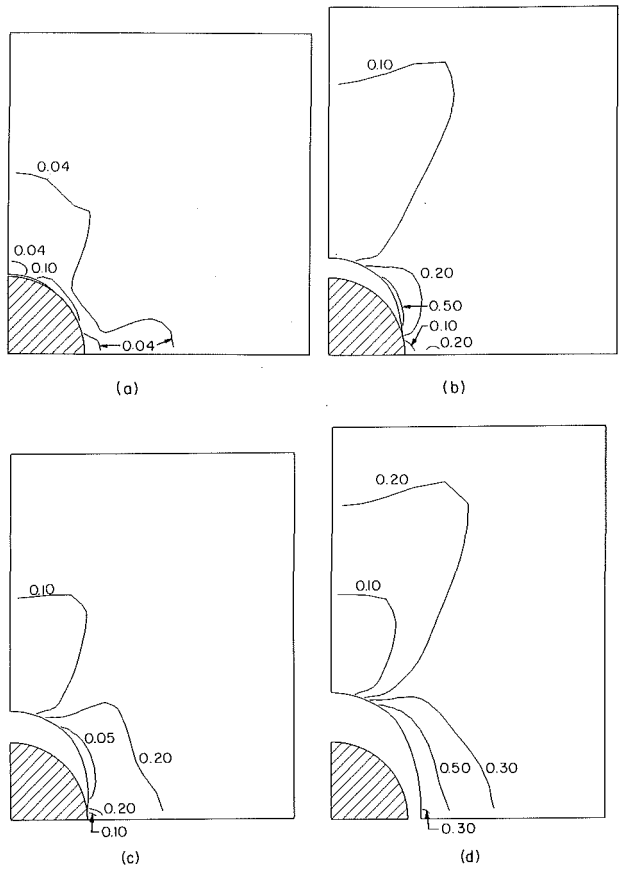


Fig. 4 Contours of constant plastic strain, $\bar{\epsilon}$, in the deformed configuration of the quadrant analyzed numerically. The rigid inclusion is shaded. The volume fraction of inclusions is 1.04 percent. The interface is characterized by $\sigma_{max} = 3\sigma_0$, $\delta/r_0 = 0.01$, $\alpha = 10.0$ and the stress triaxiality parameter, ρ, in equation (3.7) is 0.5. (a) $\epsilon_a = 0.040$; (b) $\epsilon_a = 0.121$; (c) $\epsilon_a = 0.169$; (d) $\epsilon_a = 0.240$.

Figure 4 shows contours of constant plastic strain, $\bar{\epsilon}$, at various stages of the nucleation process. While the matrix and inclusion remain bonded, the main strain concentration occurs along the inclusion surface at about 45 deg from the tensile axis. Debonding does not begin at the axis of symmetry; it begins at the end of the strain concentration nearest the symmetry axis. The crack rather rapidly propagates to the axis of symmetry and then a spherical cap void opens. Already, at the stage of deformation shown in Fig. 4(*a*), the maximum normal displacement is on the symmetry axis and there are "dead" zones at 0 deg and 90 deg. As the decohering region propagates toward the midsection, the deformation pattern changes to one where the maximum straining is near the midsection.

Curves of effective stress, Σ_e, versus axial strain, $\epsilon_a = ln(1 + U_3/b_0)$, are plotted in Fig. 5 for three values of δ/r_0, where, for the axisymmetric configuration analyzed, we define

$$\Sigma_e = |\Sigma_3 - \Sigma_1|, \quad \Sigma_h = \frac{1}{3}(\Sigma_3 + 2\Sigma_1) \tag{4.1}$$

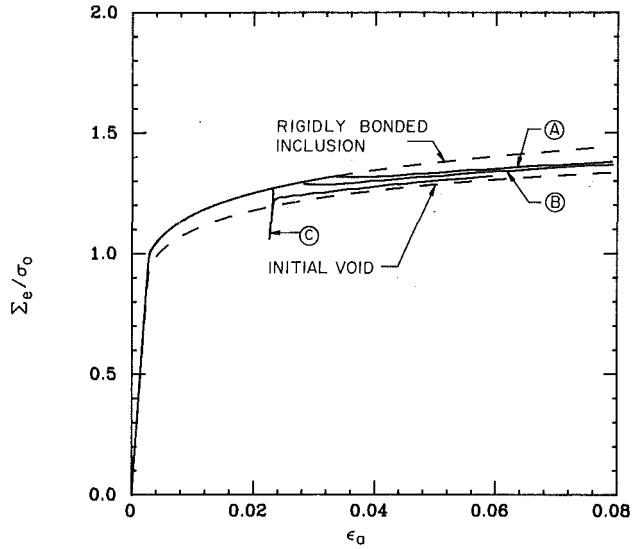


Fig. 5 Curves of aggregate effective stress, Σ_e, versus axial strain, ϵ_a, for a 1.04 percent volume fraction of inclusions with $\rho = 0.5$ in equation (3.7) and using three interface characterizations. In all three cases $\sigma_{max} = 3\sigma_0$ and $\alpha = 10.0$; (*A*) $\delta/r_0 = 0.01$; (*B*) $\delta/r_0 = 0.006$; (*C*) $\delta/r_0 = 0.002$. For comparison purposes, corresponding curves for a 1.04 percent volume fraction of perfectly bonded inclusions and for a 1.04 percent volume fraction of initial voids are also shown.

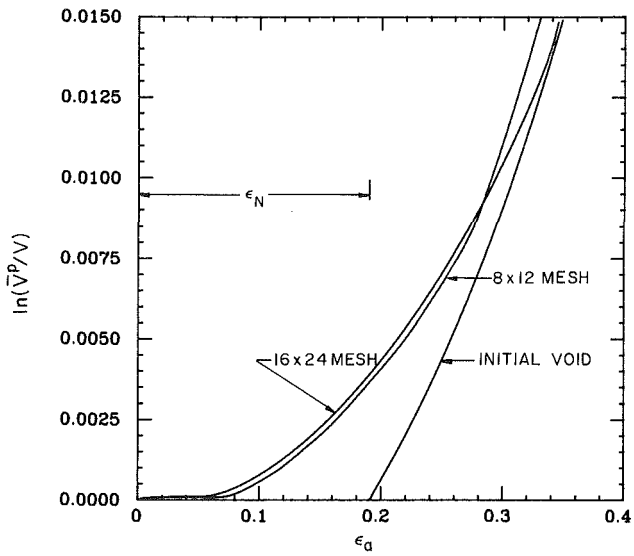


Fig. 6 Curves of normalized logarithmic plastic volume change, $ln(\bar{V}^p/V)$, versus axial strain, ϵ_a, for a 1.04 percent volume fraction of inclusions with $\rho = 0.5$ in equation (3.7) and $\sigma_{max} = 3\ \sigma_0$, $\alpha = 10.0$ and $\delta/r_0 = 0.04$. Numerical results are shown obtained from a coarse 8 × 12 mesh as well as results obtained using the 16 × 24 mesh shown in Fig. 3. The normalized logarithmic plastic volume change, $ln(\bar{V}^p/V)$, versus axial strain, ϵ_a, curve for a 1.04 percent volume fraction of initial voids is shown shifted by an amount ϵ_N along the strain axis.

With regard to iron carbide particles in spheroidized carbon steels, these values of δ/r_0 correspond to particle sizes of the order of 1 micron (10^{-6} m).

The variation of δ/r_0 can be regarded either in terms of a variation in ϕ_{sep} at fixed particle size or as a variation in particle size at fixed work of separation. A sufficiently small value of δ/r_0 gives rise to "brittle" interface behavior, while larger values lead to a more ductile mode of separation. For comparison purposes, the corresponding curves for a rigidly bonded inclusion and for an initial void are shown.

As debonding progresses, the overall stress-strain behavior changes from that characteristic of a matrix reinforced by rigid inclusions to one weakened by an equal volume fraction of voids. For the larger two values of δ/r_0, this transition takes place gradually and with increasing extension. As δ/r_0 decreases, the stress drop becomes more abrupt and, for the case with $\delta/r_0 = 0.002$, the stress drop cannot be affected with continued plastic loading. Even though explicit elastic unloading is not incorporated into the material description, the material response is essentially linear elastic during this abrupt stress drop and, as can be seen in Fig. 5, the stress drop occurs with the initial elastic slope. Initial debonding occurs in the element nearest the axis of symmetry, in contrast to the situation for a more ductile interface, where debonding initiates off the axis. The stress drop occurs before initial debonding, when the traction across this interface segment is on the descending branch of the traction versus displacement curve in Fig. 1. After initial debonding, the stress increases, although, as can be seen in Fig. 5, there are slight oscillations (which may be an artifact of the numerics) in the overall stress-strain curve as debonding propagates along the interface.

As illustrated in Fig. 4, void nucleation is a process that occurs over a range of strain. A void nucleation strain, ϵ_N, can be defined in various ways, with the appropriate definition depending on the context in which it is to be used. For example, the void nucleation strain can be identified with the strain at which initial debonding takes place or with the strain at which complete separation occurs. In the phenomenological constitutive framework of Gurson (1975, 1977), a void created by inclusion debonding has been regarded as equivalent to a void occupying the same fraction as the inclusion being abruptly introduced into the material at ϵ_N. To define a nucleation strain in this context, the normalized logarithmic plastic volume change, $ln\ (\bar{V}^p/V)$, is calculated via

$$ln\left(\frac{\bar{V}^p}{V}\right) = ln\left(\frac{b}{b_0}\right) + 2ln\left(\frac{R}{R_0}\right) - \frac{1-2\nu}{E}\,\Sigma_h \qquad (4.2)$$

The expression (4.2) is approximate because the effect of the inclusions is not accounted for in the elastic volume change term, but this is not of significance for the present purpose.

In Fig. 6, plastic volume change versus axial strain curves are plotted using two different finite element meshes; one is a coarse 8 × 12 mesh, while the other is the 16 × 24 mesh shown in Fig. 3 and used in all the remaining calculations reported on here. The results in Fig. 6 are for a rather "ductile" interface; $\delta/r_0 = 0.04$. Initial debonding takes place at $\epsilon_a = 0.068$ with the 16 × 24 mesh and at $\epsilon_a = 0.08$ with the coarse mesh. On the other hand, complete separation occurs somewhat earlier for the coarse mesh; at $\epsilon_a = 0.29$ as compared with $\epsilon_a = 0.34$ with the finer mesh. The more rapid separation with the coarse mesh is expected since the last points to debond are in the low strain region near the x^1 axis and the strain depression is resolved better in the fine mesh calculation. In the 16 × 24 mesh calculation complete debonding has occurred in all elements except the last one at $\epsilon_a = 0.31$. The strain interval over which debonding occurs depends on the value of δ/r_0. With $\delta/r_0 = 0.01$, but all other parameters as in Fig. 6, initial debonding takes place at $\epsilon_a = 0.034$ and complete separation at $\epsilon_a = 0.18$.

A plastic volume change versus axial strain curve is also shown in Fig. 6 (using the 16 × 24 mesh) for an initial void of the same size as the inclusion. The curve of $ln(\bar{V}^p/V)$ versus axial strain for the void is shifted an amount ϵ_N. The value of ϵ_N for which the $ln(\bar{V}^p/V)$ versus ϵ_a curves nearly coincide at the larger volume changes shown (i.e., volume changes of the order of 1 percent) is taken as the nucleation strain. The fine mesh calculation gives a nucleation strain of 0.19, as shown, while the coarse mesh calculation implies $\epsilon_N = 0.17$.

This definition of nucleation strain is inherently imprecise. The sensitivity to the particular volume fraction at which the values of $ln(\bar{V}^p/V)$ are matched depends on the "ductility" of

the interface, with more ductile interfaces being more sensitive to the value of $ln(\bar{V}^p/V)$ chosen. In Fig. 6 where $\delta/r_0 = 0.04$, the above values of ϵ_N were based on matching the response at $ln(\bar{V}^p/V) = 0.015$. Using $ln(\bar{V}^p/V) = 0.01$ gives nucleation strains of 0.172 and 0.176 for the 8×12 and 16×24 meshes, respectively. With $\delta/r_0 = 0.01$, but all other parameters as in Fig. 6, defining ϵ_N based on $ln(\bar{V}^p/V) = 0.015$, 0.01, and 0.005 leads to $\epsilon_N = 0.084$, 0.082, and 0.079.

Some calculations were carried out to explore the dependence of the nucleation strain, as defined above, on interface properties. A series of calculations were carried out using three values of δ/r_0, i.e., for three sizes of inclusion. The model predicts a definite size effect and the results will be discussed in connection with void nucleation criteria in the context of Gurson's (1975, 1977) constitutive framework. There is a stronger dependence of nucleation strain on interfacial strength than on size. To illustrate this, two sets of interface parameters were chosen that have values of ϕ_{sep} in (2.8) 60 percent of that for the case with $\sigma_{max} = 3\ \sigma_0$ and $\delta/r_0 = 0.01$; in one case $\sigma_{max} = 3\sigma_0$ and $\delta/r_0 = 0.006$, while in the other case $\sigma_{max} = 1.8\ \sigma_0$ and $\delta/r_0 = 0.01$. These give rise to nucleation strains of 0.060 and 0.024, respectively, while in the reference case $\epsilon_N = 0.084$.

The role of the shear stiffness parameter, α, was investigated for the case $\rho = 0.4$, $\sigma_{max} = 3\sigma_0$, $\delta/r_0 = 0.02$. With $\alpha = 10.0$, the nucleation strain is 0.23. Increasing α to 50 increases the nucleation strain to 0.25, while with $\alpha = 1$, $\epsilon_N = 0.20$. In fact, taking $\alpha = 0$ in this case gives a nucleation strain of 0.18. Hence, for the geometry and loading conditions here, the interface shear stiffness plays a relatively minor role. This may not be the case for other inclusion geometries and for imposed stress states with a large shear component.

5 Void Nucleation Criterion

Within the constitutive framework for progressively cavitating solids introduced by Gurson (1975, 1977), the voids are represented in terms of a single parameter, the void volume fraction, f. The evolution equation for the void volume fraction includes contributions from both void growth and void nucleation,

$$\dot{f} = \dot{f}_{\text{growth}} + \dot{f}_{\text{nucleation}} \tag{5.1}$$

The void growth contribution is determined from the plastic flow rule using the condition that the matrix material is plastically incompressible whereas the void nucleation contribution is specified separately. Although various void nucleation criteria can be formulated within this framework, two have been used in practice (Gurson, 1975, 1977; Needleman and Rice, 1978). One is a plastic strain criterion for which

$$\dot{f}_{\text{nucleation}} = D\dot{\bar{\epsilon}} \tag{5.2}$$

while the other is the stress-based criterion

$$\dot{f}_{\text{nucleation}} = B(\dot{\Sigma}_e + \dot{\Sigma}_h) \tag{5.3}$$

In equation (5.3), Σ_e is identified with the matrix effective stress appearing in the Gurson (1975, 1977) flow potential. As long as the void volume fraction is zero, as it is prior to nucleation, the matrix effective stress and the macroscopic effective stress are equal. In (5.2), D is considered a function of $\bar{\epsilon}$, while analogously in equation (5.3), B is taken to be a function of $(\Sigma_e + \Sigma_h)$ so that the quantity $(\Sigma_e + \Sigma_h)$ plays the role of a nucleation stress.

Analyses of localization, carried out within the Gurson (1975, 1977) framework, indicate that equations (5.2) and (5.3) can lead to quite different predictions of macroscopic ductility (Needleman and Rice, 1978; Saje et al. 1982). What is of particular significance in this regard is that the hydrostatic stress dependence of void nucleation in equation (5.3) leads to a strong nonnormality in the plastic flow rule which promotes early flow localization.

Table 1 Nucleation strain and stress for various values of stress triaxiality. The interface is characterized by $\sigma_{max}/\sigma_0 = 3$, $\alpha = 10$ and $\delta/r_0 = 0.02$.

ρ	Σ_h/Σ_e	ϵ_N	$(\Sigma_e)_N/\sigma_0$	$(\Sigma_e + c\Sigma_h)_N/\sigma_0$
				$c = 0.35$
0.250	0.667	0.56	1.732	2.137
0.333	0.833	0.33	1.627	2.102
0.400	1.00	0.23	1.555	2.099
0.500	1.33	0.12	1.443	2.117
0.625	2.00	0.048	1.268	2.156
			Ave.	2.122

Table 2 Nucleation strain and stress for various values of stress triaxiality. The interface is characterized by $\sigma_{max}/\sigma_0 = 3$, $\alpha = 10$ and $\delta/r_0 = 0.01$.

ρ	Σ_h/Σ_e	ϵ_N	$(\Sigma_e)_N/\sigma_0$	$(\Sigma_e + c\Sigma_h)_N/\sigma_0$	
				$c = 0.35$	$c = 0.39$
0.250	0.667	0.50	1.723	2.125	2.171
0.333	0.833	0.27	1.594	2.058	2.111
0.400	1.00	0.17	1.509	2.038	2.098
0.500	1.33	0.084	1.388	2.036	2.110
0.625	2.00	0.024	1.192	2.027	2.122
			Ave.	2.057	2.122

Table 3 Nucleation strain and stress for various values of stress triaxiality. The interface is characterized by $\sigma_{max}/\sigma_0 = 3$, $\alpha = 10$ and $\delta/r_0 = 0.04$.

ρ	Σ_h/Σ_e	ϵ_N	$(\Sigma_e)_N/\sigma_0$	$(\Sigma_e + c\Sigma_h)_N/\sigma_0$	
				$c = 0.35$	$c = 0.314$
0.250	0.667	0.61	1.752	2.161	2.119
0.333	0.833	0.40	1.664	2.149	2.099
0.400	1.00	0.31	1.605	2.167	2.109
0.500	1.33	0.19	1.501	2.201	2.129
0.625	2.00	0.070	1.322	2.248	2.152
			Ave.	2.185	2.122

In order to explore the predicted hydrostatic stress dependence of the void nucleation strain, ϵ_N, calculations were carried out for various values of the stress ratio ρ in equation (3.7). In each case ϵ_N is defined in the manner sketched in Fig. 6. Tables 1 to 3 illustrate the hydrostatic stress dependence of ϵ_N for interfaces characterized by three values of δ/r_0. The other interface parameters are kept fixed at $\sigma_{max} = 3\sigma_0$ and $\alpha = 10$.

For low stress triaxiality, $\Sigma_h/\Sigma_e = 0.667$, the nucleation strain varies between 0.50 and 0.61 as δ/r_0 is increased by a factor of four, from 0.01 to 0.04. At higher values of the stress triaxiality, the absolute magnitude of the variation in ϵ_N is smaller, but the relative variation is greater; for example, when $\Sigma_h/\Sigma_e = 2.00$, the nucleation strain increases from 0.024 to 0.070 as δ/r_0 is varied over the same range. With δ regarded as fixed, this corresponds to a decrease in nucleation strain with increasing particle size at fixed volume fraction.

Also shown in Tables 1 to 3 is the outcome of correlating the results in terms of an effective nucleation stress, written as

$$\Sigma_N = \Sigma_e + c\Sigma_h \tag{5.4}$$

where Σ_e and Σ_h are obtained from equation (4.1) at $\epsilon_a = \epsilon_N$.

For the case $\delta/r_0 = 0.02$, $c = 0.35$ leads to a mean nucleation stress, Σ_N, of $2.122\ \sigma_0$. If equation (5.4) held precisely, the value of Σ_N would be independent of ρ in Table 1. The maximum deviation from the mean is 1.5 percent and this occurs for the lowest nucleation strain where the work hardening

is highest; more significant is the deviation from the mean for $\Sigma_h/\Sigma_e = 0.667$ since this occurs at large strains where the work hardening rate is low.

Tables 2 and 3 show the results of calculating nucleation stresses for $\delta/r_0 = 0.01$ and $\delta/r_0 = 0.04$ in two ways; one calculation uses a constant c, while in the other calculation c is chosen so that the average value of Σ_N is the same for all three cases. A somewhat better correlation is obtained by varying c suggesting that more "brittle" interfaces (larger particles) may be characterized by a more strongly hydrostatic stress dependent nucleation stress.

With $c = 1$ in equation (5.4), the nucleation stress appearing in equation (5.3) is recovered, and with $c = 1$ and $\Sigma_h/\Sigma_e = 0.667$ a nucleation stress of 2.887 is obtained; at $\Sigma_h/\Sigma_e = 1.00$, Σ_N $(c=1) = 3.109$ and at $\Sigma_h/\Sigma_e = 2.00$, Σ_N $(c=1) = 3.806$. Although it is interesting to note that these values, particularly for the cases with lower stress triaxiality, are reasonably close to $\sigma_{\max}$, use of $c = 1$ in equation (5.4) gives rise to a strongly hydrostatic stress dependent nucleation stress. Values of c less than unity in equation (5.4) can be thought of as due to part of the remote hydrostatic stress being "converted" to local shearing stresses around the inclusion, the magnitude of which are limited by the work hardening capacity of the material.

6 Concluding Remarks

The cohesive zone interface model developed here provides a unified description of void initiation from initial debonding through complete separation and subsequent void growth. This cohesive zone model is particularly attractive when, as is often the case, interfacial strengths are relatively weak, say of the order of several times the yield strength of the matrix material. Then the very high stress gradients associated with cracks in homogeneous bodies do not develop and standard finite-strain finite-element methods can be extended to incorporate the interface integrals. The model is a purely continuum one, so that discrete dislocation effects are not accounted for, but the formulation provides a framework for analyzing the effects of matrix and inclusion material properties, inclusion size and shape, and imposed stress state and loading rate on the nucleation process.

Dimensional considerations introduce a characteristic interface length into the model and numerical results exhibit a ductile to brittle transition in the mode of separation. For sufficiently large inclusions (relative to the characteristic length) equilibrium solutions do not exist for increasing extension during debonding. The interface debonds in a "brittle" manner, with an abrupt stress drop. One can speculate that if this were to occur at a particular weak inclusion, the stress drop could lead to load shedding to nearby inclusions. The increased stress could then precipitate further nucleation, leading to another stress redistribution and so on, so that a profusion of voids are nucleated over a rather narrow strain interval. In actuality, in such a case, debonding would occur dynamically and dynamic effects may well play a significant role in the mechanics of the stress redistribution. By way of contrast, the smooth load drop associated with a more "ductile" interface (smaller inclusions) suppresses this mechanism of void profusion. In this regard it is important to note that the range of strain over which voids nucleate can affect stability against flow localization; void profusion over a narrow range of strain is potentially destabilizing (Needleman and Rice, 1978; Saje et al., 1980).

The onset of nucleation at various levels of triaxiality of the imposed stress state has been correlated, within the framework of Gurson's (1975, 1977) constitutive relation for progressively cavitating solids, in terms of a critical nucleation stress. This critical nucleation stress depends linearly on the hydrostatic tension, but with a coefficient that is less than unity.

Acknowledgments

The support of the National Science Foundation (Solid Mechanics Program) through grant MSM-8419338 is gratefully acknowledged. The computations reported on here were carried out at the John von Neumann Center for Scientific Computing and at the Pittsburgh Supercomputer Center. Access to these facilities was made possible through a National Science Foundation program to provide supercomputer access to researchers.

References

Argon, A. S., Im, J., and Safoglu, R., 1975, "Cavity Formation from Inclusions in Ductile Fracture," *Met. Trans.*, Vol. 6A, pp. 825–837.

Barenblatt, G. I., 1962, "Mathematical Theory of Equilibrium Cracks," *Adv. Appl. Mech.*, Vol. 7, pp. 56–129.

Brown, L. M., and Stobbs, W. M., 1971, "The Work-Hardening of Copper-Silica II. The Role of Plastic Relaxation," *Phil. Mag.*, Vol. 23, pp. 1201–1233.

Brownrigg, A., Spitzig, W. A., Richmond, O., Tierlinck, D., and Embury, J. D., 1983, "The Influence of Hydrostatic Pressure on the Flow Stress and Ductility of a Spheroidized 1045 Steel," *Acta. Metall.*, Vol. 31, pp. 1141–1150.

Chang, Y. W., and Asaro, R. J., 1978, "Bauschinger Effects and Work Hardening in Spheroidized Steels," *Metal Science*, Vol. 12, pp. 277–284.

Cialone, H., and Asaro, R. J., 1978, "The Role of Hydrogen in the Ductile Fracture of Plain Carbon Steels," *Met. Trans.*, Vol. 10A, pp. 367–375.

Dugdale, D. S., 1960, "Yielding of Steel Sheets Containing Slits," *J. Mech. Phys. Solids*, Vol. 8, pp. 100–104.

Fisher, J. R., and Gurland, J., 1981, "Void Nucleation in Spheroidized Carbon Steels; Part 2: Model," *Metal Science*, Vol. 15, pp. 193–202.

Goods, S. H., and Brown, L. M., 1979, "The Nucleation of Cavities by Plastic Deformation," *Acta. Metall.*, Vol. 27, pp. 1–15.

Gurson, A. L., 1975, "Plastic Flow and Fracture Behavior of Ductile Materials Incorporating Void Nucleation, Growth and Interaction," Ph.D. Thesis, Brown University.

Gurson, A. L., 1977, "Continuum Theory of Ductile Rupture by Void Nucleation and Growth; Part I: Yield Criterion and Flow Rules for Porous Ductile Materials," *J. Engr. Mat. Tech.*, Vol. 99, pp. 2–15.

Keer, L. M., Dundurs, J., and Kiattikomol, K., 1973, "Separation of a Smooth Circular Inclusion form a Matrix," *International Journal of Engineering Science*, Vol. 11, pp. 1221–1233.

Needleman, A., 1972, "A Numerical Study of Necking in Circular Cylindrical Bars," *J. Mech. Phys. Solids*, Vol. 20, pp. 111–120.

Needleman, A., 1982, "Finite Elements for Finite Strain Plasticity Problems," *Plasticity of Metals at Finite Strain: Theory, Computation and Experiment*, E. H. Lee and R. L. Mallett, eds., pp. 387–436.

Needleman, A., and Rice, J. R., 1978, "Limits to Ductility Set by Plastic Flow Localization," *Mechanics of Sheet Metal Forming*, Koistinen, D. P., and Wang, N.-M., eds., pp. 237–265.

Needleman, A., and Rice, J. R., 1980, "Plastic Creep Flow Effects in the Diffusive Cavitation of Grain Boundaries," *Acta Metall.*, Vol., 28, pp. 1313–1332.

Peirce, D., Shih, C. F., and Needleman, A., 1984, "A Tangent Modulus Method for Rate Dependent Solids," *Comp. Struct.*, Vol. 18, pp. 875–887.

Saje, M., Pan, J., and Needleman, A., 1982, "Void Nucleation Effects on Shear Localization in Porous Plastic Solids," *Int. J. Fract.*, Vol. 19, pp. 163–182.

Taya, M., and Patterson, W. G., 1982, "Growth of a Debonded Void at a Rigid Secondary Particle in a Viscous Metal," *J. Matl. Sci.*, Vol. 17, pp. 115–120.

Tvergaard, V., 1976, "Effect of Thickness Inhomogeneities in Internally Pressurized Elastic-Plastic Spherical Shells," *J. Mech. Phys. Solids*, Vol. 24, pp. 291–304.

Tvergaard, V., 1982, "On Localization in Ductile Materials Containing Spherical Voids," *Int. J. Fract.*, Vol. 18, pp. 237–252.

An Elastic-Viscoplastic Model for Metals Subjected to High Compression

M. B. Rubin
Faculty of Mechanical Engineering,
Technion—Israel Institute of Technology,
Haifa 32000, Israel

Specific constitutive equations are proposed for a material exhibiting isotropic-elastic response in its reference configuration, strain-rate, temperature and density dependent plastic flow with isotropic and directional hardening, and thermal recovery of hardening. The shear modulus is temperature and density dependent and it vanishes when the temperature reaches the density dependent melting temperature. These equations include modifications, relative to those proposed by Rubin (1986), which are appropriate to describe metals subjected to high compression. The constitutive functions characterizing pressure are determined by comparison with a Mie-Grüneisen equation of state which includes functions that are obtained from common shock-wave experiments. To examine some of the features of these equations at high compression we consider an example of homogeneous uniaxial strain and show that the deviatoric stress may be quite large at ultra high compression rates and high compression.

Introduction

Recently, Rubin (1986) considered a rather general class of constitutive equations modeling elastic-viscoplastic behavior of metals. Restrictions on these equations were obtained to ensure consistency with the thermodynamic procedures proposed by Green and Naghdi (1977, 1978). In addition, specific constitutive equations were proposed for a material exhibiting isotropic-elastic response in its reference configuration, strain-rate and temperature dependent plastic flow with isotropic and directional hardening, and thermal recovery of hardening. These specific equations represent a generalization to the nonlinear region of the works of Bodner and Partom (1972) and Bodner (1984, 1985).

The objective of this paper is to discuss specific constitutive equations which characterize the elastic-viscoplastic behavior of metals subjected to high compression. Modifications of the specific equations proposed by Rubin (1986) are presented which are motivated by the interest in shock-wave plate impact experiments (Clifton, 1983) which are used to obtain material properties at high strain rates. Among other modifications, we note that the Helmholtz free energy (equation 1(*a*)) is modified to include a temperature and density dependent shear modulus which vanishes when the temperature reaches the density dependent melting temperature. The flow rule (equation 16(*a*)) and the requirement of plastic incompressibility are also modified. We emphasize that these modified equations are within the scope of the general equations considered in Rubin (1986) so it is not necessary to reexamine their consistency with the first law of thermodynamics. However, the various restrictions associated with the second law of thermodynamics must be reconsidered.

An important feature of these constitutive equations is the equation for stress (3*a*) which is a hyper-elastic equation relating stress to deformation quantities. In particular, stress is not calculated using a hypo-elastic equation for a stress rate as is common. For a recent review of the state of hypo-elastic equations see Reed and Atluri (1985).

In the following sections we record the modified constitutive equations and discuss each of the modifications. Then, we show how some of the constitutive functions can be estimated by comparing with a Mie-Grüneissen equation of state (equation (24)) which is used often to analyze shock-wave experiments. To examine some of the features of these equations at high compression we consider an example of homogeneous uniaxial strain and show that the deviatoric stress may be quite large at ultra high compression rates and high compression.

Specific Constitutive Equations

In this section, we record the modified constitutive equations and discuss each of the modifications (relative to the equations proposed in Rubin, 1986). We refer all quantities to the reference configuration because the relevant quantities referred to the reference configuration are trivially invariant under superposed rigid body motions. An important manifestation of this invariance property is that the evolution equations may be formulated without introducing special invariant rates like the Jaumann rate.

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April 1, 1986; final revision, November 18, 1986.

Here, we consider a finite body with material points which are located by their position vector $\mathbf{X}$ in the reference configuration. A motion of the body is defined by a sufficiently smooth vector function χ, which assigns position $x=\chi(\mathbf{X},t)$ to each material point $\mathbf{X}$ at each instant of time t. The model under consideration here may be characterized by specific constitutive equations for the specific (per unit mass) Helmholtz free energy ψ, the entropy flux $\mathbf{P}$ (per unit area in the reference configuration), and the specific internal rate of entropy production ξ which take the form[1]:

$$2\rho_0\psi = -2\rho_0 h(\theta) - (\theta-\theta_0)f_1(I_3) + f_2(I_3) + 2\rho_0\psi', \tag{1a}$$

$$2\rho_0\psi' = \hat{\mu}\ell n\alpha, \quad \alpha = \left(\frac{\mathbf{C}\cdot\mathbf{C}_p^{-1}}{3}\right)^3 \frac{I_{3p}}{I_3}, \tag{1b,c}$$

$$\mathbf{P} = -\frac{K(I_3,\theta)}{\theta} I_3^{1/2}\mathbf{C}^{-1}\mathbf{G}, \tag{1d}$$

$$\rho_0\theta\xi = -\mathbf{P}\cdot\mathbf{G} + \rho_0\theta\xi', \tag{1e}$$

$$I_3 = \det\mathbf{C}, \quad I_{3p} = \det\mathbf{C}_p, \tag{1f,g}$$

$$\hat{\mu} = \hat{\mu}(I_3,\theta), \quad \mathbf{G} = \frac{\partial\theta}{\partial\mathbf{X}}, \tag{1h,i}$$

In these equations: $\mathbf{F}=\partial\mathbf{x}/\partial\mathbf{X}$ is the deformation gradient; $\mathbf{C}=\mathbf{F}^T\mathbf{F}$ is the Cauchy-Green deformation tensor; $\mathbf{C}_p$ is the plastic deformation tensor which will be defined later through a flow rule; θ is the absolute temperature and θ_0 is its reference value; $\mathbf{G}$ is the temperature gradient with respect to $\mathbf{X}$; $I_3=(\rho_0/\rho)^2$ is a pure measure of total dilatation with ρ being the mass density in the present configuration and ρ_0 being its reference value; I_{3p} is a pure measure of plastic dilatation; $\hat{\mu}$ is a shear modulus and μ_0 is its reference value; K is the heat conduction coefficient; ξ' is related to plastic dissipation; and h, f_1, f_2 are functions to be specified.

The part ψ' in equation (1*b*) of the Helmholtz free energy represents the strain energy of elastic distortion and has been modified (relative to Rubin, 1986) by introducing the scalar α defined by equation (1*c*). The quantity α is a pure measure of elastic distortion which is insensitive to changes in total dilatation I_3 or plastic dilatation I_{3p} (i.e., $\mathbf{C}$ may be replaced by $a^2\mathbf{C}$, and $\mathbf{C}_p$ may be replaced by $b^2\mathbf{C}_p$ without changing the value of α). Furthermore, the quantity α attains the value unity when there is no elastic distortion and $\mathbf{C}$ differs from $\mathbf{C}_p$ by only a uniform dilatation ($\mathbf{C}=a^2\mathbf{C}_p$).

The entropy flux $\mathbf{P}$ in (1*d*) is also modified (relative to Rubin, 1986) to be consistent with the assumption of Fourier heat conduction in the present configuration. That is, equation (1*d*) is consistent with

$$\mathbf{p} = \frac{1}{\theta}\mathbf{q} = -\frac{K(I_3,\theta)}{\theta}\mathbf{g}, \quad \mathbf{g} = \frac{\partial\theta}{\partial\mathbf{x}} \tag{2a,b}$$

where $\mathbf{p}$ is the entropy flux and $\mathbf{q}$ is the heat conduction vector, each per unit present area; and $\mathbf{g}$ is the temperature gradient with respect to $\mathbf{x}$. In equations (1*d*) and (2*a*) we have included dependence of the heat conduction coefficient K on the dilatation I_3 for generality in dealing with high compression situations.

By considering general constitutive equations of the type (1) restrictions may be obtained (Rubin, 1986) to ensure consistency with the first law of thermodynamics. In particular, the symmetric Piola-Kirchhoff stress $\mathbf{S}$, the specific entropy η, and the rate of plastic dissipation $\rho_0\theta\xi'$ are related to derivatives of the Helmholtz free energy ψ. These restrictions, together with an expression for the specific internal energy ϵ may be summarized in the form:

$$\mathbf{S} = 2\rho_0\frac{\partial\psi}{\partial\mathbf{C}}, \tag{3a}$$

$$2\rho_0\eta = -2\rho_0\frac{\partial\psi}{\partial\theta} = 2\rho_0\frac{dh}{d\theta} + f_1(I_3) + 2\rho_0\eta', \tag{3b}$$

$$2\rho_0\eta' = -\frac{\partial\hat{\mu}}{\partial\theta}\ell n\alpha, \tag{3c}$$

$$2\rho_0\epsilon = 2\rho_0(\psi+\theta\eta) = 2\rho_0(\epsilon_1+\epsilon'), \tag{3d}$$

$$2\rho_0\epsilon_1(I_3,\theta) = 2\rho_0\left(\theta\frac{dh}{d\theta} - h\right) + \theta_0 f_1(I_3) + f_2(I_3), \tag{3e}$$

$$2\rho_0\epsilon' = \left(\hat{\mu} - \theta\frac{\partial\hat{\mu}}{\partial\theta}\right)\ell n\alpha, \tag{3f}$$

$$\rho_0\theta\xi' = -\rho_0\frac{\partial\psi}{\partial\mathbf{C}_p}\cdot\dot{\mathbf{C}}_p, \tag{3g}$$

where a superposed dot denotes material time differentiation holding $\mathbf{X}$ fixed, and $\mathbf{A}\cdot\mathbf{B}=tr(\mathbf{A}\mathbf{B}^T)$ denotes the inner product of two tensors.

The Cauchy stress $\mathbf{T}$ and the symmetric Piola-Kirchhoff stress $\mathbf{S}$ are related by the expression

$$\mathbf{T} = I_3^{-1/2}\mathbf{F}\mathbf{S}\mathbf{F}^T \tag{4}$$

For many applications, and in particular for metals which are isotropic in their reference configurations, it is desirable to decompose the Cauchy stress $\mathbf{T}$ into a pressure p and a deviatoric part $\mathbf{T}'$. Using equation (4) it follows that both $\mathbf{T}$ and $\mathbf{S}$ admit the unique decompositions

$$\mathbf{T} = -p\mathbf{I} + \mathbf{T}', \qquad \mathbf{T}'\cdot\mathbf{I} = 0, \tag{5a,b}$$

$$\mathbf{S} = -pI_3^{1/2}\mathbf{C}^{-1} + \mathbf{S}', \qquad \mathbf{S}'\cdot\mathbf{C} = 0, \tag{5c,d}$$

$$\mathbf{T}' = I_3^{-1/2}\mathbf{F}\mathbf{S}'\mathbf{F}^T, \tag{5e}$$

where $\mathbf{I}$ is the identity tensor. Note that $\mathbf{S}'$ is the counterpart of $\mathbf{T}'$ even though it is not a deviatoric tensor. An interesting manifestation of the kinematic separation of pure dilatation I_3 from pure elastic distortion α is that the stress $\mathbf{S}$ naturally separates into the form (5*c*). To see this we substitute equation (1*a*) into equation (3*a*) to deduce that

$$p = p_1(I_3,\theta) + p', \tag{6a}$$

$$p_1(I_3,\theta) = \left[(\theta-\theta_0)\frac{df_1}{dI_3} - \frac{df_2}{dI_3}\right]I_3^{1/2}, \tag{6b}$$

$$p' = -I_3^{1/2}\frac{\partial\hat{\mu}}{\partial I_3}\ell n\alpha, \tag{6c}$$

$$\mathbf{S}' = \hat{\mu}\bar{\mathbf{S}}' = \hat{\mu}\left[\left(\frac{3}{\mathbf{C}\cdot\mathbf{C}_p^{-1}}\right)\mathbf{C}_p^{-1} - \mathbf{C}^{-1}\right] \tag{6d}$$

Furthermore, it may be shown that the rate of plastic dissipation may also be written in the form

$$\rho_0\theta\xi' = (\mathbf{C}_p^{-1}\mathbf{C})\mathbf{S}'\cdot\dot{\mathbf{E}}_p, \quad \mathbf{E}_p = \frac{1}{2}(\mathbf{C}_p - \mathbf{I}), \tag{7a,b}$$

where $\mathbf{E}_p$ is the plastic strain.

This constitutive assumption for stress was motivated by the physical notion that a plastic material flows somewhat like a fluid. In this regard, we observe that if the shear modulus $\hat{\mu}$ in equation (6*d*) vanishes then equations (5*a,c*) yield constitutive equations for an ideal fluid ($\mathbf{T}=-p\mathbf{I}$) with $\mathbf{S}$ proportional to $\mathbf{C}^{-1}$. Motivated by this result we sought a form for the kinematic variable α in equation (1*c*) which led to the result (6*d*) with $\mathbf{S}'$ being proportional to $\mathbf{C}^{-1}$ and $\mathbf{C}_p^{-1}$. The quantity $\mathbf{S}'$ and the deviatoric Cauchy stress $\mathbf{T}'$ are controlled by the amount of elastic distortion. Specifically, these quantities vanish when elastic distortion vanishes ($\alpha=1$) and the elastic deformation is characterized by a pure dilatation ($\mathbf{C}=a^2\mathbf{C}_p$).

[1]The prime used here should not be confused with the use of prime in a different context in Rubin (1986).

Furthermore, it is of interest to note that if equation (6*d*) were rewritten in terms of the total strain $\mathbf{E}=(1/2)(\mathbf{C}-\mathbf{I})$ and the plastic strain $\mathbf{E}_p$ (equation 7(*b*)) then $\mathbf{S}'$ cannot be expressed as a function of $\mathbf{E}-\mathbf{E}_p$ only. This means that the quantity $\mathbf{E}-\mathbf{E}_p$ does not have the meaning of elastic strain in the same sense that it does in the linear theory.

Substituting equation (6*d*) into equation (5*e*) we may write

$$\mathbf{T}'=\hat{\mu}\bar{\mathbf{T}}'=\mu\bar{\sigma}', \quad \bar{\mathbf{T}}'=I_3^{-1/3}\bar{\sigma}', \quad \mu=\hat{\mu}I_3^{-1/3}, \qquad (8a,b,c)$$

$$\bar{\sigma}'=\mathbf{F}\bar{\mathbf{S}}'\mathbf{F}^T=\left[\left(\frac{3}{\mathbf{C}\cdot\mathbf{C}_p^{-1}}\right)\mathbf{F}\mathbf{C}_p^{-1}\mathbf{F}^T-\mathbf{I}\right]. \qquad (8d)$$

From equation (8*d*) we observe that $\bar{\sigma}'$ depends only on elastic distortion and is independent of both total and plastic dilatations. Consequently, the effective shear modulus μ in equations (8*a*,*c*) contains the complete dependence of $\mathbf{T}'$ on dilatation. Further, from equation (8*c*) we observe that the effective shear modulus depends on the total dilatation I_3 even if $\hat{\mu}$ is independent of I_3. We emphasize that the results (6) are a direct consequence of the first law of thermodynamics. Therefore, if the shear modulus has a nontrivial dependence on I_3 ($\partial\hat{\mu}/\partial I_3\neq 0$) then the pressure must depend on the elastic distortion through p' in equation (6*c*). In other words, p' is the pressure response to elastic distortion.

In the study of shock waves in solids it has become common to assume that the pressure depends on the dilatation I_3 and temperature θ only, even if the shear modulus has a nontrivial dependence on I_3. In view of the discussion above, this assumption is valid only when p' can be neglected. For these constitutive equations p' is second order in elastic distortional strain. Consequently, for most applications at low, medium, and high strain rates p' is negligible relative to p_1. However, for large elastic distortion which can occur at ultra high strain rates, p' may be significant (see the example in the last section).

Steinberg et al. (1980) have proposed a model for the shear modulus, which depends on pressure, density, and temperature, and a model for melting. They show consistency of their model with certain theoretical limits and certain experimental data which are referenced in their paper. Here, we present different models for the shear modulus and melting which also appear to be consistent with these theoretical limits and experimental data.

In their model Steinberg et al. (1980) propose that the shear modulus depends linearly on pressure and temperature, and nonlinearity on density. Since pressure is usually assumed to depend linearly on internal energy and internal energy is usually assumed to depend linearly on temperature, this is equivalent to assuming that the shear modulus depends linearly on temperature θ and nonlinearly on the dilatation I_3, which is a special case of equations (1*h*) and (8*c*). For our purposes, we also assume that when the temperature reaches the melting temperature $\theta_M(I_3)$ the shear modulus vanishes and the material ceases to support shear stress. Here, we propose the form

$$\hat{\mu}=\mu_0 f_3(I_3)f_4(\bar{\theta}) \quad \text{for } \bar{\theta}>0, \qquad (9a)$$

$$\hat{\mu}=0 \quad \text{for } \bar{\theta}\leqq 0, \qquad (9b)$$

$$\bar{\theta}=\frac{\theta_M(I_3)-\theta}{\theta_{M0}-\theta_0}, \qquad (9c)$$

where the melting temperature may be characterized by (Steinberg et al., 1980)

$$\theta_M(I_3)=\theta_{M0}\exp[2a_1(1-I_3^{1/2})]I_3^{-(\Gamma_0-a_1-1/3)}. \qquad (10)$$

In equation (10), Γ_0 and a_1 are constants related to the Grüneisen gama Γ which is given by

$$\Gamma(I_3)=\Gamma_0-a_1(1-I_3^{1/2}). \qquad (11)$$

To ensure that at melting ($\theta=\theta_M$, $\bar{\theta}=0$) there are no discontinuities in the shear stress, entropy and internal energy, we require

$$f_4(0), \quad \frac{df_4}{d\bar{\theta}}(0)=0. \qquad (12a,b)$$

Consistent with these conditions, we specify

$$f_4=\exp\left[\frac{a_2}{a_3}\left(1-\frac{1}{\bar{\theta}^{a_3}}\right)\right], \qquad (13)$$

where a_2 and a_3 are nonnegative constants. For small ($<<1$) values of a_3 this function f_4 exhibits nearly linear dependence on temperature and drops rather abruptly to zero near melting ($\bar{\theta}=0$), which is consistent with the observed dependence of the shear modulus. The simplest form for the function f_3 in equation (9*a*) which is consistent with the effective shear modulus μ in equation (8*c*) approaching the Thomas-Fermi limit ($\mu\to I_3^{-2/3}$ as $I_3\to 0$) becomes

$$f_3(I_3)=I_3^{-1/6}. \qquad (14)$$

In addition to the restrictions (3) various statements of the second law of thermodynamics must be satisfied. For the specific constitutive equations proposed here, these statements are satisfied provided that[2] (see Rubin, 1986)

$$\theta-\theta_1>0 \text{ whenever } \left[\theta\frac{dh}{d\theta}-h(\theta)\right]$$

$$-\left[\theta_1\frac{dh}{d\theta}(\theta_1)-h(\theta_1)\right]>0, \qquad (15a)$$

$$K(I_3,\theta)\geqq 0, \quad \xi'\geqq 0. \qquad (15b,c)$$

Equation (15*a*) requires the specific heat at constant volume and zero elastic distortion to be positive; (15*b*) requires heat to flow from hot to cold; and (15*c*) requires plastic dissipation to be nonnegative.

From equations (6*d*) and (8*a*,*d*) it is observed that deviatoric stress vanishes when $\mathbf{C}$ is parallel to $\mathbf{C}_p$; but not necessarily equal to $\mathbf{C}_p$, as was required by the equations in Rubin (1986). This distinction is essential when considering metals at high compression. To see this we observe that when a material is severely compressed the dilatation $I_3=\det\mathbf{C}=(\rho_0/\rho)^2$ is appreciably different from one. However, the notion of plastic incompressibility suggests that $I_{3p}=\det\mathbf{C}_p=1$. Consequently, if $\mathbf{C}=\mathbf{C}_p$ then the material cannot be both severely compressed and plastically incompressible. Motivated by this observation we modified the flow rule to take the form:

$$\dot{\mathbf{C}}_p=g\mathbf{A}, \quad \mathbf{A}=\frac{(I_3^{-1/3}\mathbf{C}-\mathbf{C}_p)}{|I_3^{-1/3}\mathbf{C}-\mathbf{C}_p|} \qquad (16a,b)$$

$$g=2\sqrt{2}D_0\exp\left[-\frac{1}{2}\left(\frac{[\bar{Z}R(I_3,\theta)]^2}{3\bar{J}_2}\right)^{n(I_3,\theta)}\right], \qquad (16c)$$

$$J_2=\frac{1}{2}\mathbf{T}'\cdot\mathbf{T}'=\hat{\mu}^2\bar{J}_2=\frac{1}{2}\mu^2\bar{\sigma}'\cdot\bar{\sigma}', \qquad (16d)$$

$$\bar{J}_2=\frac{1}{2}I_3^{-1}\bar{\sigma}'\cdot\bar{\sigma}'=\frac{1}{2}I_3^{-1}(\mathbf{C}\bar{\mathbf{S}}'\cdot\bar{\mathbf{S}}'\mathbf{C}) \qquad (16e)$$

$$\bar{Z}=\bar{\kappa}+\bar{\beta}, \quad \bar{\beta}=\bar{\boldsymbol{\beta}}\cdot\mathbf{U}, \qquad (16f,g)$$

$$\mathbf{U}=\frac{\mathbf{S}}{|\mathbf{S}|}. \qquad (16h)$$

In equations (16) D_0 is a constant; $\bar{Z}$ is a scalar measure of hardening which is additively separated into isotropic hardening $\bar{\kappa}$ and a scalar measure of directional hardening $\bar{\beta}$; $\bar{\boldsymbol{\beta}}$ is a tensor measure of directional hardening (which models the

[2] In obtaining the restriction (15*a*) we have interpreted thermal equilibrium to be a state with zero elastic distortion, so that $\alpha=1$ and plastic deformation rate vanishes.

Bauschinger effect); and $|\mathbf{A}| = (\mathbf{A}\cdot\mathbf{A})^{1/2}$ denotes the magnitude of the tensor $\mathbf{A}$. Further, the quantities $\bar{Z}$, $\bar{\kappa}$, $\bar{\beta}$, $\bar{\beta}$ are nondimensional quantities which should not be confused with similar unbarred quantities used in Rubin (1986).

The modified flow rule (16*a*) states that plastic deformation $\mathbf{C}_p$ evolves in the direction of total deformation $\mathbf{C}$. This ensures that plastic deformation evolves towards a value which will make the deviatoric stress $\mathbf{T}'$ in equations (8) vanish. In equation (16*a*) plastic incompressibility is not enforced in a strong sense by requiring $\dot{I}_{3p}$ to vanish as was done in Rubin (1986). Instead, plastic incompressibility is enforced in a weak sense by requiring plastic deformation to evolve towards a tensor whose determinant is unity [$\det(I_3^{-1/3}\mathbf{C}) = 1$]. In general, this allows small deviations of I_{3p} from unity. Further, it is worth mentioning that it is possible to model plastic compressibility of geological materials by replacing $I_3^{-1/3}$ in equation (16*b*) with a function of I_3 and I_{3p}.

For definiteness, let the reference configuration be a stress free configuration which is specified by

$$\mathbf{x} = \mathbf{X}, \quad \mathbf{C} = \mathbf{I}, \quad \mathbf{C}_p = \mathbf{I}, \quad \theta = \theta_0, \tag{17a,b,c,d}$$

$$\bar{\kappa} = \bar{Z}_2, \quad \bar{\beta} = 0, \tag{17e,f}$$

$$\psi = 0, \quad \eta = 0, \quad \mathbf{S} = 0. \tag{17g,h,i}$$

It follows from equations (1*a*,*b*,*c*), (3*a*–*c*), (6), and (17) that we may specify

$$f_1(1) = 0, \quad f_2(1) = 0, \quad \frac{df_2}{dI_3}(1) = 0, \tag{18a,b,c}$$

$$h(\theta_0) = 0, \quad \frac{dh}{d\theta}(\theta_0) = 0. \tag{18d,e}$$

Using equations (1*c*) and (6*b*–*d*), we deduce that the stress-free configuration attainable after an arbitrary plastic deformation is characterized by

$$\theta = \theta(I_3) = \theta_0 + \frac{df_2}{dI_3}\Big/\frac{df_1}{dI_3}, \quad \mathbf{C} = \gamma\mathbf{C}_p, \tag{19a,b}$$

where γ is an arbitrary scalar. We emphasize that unloading from a given plastic state may require additional plastic deformation to attain equation (19*b*) since, unlike the total deformation $\mathbf{C}$, plastic deformation $\mathbf{C}_p$ is not necessarily derivable from a displacement field and hence is not necessarily associated with an attainable configuration.

The function g in equations (16*a*,*c*) is consistent with a kinetic equation of the form

$$|\dot{\mathbf{C}}_p| = g(I_3, \theta, \bar{J}_2, \bar{Z}) \tag{20}$$

and causes yield-like behavior in the sense that $\dot{\mathbf{C}}_p$ is vanishingly small for small values of $\bar{J}_2$ and increases rapidly when $\bar{J}_2$ attains a value of the order of $(\bar{Z}R)^2$. The quantity D_0 in equation (16*c*) corresponds to the maximum value of plastic shear strain rate (say, $\dot{E}^p_{12} = \frac{1}{2}\dot{C}^p_{12}$). This requirement that $|\dot{\mathbf{C}}_p|$ be bounded from above has particular significance at ultra high strain rates which are much larger than D_0 since the predicted material response is nearly elastic (see the example in the last section). An important modification of g in equation (16*c*) relative to g in Rubin (1986) is that Z in Rubin (1986) has been replaced by $\hat{\mu}\bar{Z}R$. The effect of this modification is that at melting $\bar{J}_2$ does not necessarily vanish even though J_2 does. This means that at melting plastic deformation will continue to evolve towards total deformation so that the stress free configuration (apart from density changes) will never be far from the present configuration. Furthermore, in equation (16*c*), the quantity $\bar{Z}R$ controls the normalized shear strength of the material and n controls the strain-rate sensitivity of the material. For high compression applications it is possible that n and R will be functions of I_3 and θ instead of n being a function of θ only and R being unity as was taken in Rubin (1986). The dependence of R and n on I_3 effectively introduces a dependence of plastic flow on pressure. In this regard, we note that even if R and n in equation (16*c*) are independent of the total dilatation I_3 and $\bar{J}_2$ is relatively insensitive to changes in I_3, the effective flow stress $(3J_2)^{1/2}$ will depend on the dilatation I_3 through the dependence of the shear modulus $\hat{\mu}$ (see equation (16*d*)).

The hardening variables $\bar{\kappa}$ and $\bar{\beta}$ are determined by evolution equations of the form:

$$\dot{\bar{\kappa}} = m_1(\rho_0\theta\xi')(\bar{Z}_1 - \bar{\kappa}) - A_1(\theta)\bar{Z}_1\left(\frac{\bar{\kappa} - \bar{Z}_2}{\bar{Z}_1}\right)^{r_1}, \tag{21a}$$

$$\dot{\bar{\beta}} = m_2(\rho_0\theta\xi')(\bar{Z}_3\mathbf{U} - \bar{\beta}) - A_2(\theta)\bar{Z}_1\left(\frac{|\bar{\beta}|}{\bar{Z}_1}\right)^{r_2}\mathbf{V}, \tag{21b}$$

$$\mathbf{V} = \frac{\bar{\beta}}{|\bar{\beta}|}. \tag{21c}$$

The first terms in these evolution equations represent hardening and the second terms represent thermal recovery of hardening. In equations (21) m_1 and m_2 are constants determining the rate of hardening; $\bar{Z}_1$, $\bar{Z}_3$, and $\bar{Z}_2$ are nondimensional constants representing the saturation values of $\bar{\kappa}$ and $\bar{\beta}$, and the annealed value of $\bar{\kappa}$, respectively; and the constants r_1, r_2, and functions $A_1(\theta)$, $A_2(\theta)$ control the rate of thermal recovery. The constants $\bar{Z}_1$, $\bar{Z}_2$, $\bar{Z}_3$ together with the reference value $\bar{\kappa}_0$ of $\bar{\kappa}$ are normalized values of Z_1, Z_2, Z_3, κ_0 in Rubin (1986) and are defined by

$$\bar{Z}_1 = \frac{Z_1}{\mu_0}, \quad \bar{Z}_2 = \frac{Z_2}{\mu_0}, \quad \bar{Z}_3 = \frac{Z_3}{\mu_0}, \quad \bar{\kappa}_0 = \frac{\kappa_0}{\mu_0}, \tag{22a,b,c,d}$$

where μ_0 is the reference value of the shear modulus $\hat{\mu}$. When the material melts the quantity $\rho_0\theta\xi'$, which represents the rate of plastic dissipation, vanishes and hardening ceases.

In the following, we briefly discuss how the nonlinear theory presented above may be reduced to one consistent with the usual linear theory. For the linearized theory quadratic terms in the quantities $\mathbf{E}$, $\mathbf{E}_p$ and $(\theta - \theta_0)$ are neglected. With this approximation it can be shown that equations (1), (3), (6), (7), (16), and (21) reduce to the usual linearized equations and are equivalent to those in Bodner (1985) provided that

$$\frac{df_1}{dI_3}(1) = 3k_0\alpha_0, \quad \frac{d^2f_2}{dI_3^2}(1) = \frac{k_0}{2}, \tag{23a,b}$$

$$\frac{d^2h}{d\theta^2}(\theta_0) = \frac{C_v}{\theta_0}, \tag{23c}$$

$$\hat{\mu}(1, \theta_0) = \mu_0, \quad n = n(\theta), \quad R = 1, \tag{23d,e,f}$$

where k_0, α_0, C_v are the reference values of the isothermal Bulk modulus, the coefficient of linear thermal expansion and the specific heat at constant volume, respectively. The condition (23*d*) is satisfied by the specifications (9)–(14).

Finally, we note that unlike the linearized form of plastic dissipation (7*a*), the nonlinear form of (7*a*) is different from the usual expression for rate of plastic work ($\mathbf{S}\cdot\dot{\mathbf{E}}_p$). We emphasize that in most developments the form for plastic dissipation is assumed, whereas here the expression (7*a*) is a direct consequence of the first law of thermodynamics. At present it does not seem possible to prove that (7*a*) satisfies the restriction (15*c*) for all thermodynamic processes. Consequently, we suggest monitoring the value of ξ' and limiting any calculation to that range for which ξ' remains nonnegative continuously. We note, however, that for the linearized theory ξ' is nonnegative for all processes.

Determination of Certain Constitutive Functions

In this section we show how the constitutive functions h, f_1, f_2 in equation (1*a*) may be estimated by comparison with a Mie-Grüneisen equation of state of the form

$$p_1 - p_H(I_3) = \rho_0 I_3^{-1/2}\Gamma(I_3)[\epsilon_1 - \epsilon_H(I_3)], \tag{24}$$

which relates those portions p_1 and ϵ_1 of the pressure and internal energy which depend only on the dilatation I_3 and temperature θ. In equation (24), p_H, ϵ_H, Γ are functions of I_3 only; p_H and ϵ_H are determined by Hugoniot shock-wave data using the fluid approximation which neglects deviatoric stress; and Γ is the Grüneisen gamma.

To determine the function h in equation (1a) we observe from equations (3d,e) that the specific heat C_v at constant deformation $\mathbf{C}$ and zero elastic distortion ($\alpha = 1$) is given by

$$\frac{\partial \epsilon}{\partial \theta}\bigg|_{\mathbf{C},\alpha=1} = \theta \frac{d^2 h}{d\theta^2} = C_v(\theta). \tag{25}$$

Integration of (25) subject to the conditions (18d,e) yields

$$h(\theta) = \int_{\theta_0}^{\theta} \int_{\theta_0}^{\lambda} \frac{C_v(T)}{T} dT d\lambda, \tag{26}$$

which shows that h is determined once the specific heat is measured. When the specific heat is constant, equations (26) and (3b,e) yield

$$h = C_v\left[\theta \ell n\left(\frac{\theta}{\theta_0}\right) - (\theta - \theta_0)\right], \tag{27a}$$

$$2\rho_0 \eta = 2\rho_0 C_v \ell n\left(\frac{\theta}{\theta_0}\right) + f_1(I_3) + 2\rho_0 \eta', \tag{27b}$$

$$2\rho_0 \epsilon_1 = 2\rho_0 C_v(\theta - \theta_0) + \theta_0 f_1(I_3) + f_2(I_3). \tag{27c}$$

Since the specific heat of a solid is nearly constant we specify h, η, ϵ_1 by equations (27) and note that this specification is different from that in Rubin (1986).

To determine the functions f_1, f_2 we solve equation (27c) for $\theta - \theta_0$ and rewrite (6b) in the form

$$p_1 = \left(\frac{I_3^{1/2}}{C_v}\frac{df_1}{dI_3}\right)\epsilon_1 - I_3^{1/2}\left[\frac{df_2}{dI_3} + \frac{1}{2\rho_0 C_v}(\theta_0 f_1 + f_2)\frac{df_1}{dI_3}\right]. \tag{28}$$

Comparison of equation (28) with equation (24) yields two differential equations of the forms

$$\frac{df_1}{dI_3} = \rho_0 C_v I_3^{-1} \Gamma(I_3), \tag{29a}$$

$$\frac{df_2}{dI_3} + \frac{\Gamma}{2I_3} f_2 = G(I_3), \tag{29b}$$

$$G(I_3) = -\left(\frac{\theta_0 \Gamma}{2I_3}\right) f_1 + I_3^{-1/2}(\rho_0 I_3^{-1/2} \Gamma \epsilon_H - p_H), \tag{29c}$$

to determine the functions f_1 and f_2, subject to the conditions (18a, b). The condition (18c) is satisfied because both $\epsilon_h(1)$ and $p_H(1)$ vanish. Using the expression (11) for the Grüneisen gamma, equations (29a,b) may be integrated to obtain

$$f_1(I_3) = \rho_0 C_v f(I_3), \tag{30a}$$

$$f(I_3) = \Gamma_0 \ell n I_3 - a_1[\ell n I_3 + 2\phi], \tag{30b}$$

$$\phi = 1 - I_3^{1/2}, \tag{30c}$$

$$f_2(I_3) = -e^{-\frac{1}{2}f(I_3)} \int_{I_3}^{1} G(\lambda) e^{\frac{1}{2}f(\lambda)} d\lambda. \tag{30d}$$

Finally, we assume that for a planar shock the shock velocity U is a linear function of the particle velocity u, such that

$$U = C_0 + s_1 u, \tag{31}$$

where C_0 is the low pressure wave speed and s_1 is a constant. Then, using the fluid approximation (deviatoric stress is neglected) the jump conditions yield

$$p_H = \frac{\rho_0 C_0^2 \phi}{(1 - s_1\phi)^2}, \quad \epsilon_H = \frac{C_0^2 \phi^2}{2(1 - s_1\phi)^2}. \tag{32a,b}$$

Table 1 Uniaxial strain data for the compression $\phi = 0.5$ ($\theta_0 = 300$ K, $\theta_{M0} = 1220$ K, $\mu_0 = 27.6$ GPa)

	$a = 10^4\ s^{-1}$	$a = 10^8\ s^{-1}$	$a = 10^{12}\ s^{-1}$
t_{11} (GPa)	-222	-227	-306
t'_{11} (GPa)	-0.432	-3.95	-67.0
p (GPa)	222	223	239
p' (GPa)	0.529×10^{-3}	0.0461	19.0
θ (K)	1060	1270	730
θ_M (K)	6610	6610	6610
$\alpha - 1$	14.2×10^{-6}	1.22×10^{-3}	0.687
$I_{3p}^{\frac{1}{2}} - 1$	8.14×10^{-6}	0.487×10^{-3}	17.8×10^{-6}
μ (GPa)	99.3	98.6	100

Uniaxial Strain

In this example we calculate the stresses and temperature produced by homogeneous uniaxial strain at high (10^4, 10^8) and ultra high (10^{12}) compression rates. Referring the position vectors $\mathbf{x}$ and $\mathbf{X}$ to the base vectors $\mathbf{e}_i$ ($i = 1,2,3$) of a fixed Cartesian coordinate system we take

$$x_1 = (1 - at)X_1, \quad x_2 = X_2, \quad x_3 = X_3, \tag{33a,b,c}$$

$$a = 10^4 s^{-1}, \quad 10^8 s^{-1}, \quad 10^{12} s^{-1}, \tag{33d}$$

where a is the constant rate of loading.

Since the loading is very rapid we assume that the process is adiabatic and neglect heat conduction and thermal recovery of hardening so that

$$K = 0, \quad A_1 = 0, \quad A_2 = 0. \tag{34a,b,c}$$

It follows that the temperature may be calculated by the balance of entropy equation which reduces to

$$\rho_0 \dot{\eta} = \rho_0 \xi'. \tag{35}$$

Here, we consider typical material properties of Aluminum which are specified by

$$\mu_0 = 27.6\ \text{GPa}, \quad \rho_0 = 2.71\ \text{Mg/m}^3, \quad C_0 = 5.38\ \text{km/s}, \tag{36a,b,c}$$

$$\Gamma_0 = 1.97, \quad a_1 = 1.5, \quad a_2 = 0.2, \quad a_3 = 0.01, \tag{36d,e,f,g}$$

$$s_1 = 1.35, \quad C_v = 0.862\ \text{J/gK}, \tag{36h,i}$$

$$\theta_0 = 300\ \text{K}, \quad \theta_{M0} = 1220\ \text{K}, \tag{36j,k}$$

$$D_0 = 10^8 \text{s}^{-1}, \quad \bar{\kappa}_0 = 1.63 \times 10^{-2}, \quad \bar{Z}_1 = 3.62 \times 10^{-2}, \tag{36ℓ,m,n}$$

$$m_1 = 220(\text{GPa})^{-1}, \quad m_2 = 0, \quad \bar{Z}_3 = 0, \tag{36o,p,q}$$

$$R = 1, \quad n = n_0\left(\frac{\theta_0}{\theta}\right), \quad n_0 = 5.0. \tag{36r,s,t}$$

Since no set of consistent data is available these values have been collected from different sources and are representative of different aluminum alloys. Specifically, from Bodner (1984) we took the values of D_0, n_0 (Table 3, p. 49), and m_1 (Table 1, p. 47). The values of μ_0, Γ_0, a_1, a_2, a_3, θ_0, θ_{M0}, $\bar{\kappa}_0$, $\bar{Z}_1$ are consistent with the data in (Steinberg et al., 1980, p. 1499) with a_2,

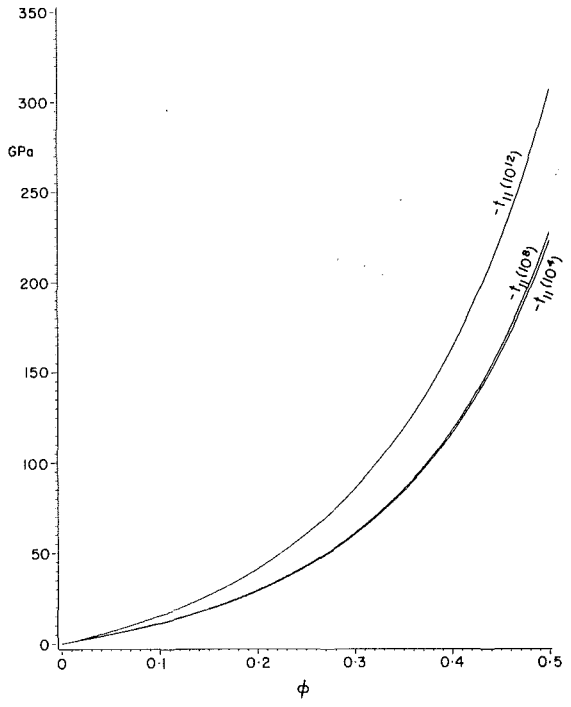


Fig. 1 Cauchy stress t_{11} versus compression ϕ for the compression rates $a = 10^4, 10^8, 10^{12}\ s^{-1}$

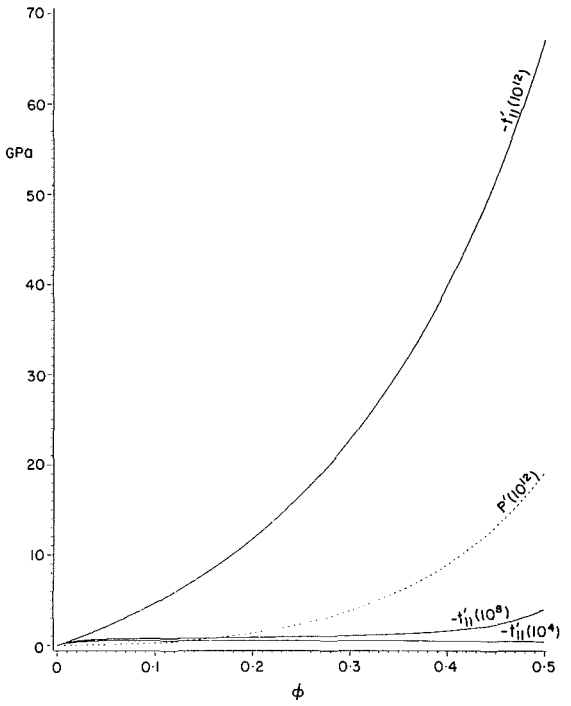


Fig. 2 Deviatoric Cauchy stress t'_{11} and pressure p' versus the compression ϕ for the compression rates $a = 10^4, 10^8, 10^{12}\ s^{-1}$

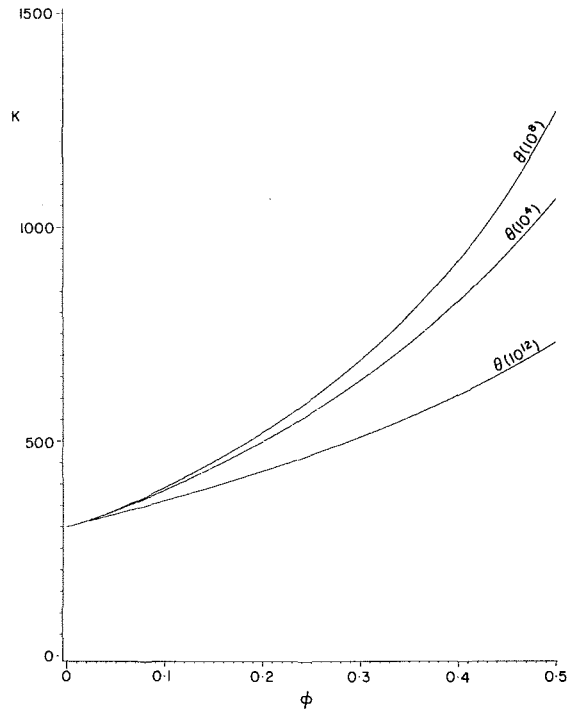


Fig. 3 Temperature θ versus the compression ϕ for the compression rates $a = 10^4, 10^8, 10^{12}\ s^{-1}$

a_3 characterizing the temperature dependence of the shear modulus. The values $\bar{\kappa}_0$ and $\bar{Z}_1$ predict the appropriate initial and saturation flow stresses in uniaxial tension at a strain rate of about $10^{-4}\ s^{-1}$. The values of ρ_0, C_0, s_1 were taken from Pugh (1970, p. 692) and the value of C_v was taken from Carslaw and Jaeger (1973, p. 497). Furthermore, directional hardening was neglected (m_2, $\bar{Z}_3$ vanish), the value of R was taken to be unity for simplicity, and the functional dependence of n on θ was chosen to be consistent with observations that the strain rate sensitivity increases with increased temperature. The remaining constants are not specified because they are not needed for the calculation.

For this process we have neglected directional hardening ($\bar{\beta} = 0$) so the flow rule (16*a*), the evolution equation (21*a*) and the balance of entropy (35) reduce to four equations to determine the quantities C^p_{11}, C^p_{22} ($C^p_{33} = C^p_{22}$), $\bar{\kappa}$ and θ. These equations were integrated numerically subject to the initial conditions

$$C^p_{11} = C^p_{22} = 1, \quad \bar{\kappa} = \bar{\kappa}_0, \quad \theta = \theta_0. \qquad (37a,b,c)$$

Table 1 summarizes the values of various parameters at the end of the high compression ($\phi = 1 - I_3^{1/2} = 0.5$) for three values of the compression rate a. In this table t_{11} and t'_{11} are the components of the Cauchy stress and Cauchy deviatoric stress, respectively, in the $\mathbf{e}_1$ direction, $\alpha - 1$ is a measure of the change in elastic distortion, and $I_{3p}^{1/2} - 1$ is a measure of plastic volume change. From this table we observe that relative to the values of t_{11} and p, the values of t'_{11} and p' are insignificant for $a = 10^4 s^{-1}$, are minimally significant for $a = 10^8\ s^{-1}$ and are quite significant for $a = 10^{12} s^{-1}$. For the high compression rates (10^4, 10^8) plasticity is important whereas for the ultra high rate (10^{12}) the material response is essentially elastic. This is because the ultra high compression rate is four orders of magnitude larger than the maximum value of plastic strain rate. Consequently, even though the plastic strain rate is quite large ($\sim 10^8$) the compression time is too short for $\mathbf{C}_p$ to evolve appreciably towards $I_3^{-1/3}\mathbf{C}$. This is manifested in the large value of the change in the elastic distortion $\alpha - 1$ and the low value of the temperature θ. The values for θ are higher for the high rates (10^4, 10^8) then for the ultra high rate (10^{12}) since at the high rates there is sufficient time for plastic dissipation to increase the material temperature. It is also worth noting that the plastic volume change ($I_{3p}^{1/2} - 1$) remains small for all compression rates. Furthermore, in each of these calculations the plastic dissipation $\rho_0 \theta \xi'$ was monitored and found to be non-negative for the entire process, which is consistent with the restriction (15*c*).

Figures 1, 2, 3 show graphically the values of t_{11}, t'_{11}, p', θ for values of the compression ϕ from 0 to 0.5 and for three values of the compression rate a. The values of p' for the high compression rates (10^4, 10^8) are not plotted in Fig. 2, because they are too small.

In typical plate impact experiments (see Clifton, 1983; or Steinberg et al., 1980) the particle velocity increases rapidly from zero to a constant value during loading. It follows that compression rate increases from zero to a maximum and back to zero during this loading. Since our calculations are for constant compression rate they qualitatively correspond to an average loading rate. More specifically, the viscoplastic behavior of our model would predict a relaxation of the deviatoric stress as the compression rate reduces to zero but this is not exhibited in our calculation because the compression rate is constant.

In summary, we recall that when the shear modulus has a nontrivial dependence on the dilatation I_3 then the pressure is not a function of I_3 and the temperature θ only. Specifically, the pressure also includes a dependence on the elastic distortion through the function p' in equation (6*c*). When the compression rate is smaller than the maximum value of plastic strain rate ($\sim D_0$) then plasticity limits the increase of elastic distortion and both the deviatoric stress t'_{11} and the pressure p' are small relative to the value of t_{11} at large compression. However, for ultra high compression rates ($> D_0$) the material response is essentially elastic and both the deviatoric stress t'_{11} and the pressure p' become significant at large compression since the elastic distortion is also large. We emphasize that the numerical results presented here are qualitative in the sense that the functional dependence of ψ' in equation (1*b*) on α, the functional dependence of R and n in equation (16*c*) on I_3 and θ, and the value of D_0 in equation (16*c*), are not known for aluminum at ultra high compression rates.

Finally, the results of this paper suggest that, in order to interpret data from shock-wave experiments which typically span the full range of compression rates from zero to ultra high, it is essential to determine ψ', R, n, and D_0.

Acknowledgment

The author would like to thank Drs. W. C. Moss and L. Glenn for helpful discussions.

References

Bodner, S. R., 1984, "Review of a Unified Elastic-Viscoplastic Theory (The Bodner Equations)," Interim Scientific Report prepared for the US Air Force Office of Scientific Research and the European Office of Aerospace Research and Development, to be published as a chapter in the book: *Unified Constitutive Equations for Plastic Deformation and Creep of Engineering Alloys*. Miller, A. K., ed., Elsevier Applied Science, London.

Bodner, S. R., 1985, "Evolution Equations for Anisotropic Hardening and Damage of Elastic-Viscoplastic Materials," *Plasticity Today—Modelling, Methods and Applications*, Sawczuk, A., and Bianchi, G., eds., Elsevier, London, pp. 471–482.

Bodner, S. R., and Partom, Y., 1972, "A Large Deformation Elastic-Viscoplastic Analysis of a Thick-Walled Spherical Shell," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 751–757.

Carslaw, H. S., and Jaeger, J. C., 1973, *Conduction of Heat in Solids*, Oxford University Press.

Clifton, R. J., 1983, "Dynamic Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 941–952.

Green, A. E., and Naghdi, P. M., 1977, "On Thermodynamics and the Nature of the Second Law," *Proc. Royal Soc. Lond.*, Vol. A 357, pp. 253–270.

Green, A. E., and Naghdi, P. M., 1978, "The Second Law of Thermodynamics and Cyclic Processes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 487–492.

Pugh, H. L. D., 1970, *Mechanical Behaviour of Materials Under Pressure*, Elsevier, Amsterdam.

Reed, K. W., and Atluri, S. N., 1985, "Constitutive Modelling and Stress Analysis for Finite Deformation Inelasticity," *Constitutive Equations: Macro and Computational Aspects*, William, K. J., ed., pp. 111–129, and presented at the Winter Annual Meeting of ASME, New Orleans, LA, December 9–14, 1984.

Rubin, M. B., 1986, "An Elastic-Viscoplastic Model for Large Deformation," *Int. J. Engng. Sci.*, Vol. 24, pp. 1083–1095.

Steinberg, D. J., Cochran, S. G., and Guinan, M. W., 1980, "A Constitutive Model for Metals Applicable at High-Strain Rates," *J. Appl. Phys.*, Vol. 51, No. 3, pp. 1498–1504.

Incompressibility in Dynamic Relaxation

S. A. Silling
Brown University,
Division of Engineering,
Providence, R.I. 02912
Assoc. Mem. ASME

A method is described for enforcing the incompressibility constraint in large-deformation solid mechanics computations using dynamic relaxation. The method is well-suited to explicit time-integration schemes because it does not require the solution of a system of linear equations. It is based on an analogy with thermoelasticity involving manipulation of the natural state of a solid.

1 Introduction

The imposition of kinematical constraints is difficult when numerical methods relying on explicit time-integration are used. The difficulty arises because of the inherently local nature of the algorithm. The motion of each node in each time step is computed solely from known data at nodes immediately surrounding it in the previous time step. This property of explicit time-integration greatly simplifies the coding of such a method, since it is unnecessary to solve large sets of linear algebraic equations. However, the imposition of global constraints becomes more difficult, since a given node does not receive information about distant nodes which could affect its own motion through the constraint.

An example of such a constraint is that of incompressibility in the deformation of a solid. Various means of enforcing incompressibility are available for implicit methods and other methods (see Needleman, 1978, for a summary of these approaches for finite elements). The penalty method is probably the most common approach. There are also algorithms for finite elements involving a variational formulation requiring pressure to be regarded as an additional degree of freedom. One approach to modeling incompressible bodies with an explicit integration scheme is to use the constitutive relation of a "slightly compressible" material, a material that is very stiff with respect to volume changes. Unfortunately, the use of such a slightly compressible constitutive law causes the stable time step to be controlled by high-speed dilatational waves which are an artifact of the compressibility. Therefore the slightly compressible approach is inefficient with explicit integration. All of the above methods have both advantages and disadvantages which will not be dealt with further here. In general, the available numerical methods for incompressible solid mechanics introduce either considerable added complexity over the compressible case or require much additional computer time.

The purpose of this paper is to present a simple method for enforcing the incompressibility constraint in a dynamic relaxation method which uses a Lagrangian finite-difference formulation (Silling, 1985). The method is based on an analogy with thermoelasticity in that it involves manipulation of the natural state of the material locally in each zone. It is therefore called the *pseudotemperature* method. The primary advantages of the method are that it allows the main benefits of the dynamic relaxation approach to be retained, namely simplicity, reliability, and flexibility with regard to constitutive relations. A numerical stability analysis described below shows that the pseudotemperature method does not unduly limit the stable time step, which for the compressible case is controlled by the Courant condition. Implementation of the method requires only minor modification of a computer program designed for compressible materials. Another advantage of the pseudotemperature method is that the incompressibility condition is satisfied with increasing accuracy as a run progresses, and would be satisfied exactly in the limit of an infinite number of time steps. This is not true of the penalty method or of slightly compressible models. While the present application is for a finite-difference code, extension of the method to finite-element programs using dynamic relaxation appears straightforward.

The primary disadvantage of the pseudotemperature method is that some viscous damping of the kind used in dynamic relaxation is necessary for numerical stability. Therefore, extension to truly dynamic problems is not currently possible. The method has found applications in a variety of problems in rubber elasticity, soil mechanics, and the study of phase changes in solids. Some of these results are described in Section 4 below.

2 Basic Numerical Method

This section describes the basic computer program for compressible materials which has been extended according to the techniques described in the next section to model incompressibility. The basic computer program is CHIMP (Silling, 1985), a method for plane-strain compressible finite elasticity. The program uses the Green's theorem differencing method, which is widely used in codes such as HEMP (Wilkins, 1969). The differencing method is summarized below. A detailed discussion has been provided by Herrmann and Bertholf (1983). CHIMP uses the first Piola-Kirchhoff stress and performs all differencing in the reference configuration. In this respect it differs from HEMP and related codes, which

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 15, 1987; final revision April 21, 1987.

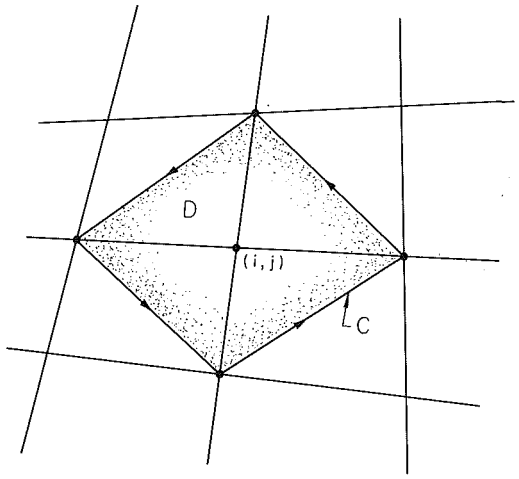


Fig. 2.1 Integration path used in Green's theorem method for differencing stress gradient components

generally use the Cauchy stress tensor and perform differencing in the current configuration. The CHIMP method is more convenient for purposes of finite elasticity, although the HEMP method is more convenient for incremental plasticity. The applicability of the pseudotemperature method is not at all dependent on this choice.

The region is discretized into a mesh with nodes indexed by (i, j). Time steps are labeled by the index n, with $n = 0$ representing the initial condition. Each node has associated with it a fixed position vector $\mathbf{x}^{i,j}$ in the reference configuration and a time-dependent position vector $\mathbf{y}^{i,j,n}$ in the current configuration. The quadrilateral region between each set of four adjacent nodes is a *zone*, labeled by $(i + \frac{1}{2}, j + \frac{1}{2})$. Node-centered quantities, in addition to $\mathbf{x}$ and $\mathbf{y}$, include the velocity vector $\mathbf{v}$ and the acceleration vector $\mathbf{a}$. Zone-centered quantities include the first Piola-Kirchhoff stress tensor $\boldsymbol{\sigma}$ and the deformation gradient tensor $\mathbf{F}$.

Denote the components of $\boldsymbol{\sigma}$ in the plane by $\sigma_{\alpha\beta}$. (Greek subscripts have the range 1, 2.) The Green's theorem differencing method leads to the following approximation for the α component of $\nabla_{\mathbf{x}} \cdot \boldsymbol{\sigma}$ at node (i, j) at time step n:

$$\left(\frac{\partial \sigma_{\alpha\beta}}{\partial x_\beta}\right)^{i,j,n} = \frac{\epsilon_{\beta\gamma}}{2A^{i,j}}\Big\{\sigma_{\alpha\beta}^{i+\frac{1}{2},j+\frac{1}{2},n}(x_\gamma^{i,j+1} - x_\gamma^{i+1,j}) + \sigma_{\alpha\beta}^{i-\frac{1}{2},j+\frac{1}{2},n}(x_\gamma^{i-1,j} - x_\gamma^{i,j+1}) + \sigma_{\alpha\beta}^{i-\frac{1}{2},j-\frac{1}{2},n}(x_\gamma^{i,j-1} - x_\gamma^{i-1,j}) + \sigma_{\alpha\beta}^{i+\frac{1}{2},j-\frac{1}{2},n}(x_\gamma^{i+1,j} - x_\gamma^{i,j-1})\Big\} \tag{1}$$

where $\epsilon_{\beta\gamma}$ is the two-dimensional alternator symbol defined by $\epsilon_{12} = -\epsilon_{21} = 1$, $\epsilon_{11} = \epsilon_{22} = 0$; and $A^{i,j}$ is the equivalent node area, equal to one-half the area of the shaded quadrilateral in Fig. 2.1. For a node lying on a boundary of the mesh, the areas and stress components of the missing zones are set to zero in the above difference formula. See the discussion by Herrmann and Bertholf (1983) for a more detailed analysis of the boundary conditions.

The difference formula (1) is obtained by one-point integration of the following form of Green's theorem applied to the quadrilateral region D enclosed by the curve C as shown in Fig. 2.1:

$$\int_D \frac{\partial \phi}{\partial x_\beta}\, dA = \epsilon_{\beta\gamma} \oint_C \phi\, dx_\gamma \tag{2}$$

where ϕ is any sufficiently smooth field on D, in this case the components of $\boldsymbol{\sigma}$.

An elastic constitutive relation of the form

$$\boldsymbol{\sigma} = \mathbf{h}(\mathbf{F}) \tag{3}$$

is assumed, where $\mathbf{h}$ is a tensor-valued stress response function and $\mathbf{F}$ is the deformation gradient tensor, defined by

$$F_{\alpha\beta} = \frac{\partial y_\alpha}{\partial x_\beta}. \tag{4}$$

These components are evaluated by the following difference formula:

$$\left(\frac{\partial y_\alpha}{\partial x_\beta}\right)^{i+\frac{1}{2},j+\frac{1}{2},n} = \frac{\epsilon_{\beta\gamma}}{2A^{i+\frac{1}{2},j+\frac{1}{2}}}\Big\{(x_\gamma^{i,j+1} - x_\gamma^{i+1,j})(y_\alpha^{i+1,j+1,n} - y_\alpha^{i,j,n}) - (x_\gamma^{i+1,j+1} - x_\gamma^{i,j})(y_\alpha^{i,j+1,n} - y_\alpha^{i+1,j,n})\Big\}. \tag{5}$$

The origin of equation (5) is similar to that of equation (1), but with the four zone edges taken as the contour for Green's theorem. The integrand is the mean value of y_α along a zone edge.

Dynamic relaxation is a standard method which involves the modeling of an equilibrium boundary-value problem as the large-time limit of a damped dynamic problem. This method was first introduced by Day (1965). A thorough discussion of this method has been provided by Underwood (1983), and the following summary is partly based on this source. A discretized version of the equation of motion is most conveniently represented as a form of Newton's second law:

$$\mathbf{f}_{\text{int}}^{i,j,n} + \mathbf{f}_{\text{body}}^{i,j,n} + \mathbf{f}_{\text{bdry}}^{i,j,n} = m_0^{i,j}\mathbf{a}^{i,j,n} + \beta m_0^{i,j}\mathbf{v}^{i,j,n+\frac{1}{2}}, \tag{6}$$

where β is a damping constant, $m_0^{i,j}$ is the nodal mass, $\mathbf{v}^{i,j,n+\frac{1}{2}}$ is the velocity vector, $\mathbf{a}^{i,j,n}$ is the acceleration vector, and $\mathbf{f}_{\text{int}}^{i,j,n}$, $\mathbf{f}_{\text{body}}^{i,j,n}$, and $\mathbf{f}_{\text{bdry}}^{i,j,n}$ are the nodal forces due to internal stress gradients, body forces, and boundary loads, respectively. The force on the node from internal stresses is given by

$$(\mathbf{f}_{\text{int}}^{i,j,n})_\alpha = A^{i,j}\left(\frac{\partial \sigma_{\alpha\beta}}{\partial x_\beta}\right)^{i,j,n} \tag{7}$$

which is evaluated using the difference formulas (1). The quantities $\mathbf{f}_{\text{body}}^{i,j,n}$ and $\mathbf{f}_{\text{bdry}}^{i,j,n}$ are found by multiplying the body force density field or boundary traction field by the nodal area or length.

The difference approximations used in equation (6) for acceleration and velocity are the following:

$$\mathbf{a}^{i,j,n} = (\mathbf{v}^{i,j,n+\frac{1}{2}} - \mathbf{v}^{i,j,n-\frac{1}{2}})/h^n \tag{8}$$

and

$$\mathbf{v}^{i,j,n+\frac{1}{2}} = (\mathbf{y}^{i,j,n+1} - \mathbf{y}^{i,j,n})/h^{n+\frac{1}{2}} \tag{9}$$

where h^n and $h^{n+\frac{1}{2}}$ are the time steps, related by

$$h^n = \frac{1}{2}(h^{n-\frac{1}{2}} + h^{n+\frac{1}{2}}). \tag{10}$$

The discretized equation of motion (6) is solved for $\mathbf{y}^{i,j,n+1}$ for each node using equations (8) and (9). The difference representation is an explicit differencing method, since $\mathbf{y}^{i,j,n+1}$ depends only on quantities which are known from time step n.

The difference formula (6) differs slightly from the usual one for dynamic relaxation in that the damping term uses the velocity value centered at $n + \frac{1}{2}$ rather than n, which appears in the inertial term. This inconsistency causes an increase in truncation error, but it is used here because it makes the stability condition for the pseudotemperature method simpler. Of course, the error in the time-differencing is of no concern in dynamic relaxation provided the method is stable. A sufficient condition for the numerical stability of the method is the Courant stability condition:

$$h^{n+\frac{1}{2}} \le \min_{i,j}\{\epsilon/c\}^{i+\frac{1}{2},j+\frac{1}{2},n}. \tag{11}$$

Here c is the local speed in the reference configuration of whatever type of infinitesimal wave moves fastest. The zone width ϵ is the minimum of the lengths of the zone edges and diagonals. In order to accommodate meshes in which ϵ varies widely, CHIMP equalizes the stable time step throughout the mesh by assigning a suitable fictitious mass density (Welsh, 1967) which influences the local value of c.

For the quickest approach to an equilibrium solution, the damping coefficient β is chosen in such a way as to provide critical damping of the fundamental mode of the mesh. Hourglassing is controlled using small resistive forces found from an algorithm similar to the one proposed by Hancock (1979), although in CHIMP the algorithm is based on displacement rather than velocity.

3 Pseudotemperature Method for Incompressibility

CHIMP uses a new algorithm which enforces incompressibility by adjusting the hydrostatic pressures in the zones in such a manner that the zone volume change tends to zero for large times. So far the method has been used only for large-deformation elasticity, although extension to other types of materials appears straightforward.

If a material is incompressible, then the constitutive relation (3) must be modified to include a pressure term:

$$\sigma = \mathbf{h}(\mathbf{F}) - P\mathbf{F}^{-T}, \text{ det } \mathbf{F} = 1 \tag{12}$$

where P is a scalar field and $\mathbf{F}^{-T}$ is the inverse of the transpose of $\mathbf{F}$. Since the first Piola-Kirchhoff stress σ is related to the Cauchy stress τ by

$$\tau = \sigma\mathbf{F}^T/J, \; J = \det \mathbf{F}, \tag{13}$$

(12) is equivalent to

$$\tau = \mathbf{h}(\mathbf{F})\mathbf{F}^T - P\mathbf{1}, \text{ det } \mathbf{F} = 1. \tag{14}$$

The incompressible stress response function $\mathbf{h}$ in equations (12) is defined only on the set of unimodular tensors (det $\mathbf{F}$ = 1). In the pseudotemperature method, $\mathbf{h}$ is formally replaced by a compressible stress response function $\mathbf{h}^*$ with the following properties:

(*a*) $\mathbf{h}^*(\mathbf{F}) = \mathbf{h}(\mathbf{F})$ whenever det $\mathbf{F} = 1$;
(*b*) h^* must repesent the response of a physically reasonable material.

Assume that the mesh is initially in a stress-free state. The calculation proceeds using $\mathbf{h}^*$ as if the material were compressible. However, the computation of stress in each zone is modified in a special way. A new zone-centered scalar field θ is introduced and stored in an array. In each zone in each time step, θ is adjusted according to how far the zone is from satisfying incompressibility:

$$\theta^{i+\frac{1}{2}, j+\frac{1}{2}, 0} = 1,$$

$$\theta^{i+\frac{1}{2}, j+\frac{1}{2}, n} = \theta^{i+\frac{1}{2}, j+\frac{1}{2}, n-1} + \alpha(1 - J^{i+\frac{1}{2}, j+\frac{1}{2}, n}), \; n \geq 1 \tag{15}$$

where α is a dimensionless nonnegative constant. For each zone in each time step, a scalar quantity p is computed:

$$p^{i+\frac{1}{2}, j+\frac{1}{2}, n} = k(\theta^{i+\frac{1}{2}, j+\frac{1}{2}, n} - J^{i+\frac{1}{2}, j+\frac{1}{2}, n}) \tag{16}$$

where k is a constant. The stress tensor for a zone is then computed from

$$\sigma^{i+\frac{1}{2}, j+\frac{1}{2}, n} = \mathbf{h}^*(\mathbf{F}^{i+\frac{1}{2}, j+\frac{1}{2}, n}) - (Jp\mathbf{F}^{-T})^{i+\frac{1}{2}, j+\frac{1}{2}, n}. \tag{17}$$

p does not by itself give the hydrostatic pressure, since there is a contribution from Tr $\mathbf{h}^*$ $\mathbf{F}^T$.

Changes in θ have the effect of adjusting the *natural state*, or unstressed state, of the zone in a manner similar to temperature in thermoelasticity. For this reason θ is called the *pseudotemperature.* Note that if $J < 1$, the zone gets "hotter" according to equations (15). This induces an increase in pressure through equation (16), leading to an expansion of the zone in subsequent time steps. The effect is to drive J to 1 in the limit of large time.

Two new constants have been introduced, α and k. These are set through considerations of numerical stability and convergence rate, i.e., the number of time steps needed to attain a solution sufficiently close to the conditions of equilibrium and incompressibility. The remainder of this section presents a derivation of stability conditions for these parameters by a von Neumann stability analysis. A detailed look at this analysis is of interest because many related methods, especially those in which P is manipulated directly (as opposed to the indirect effect of pseudotemperature) have much more severe stability conditions.

Consider an infinite mesh which is uniform and rectangular in the reference configuration. Let ϵ_1 and ϵ_2 be the zone spacings in the x_1 and x_2 directions. Let the mass density in the *reference* configuration be ρ_0. Assume that the mesh undergoes a homogeneous deformation, possibly a large one, with principal stretches λ_1 and λ_2. These principal stretches are assumed to be in the x_1 and x_2 directions, respectively.

The von Neumann stability test (see Richtmeyer, 1967) attempts to detect whether any of the vibrational modes of the mesh, if excited by a small disturbance, grows nonphysically with time. A full two-dimensional treatment of this type is prohibitively complex. However, a reliable stability condition may be derived by making a reasonable assumption about the mode that is the first mode to exhibit nonphysical growth as h is increased. The assumption is that this most sensitive mode corresponds to a dilatational wave in the direction of one of the principal stretches. This assumption is reasonable because one would expect pseudotemperature to affect only dilatational modes, rather than shear modes, since its effect is felt only through the hydrostatic pressure. Further, in isotropic materials, the fastest dilatational waves occur in the direction of one of the principal stretches.

The above assumption allows the stability analysis to be carried out as though the mesh were one-dimensional. Consider an arbitrary node (i, j). Assume a motion of the form

$$\begin{aligned} y_1^{i,j,n} &= \lambda_1 x_1^{i,j} + u^{i,j,n} \\ y_2^{i,j,n} &= \lambda_2 x_2^{i,j} \end{aligned} \tag{18}$$

where $|u^{i,j,n}| \ll \epsilon_1$. Following the usual procedure for von Neumann stability analysis, assume

$$u^{i,j,n} = v^n e^{\sqrt{-1}\kappa i} \tag{19}$$

where $\kappa > 0$ is the wave number and v is a complex constant which characterizes the growth rate of a vibrational mode. (Here the superscript in v^n signifies exponentiation.) A constant time step h is assumed.

The first aim is to express the difference formulas entirely in terms of u values. Under the present assumptions, the difference formulas (1) and (5) simplify to (omitting the j labels)

$$(u^{i,n+1} - 2u^{i,n} + u^{i,n-1}) + \beta h(u^{i,n+1} - u^{i,n}) = \frac{h^2}{\rho_0 \epsilon_1}(\sigma_{11}^{i+\frac{1}{2},n} - \sigma_{11}^{i-\frac{1}{2},n}) \tag{20}$$

After linearizing the constitutive relation (17) and again using the difference formula (5) for the deformation gradient, one finds

$$\sigma_{11}^{i+\frac{1}{2},n} = \bar{\sigma}_{11} + \frac{\rho_0 c_1^2}{\epsilon_1}(u^{i+1,n} - u^{i,n}) - \lambda_2 p^{i+\frac{1}{2},n} \tag{21}$$

where $\bar{\sigma}_{11}$ is the stress associated with the homogeneous deformation. c_1 is the speed of dilatational waves in the x_1 direction, relative to the reference configuration, through a compressible solid characterized by $\mathbf{h}^*$ which has been subjected to

the homogeneous deformation. Combining equation (21) with its analogue for time step $n-1$ leads to

$$\sigma_{11}^{i+\frac{1}{2},n}-\sigma_{11}^{i+\frac{1}{2},n-1}=\frac{\rho_0 c_1^2}{\epsilon_1}(u^{i+1,n}-u^{i+1,n-1}-u^{i,n}+u^{i,n-1}) -\lambda_2(p^{i+\frac{1}{2},n}-p^{i+\frac{1}{2},n-1}) \tag{22}$$

Using equations (15), (16), and the appropriate expressions for the J terms yields

$$p^{i+\frac{1}{2},n}-p^{i+\frac{1}{2},n-1}=k\left\{\alpha(1-\lambda_1\lambda_2) +\frac{\lambda_2}{\epsilon_1}\left[(1+\alpha)(u^{i,n}-u^{i+1,n})+u^{i+1,n-1}-u^{i,n-1}\right]\right\} \tag{23}$$

Combining equation (20) with its analogue for the previous time step while using equations (22) and (23) provides an equation in which displacements are the only variables:

$$(1+\beta h)u^{i,n+1}+(-3-2\beta h)u^{i,n}+(3+\beta h)u^{i,n-1}-u^{i,n-2} =\frac{h^2c_1^2}{\epsilon_1^2}\left\{\left(\frac{k\lambda_2^2}{\rho_0 c_1^2}(1+\alpha)+1\right)\left(u^{i+1,n}-2u^{i,n}+u^{i-1,n}\right)\right. \tag{24}$$

$$\left.-\left(\frac{k\lambda_2^2}{\rho_0 c_1^2}+1\right)\left(u^{i+1,n-1}-2u^{i,n-1}+u^{i-1,n-1}\right)\right\}$$

Note that values for four different time steps appear in equation (24). Using the complex representation (19) and making use of the identity $\cos\kappa=(\exp(\sqrt{-1}\kappa)+\exp(-\sqrt{-1}\kappa))/2$ in equation (24) yields a cubic equation for the complex number v:

$$(1+\beta h)v^3 +\left[-3-2\beta h+(2-2\cos\kappa)\left(\frac{k\lambda_2^2}{\rho_0 c_1^2}(1+\alpha)+1\right)C^2\right]v^2$$

$$+\left[3+\beta h-(2-2\cos\kappa)\left(\frac{k\lambda_2^2}{\rho_0 c_1^2}+1\right)C^2\right]v-1=0 \tag{25}$$

where the C is the Courant number, $C=hc_1/\epsilon_1$.

The condition for stability is that $|v|\le 1$ for all vibrational modes. The cubic equation (25) has no root v whose modulus exceeds unity for any mode if both of the following conditions hold:

$$\left(\frac{k\lambda_2^2}{\rho_0 c_1^2}+1\right)C^2\le 1 \tag{26}$$

and

$$\alpha\le\beta h. \tag{27}$$

The stability conditions (26) and (27) are generalized to two dimensions by replacing c_1 by c, the maximum wave speed in any direction; λ_2 by λ, the maximum principal stretch; and ϵ_1 by ϵ, the zone width previously defined. Making these substitutions and using the definition of C, the stability condition is found to be

$$h^{n+\frac{1}{2}}\le h_{cr}\equiv\min_{i,j}\left\{\frac{\epsilon}{\sqrt{c^2+k\lambda^2/\rho_0}}\right\}^{i+\frac{1}{2},j+\frac{1}{2},n}. \tag{28}$$

(27) must also be satisfied once $h^{n+\frac{1}{2}}$ is chosen.

Inspection of (28) shows that one can adjust the relative effect of pseudotemperature on the stable time step by defining a separate value of k for each zone at the start of the run:

$$k^{i+\frac{1}{2},j+\frac{1}{2}}=K(\rho_0 c^2/\lambda^2)^{i+\frac{1}{2},j+\frac{1}{2},0} \tag{29}$$

where K is a preassigned nonnegative number. In practice, a specific value for the time step length is found from $h^{n+\frac{1}{2}}=sh_{cr}$, where s is an ad hoc safety factor, $0<s<1$, and h_{cr} is found from (28). The safety factor accounts for the approximate nature of the above stability analysis.

The stability condition (28) reduces to the usual Courant condition for the case of no pseudotemperature, $k=0$. Also note that (27) implies that the pseudotemperature method would be unconditionally unstable in its present form in the absence of damping.

In summary, if the k and h values are generated by equations (29) and (28), and if α satisfies (27), then the condition for stability is satisfied. In order to maximize the convergence rate, α and s should be chosen as large as possible such that (27) holds and $s<1$. Values of $\alpha^n=0.9\beta\,h^{n+\frac{1}{2}}$ and $s=0.9$ usually work well, although a smaller value of s is sometimes necessary. Note that α^n and $h^{n+\frac{1}{2}}$ must be recomputed at each time step. The constant K determines the stiffness of the material with respect to volume changes. As was shown above, too large a value of K would reduce the time step unacceptably. On the other hand, a very small a value of K would make the pseudotemperature method ineffectual. Based on experience a value of $K=0.5$ appears to result in good convergence rate for most problems.

In applications involving very large distortions, the stability criterion (28) may become so restrictive as to make the calculation inefficient. In this case, one can reassign all the mass densities and k values using equation (29) as is done initially in a run. In this event it is best to set all the node velocities to zero in order to avoid creating kinetic energy.

When applying the pseudotemperature method to a specific incompressible material, the user must provide a compressible constitutive relation $\mathbf{h}^*$ subject to the restrictions discussed earlier. There is considerable flexibility in this choice, and constitutive relations of the form used in Section 4 for the neo-Hookean material generally work well.

The user must also provide c, the maximum sound speed in any direction as a function of $\mathbf{F}$ relative to the reference configuration. The exact value of this quantity for an elastic material is

$$c=\max_{\mathbf{m},\mathbf{n}}\sqrt{\frac{1}{\rho_0}\frac{\partial h_{\alpha\beta}^*}{\partial F_{\gamma\delta}}m_\alpha m_\gamma n_\beta n_\delta} \tag{30}$$

where $\mathbf{m}$ and $\mathbf{n}$ are arbitrary unit vectors in the plane. The computation indicated in equation (30) can be tedious, but experience has shown that a simplifying assumption leads to a reliable estimate of c which is much easier to compute. One assumes that the maximum wave speed is that of a dilatational wave in the direction of one of the principal stretches. Then

$$c=\max(c_1,c_2)=\max\left(\sqrt{\frac{1}{\rho_0}\frac{\partial h_{11}^*}{\partial F_{11}}},\sqrt{\frac{1}{\rho_0}\frac{\partial h_{22}^*}{\partial F_{22}}}\right). \tag{31}$$

The case of one or both of the partial derivatives in equation (31) being negative corresponds to an unstable material, since such a material would possess an imaginary wave speed in some direction. If for some zone in some time step one of the derivatives is zero or negative, the zone's effect on the stable time step may be ignored, since in practice the zone always reverts to a materially stable condition within a small number of time steps anyway.

The use of pseudotemperature does not affect the truncation error of the difference formulas (1) and (5), since the value of J used in equations (15) and (16) is consistent with these formulas. Since only the large-time limit of the dynamic relaxation is of interest, any effect of pseudotemperature on the error in the time integration scheme is unimportant except as it affects numerical stability, an issue which has been dealt with above.

4 Sample Problems

This section presents results of application of the pseudotemperature method to two problems in incompressible finite elasticity. The first problem, the closure of a wedge-

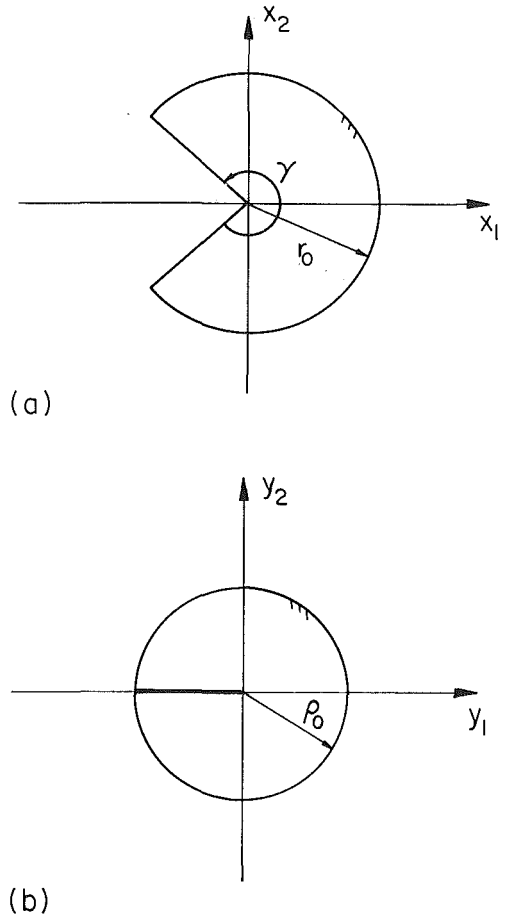


Fig. 4.1 Gap closure problem: (*a*) Reference configuration; (*b*) Deformed configuration

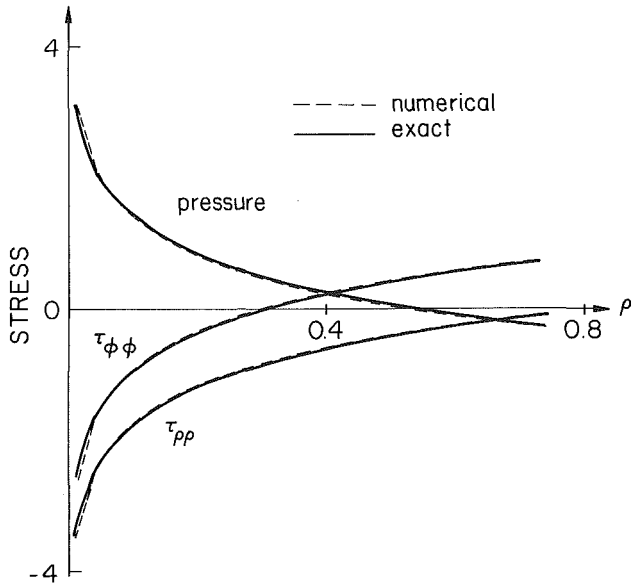


Fig. 4.2 Comparison of CHIMP results and exact values for normal Cauchy stress components and pressure in gap closure problem

shaped gap in a circular disk, demonstrates the capability of the method to predict singular solutions. The second problem is that of a rubber cylindrical rod being squeezed between two rigid walls, which demonstrates the method when contact boundary conditions are used.

4.1 Gap Closure Problem. An unstressed circular disk of radius r_0 contains a wedge-shaped gap, as shown in Fig. 4.1. Using a polar coordinate system with $\theta = 0$ on the horizontal axis, the gap is oriented so that the disk occupies the interval $-\gamma/2 \le \theta \le \gamma/2$ where γ is a constant, $0 < \gamma < 2\pi$.

The disk is composed of a neo-Hookean material, an incompressible hyperelastic solid whose Piola stress response function may be written

$$\mathbf{h}(\mathbf{F}) = \mu(\mathbf{F} - \mathbf{F}^{-T}),\ \det \mathbf{F} = 1 \tag{32}$$

where μ is a positive constant. The corresponding Cauchy stress tensor is $\tau = \mu(\mathbf{F}\mathbf{F}^T - \mathbf{1}) - P\mathbf{1}$.

Let (ρ, ϕ, ζ) be the polar coordinates of the image of a particle initially at (r, θ, z). Assume plane strain, $\zeta \equiv z$. The boundary conditions are such that the gap is welded shut and the outer circular boundary is traction-free. This problem belongs to a class of problems involving wedge-shaped regions investigated by Singh and Pipkin (1965) and separately by Klingbeil and Shield (1966). Its tractability stems from the fact that the constraint of incompressibility completely determines the deformation. The requirement of incompressibility implies that only one deformation with polar symmetry is possible:

$$\rho = (\gamma/2\pi)^{1/2} r,\ \phi = 2\pi\theta/\gamma. \tag{33}$$

Since the deformation is now determined, the stress field is also determined through (32) except for the scalar field P. Because of polar symmetry and the isotropy of the material, the Cauchy shear stress components $\tau_{\rho\phi}$, $\tau_{\zeta\phi}$, and $\tau_{\rho\zeta}$ all vanish, and the nonvanishing components are independent of ϕ. Therefore, the scalar field P may be found by integrating the Eulerian from the equilibrium equation in the deformed configuration. After applying the traction boundary condition at the outer edge, one finds that the nonzero components of the Cauchy stress fields are given by

$$\tau_{\rho\rho}(\rho) = \mu\left(\frac{2\pi}{\gamma} - \frac{\gamma}{2\pi}\right)\log\frac{\rho}{\rho_0},$$

$$\tau_{\phi\phi}(\rho) = \mu\left(\frac{2\pi}{\gamma} - \frac{\gamma}{2\pi}\right)\left(1 + \log\frac{\rho}{\rho_0}\right), \tag{34}$$

$$\tau_{\zeta\zeta}(\rho) = \mu\left(1 - \frac{\gamma}{2\pi}\right) + \mu\left(\frac{2\pi}{\gamma} - \frac{\gamma}{2\pi}\right)\log\frac{\rho}{\rho_0},\ 0 < \rho \le \rho_0$$

where $\rho_0 = (\gamma/2\pi)^{1/2} r_0$.

CHIMP was used to model this problem for the case $\gamma = 4\pi/3$, $r_0 = 1$, $\mu = 1$. The mesh for this problem used 13 rows of constant radius and 12 columns of constant angle. Exploiting symmetry, a mesh modeling only one quadrant was used, with the x_2 axis becoming a lubricated wall. The edge of gap was modeled as another lubricated wall, with its angle a ramp function of time. The ramp took about 100 time steps to reach the fully closed position.

The compressible Piola stress response function used for the pseudotemperature method was

$$\mathbf{h}^*(\mathbf{F}) = \mu\left\{\mathbf{F} + (J^2 - 2)\mathbf{F}^{-T}\right\},\ J = \det \mathbf{F} > 0 \tag{35}$$

The resulting Cauchy stress tensor is $\tau = \mu\{(\mathbf{F}\mathbf{F}^T - \mathbf{1})/J + (J - 1/J)\mathbf{1}\}$. To evaluate c as discussed at the end of the previous section, the deformation gradient tensor is written as

$$[\mathbf{F}] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{36}$$

in which the coordinate frame has the x_1 axis parallel to the direction of maximum stretch. (It is not necessary to find this direction.) The general form (36) continues to hold when a small dilatational wave in the x_1 or the x_2 direction is super-

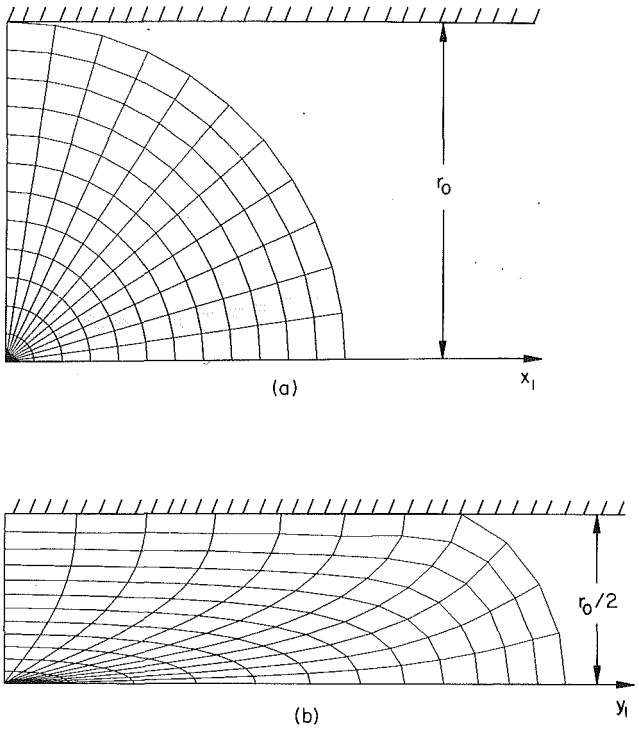


Fig. 4.3 Rubber cylinder compression problem: (*a*) Undeformed; (*b*) Deformed

posed on the large homogeneous deformation, but in the case of a wave λ_1 or λ_2 is time-dependent. Equations (35) and (36) imply

$$\sigma_{11} = h_{11}^*(\mathbf{F}) = \mu(\lambda_1 + \lambda_1\lambda_2^2 - 2/\lambda_1). \tag{37}$$

Thus

$$c_1 = \sqrt{\frac{1}{\rho_0}\frac{\partial h_{11}^*}{\partial \lambda_1}} = \sqrt{\frac{\mu}{\rho_0}(1 + \lambda_2^2 + 2/\lambda_1^2)}. \tag{38}$$

c_2 is found from equation (38) with all the subscripts 1 and 2 interchanged. Then $c = \max(c_1, c_2)$. The principal stretches λ_1 and λ_2 are found from the relations

$$J = \lambda_1\lambda_2 \text{ and } \mathrm{Tr}(\mathbf{F}\mathbf{F}^T) = \lambda_1^2 + \lambda_2^2. \tag{39}$$

The left-hand sides of (39) are easily computed from $\mathbf{F}$ in *any* basis, since they are invariants.

The pseudotemperature parameters were $\alpha = 0.9$ and $K = 0.5$. The entire simulation was run for 700 time steps. Figure 4.2 compares the numerical results against the exact solution derived above for $\tau_{\rho\rho}$, $\tau_{\phi\phi}$, and pressure as a function of ρ.

4.2 Compression of a Rubber Cylinder Between Rigid Walls. A circular cylindrical rod composed of a neo-Hookean material is compressed between two rigid lubricated walls. The final distance between the walls is the radius of the cylinder. The CHIMP mesh for this problem employed 12 radial lines and 13 circular lines (see Fig. 4.3(*a*)). The walls were initially spaced at the diameter of the cylinder and brought together over a period of 300 time steps. The pseudotemperature parameters were $\alpha = 0.9$ and $K = 0.5$. The entire simulation was run for 1000 time steps. Figure 4.3(*b*) shows the deformed mesh. The analogous problem for walls with a "rough" surface, which prevents motion of the boundary of the cylinder after contact with the wall is initially made, has also been modeled successfully.

5 Summary

Pseudotemperature has proven to be a useful means of imposing the constraint of incompressibility when computing large elastic deformations using a dynamic relaxation method. It has the advantage of requiring minimum disruption to the architecture of a computer code designed for compressible materials. It also has the property of enforcing incompressibility with greater and greater accuracy as a calculation progresses. The above discussion of stability shows that suitable values of the scalars α and K may always be chosen, provided viscous damping is present.

Future work will include application of this approach to other constraints, such as the presence of inextensible fibers. Attempts are also being made to extend these ideas to fully dynamic calculations.

References

Day, A. S., 1965, "An Introduction to Dynamic Relaxation," *The Engineer*, London, Vol. 219, pp. 218–221.

Hancock, S., 1979, "Equation for Forces in Axisymmetric Lagrange Zones," *Int. Conf. on Computational Methods in Nonlinear Mechanics*, Austin, Texas.

Herrmann, W., and Bertholf, L. D., 1983, "Explicit Lagrangian Finite-Difference Methods," *Computational Methods for Transient Analysis*, Belytschko, T., and Hughes, T. J. R., eds., Elsevier Science Publishers, Amsterdam, pp. 361–416.

Klingbeil, W. W., and Shield, R. T., 1966, "On a Class of Solutions in Plane Finite Elasticity," *J. Appl. Math. and Phys.* (ZAMP), Vol. 17, pp. 489–511.

Needleman, A., and Shih, C. F., 1978, "A Finite Element Method for Plane Strain Deformations of Incompressible Solids," *Computer Methods in Applied Mechanics and Engineering*, Vol. 15, pp. 223–240.

Richtmeyer, R. D., and Morton, K. W., 1967, *Difference Methods for Initial-Value Problems*, 2nd ed., Interscience Publishers, New York, p. 9.

Silling, S. A., 1985, "CHIMP—A computer program for finite elastostatics," California Institute of Technology Division of Engineering and Applied Science report 54.

Singh, M., and Pipkin, A. C., 1965, "Note on Ericksen's Problem," *J. Appl. Math. and Phys.* (ZAMP), Vol. 16, pp. 706–709.

Underwood, P., 1983, "Dynamic Relaxation," *Computational Methods for Transient Analysis*, Belytschko, T., and Hughes, T. J. R., eds., Elsevier Science Publishers, Amsterdam, Holland, pp. 245–265.

Welsh, A. K., 1967, "Discussion on Dynamic Relaxation," *Proc. Inst. Civil Engrs.*, Vol. 37, pp. 723–750.

Wilkins, M. L., 1969, "Calculation of Elastic-Plastic Flow," Lawrence Radiation Laboratory report UCRL-7322.

Plane Strain Dislocations in Linear Elastic Diffusive Solids

J. W. Rudnicki
Department of Civil Engineering,
Northwestern University,
Evanston, IL 60201

Solutions are obtained for the stress and pore pressure due to sudden introduction of plane strain dislocations in a linear elastic, fluid-infiltrated, Biot, solid. Previous solutions have required that the pore fluid pressure and its gradient be continuous. Consequently, the antisymmetry (symmetry) of the pore pressure p *about* y = 0 *requires that this plane be permeable* (p = 0) *for a shear dislocation and impermeable* ($\partial p/\partial y = 0$) *for an opening dislocation. Here Fourier and Laplace transforms are used to obtain the stress and pore pressure due to sudden introduction of a shear dislocation on an impermeable plane and an opening dislocation on a permeable plane. The pore pressure is discontinuous on* y = 0 *for the shear dislocation and its gradient is discontinuous on* y = 0 *for the opening dislocation. The time-dependence of the traction induced on* y = 0 *is identical for shear and opening dislocations on an impermeable plane, but differs significantly from that for dislocations on a permeable plane. More specifically, the traction on an impermeable plane does not decay monotonically from its short-time (undrained) value as it does on a permeable plane; instead, it first increases to a peak in excess of the short-time value by about 20 percent of the difference between the short and long time values. Differences also occur in the distribution of stresses and pore pressure depending on whether the dislocations are emplaced on permeable or impermeable planes.*

Introduction

The presence of an infiltrating fluid that can diffuse in response to an inhomogeneous mean stress field can introduce time-dependence into the response of an otherwise linear-elastic solid. Although a linear theory is obviously an approximation to actual behavior, this theory is rich enough to provide insight into the nature of coupling between deformation and diffusion and guidance into more complicated nonlinear problems. Moreover, there is often insufficient data to warrant the construction of a more elaborate theory.

The equations describing the response of a linear elastic, diffusive solid were first formulated by Biot (1941a) within the context of a fluid-saturated porous elastic solid. More recently, Rice and Cleary (1976) reformulated these equations in a way that is often more convenient. Solutions to these equations have been widely used in consolidation theory (e.g. Biot, 1941b; Biot and Clingan, 1941) and, more recently, in studying the role of coupling between deformation and diffusion of ground water on earth faulting (see Rudnicki, 1985, for a review). The equations have also been applied to biological materials (e.g., Kuei, 1977; Mow and Lai, 1980). Indeed, the formulation is sufficiently general to describe the linearized response of any solid containing a diffusing species that can be characterized by a relation between two scalar variables, for example, pressure and fractional volume change in the case of groundwater. The formal analogy of these equations to completely coupled thermoelasticity has also been noted (Biot, 1956; Rice and Cleary, 1976; Rice 1979).

This paper considers the problem of plane strain (edge) dislocations in a linear elastic diffusive solid. Booker (1974), using the stress function formulation of McNamee and Gibson (1960a,b) and integral transforms, obtained the solution for a shear (gliding edge) dislocation in the special case that both solid and fluid constituents are incompressible. Rice and Cleary (1976), using a complex variable formulation, derived the solution for arbitrarily compressible constituents. These solutions correspond to the case in which the glide plane of the dislocation (the plane containing the dislocation line and the Burger's vector) is permeable to the diffusing species. Although neither author emphasizes this feature, it results because the mean stress and pore pressure are antisymmetric about the glide plane. If the pore pressure is continuous, then it must be zero on the glide plane. Another possibility, however, is that glide plane is impermeable to the diffusing species. Now the pore pressure can be discontinuous on this plane.

The stresses and displacements for an opening (climbing edge) dislocation can also be obtained from the results of Rice and Cleary (1976) although they do not explicitly display this solution. In this case the boundary condition on the pore fluid pressure corresponds to no flow across the plane containing the Burger's vector and the dislocation line. Again, however,

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 16, 1986; final revision, November 15, 1986.

there is another possibility: no change in pore fluid pressure on this plane. Now, the symmetry of the mean-stress and pore pressure requires that the gradient of the pore fluid pressure be discontinuous.

In this paper, Fourier and Laplace transforms are used to derive the stresses and pore fluid pressure due to sudden introduction of a plane strain shear dislocation on an impermeable plane and of a plane strain opening dislocation on a permeable plane. These solutions are compared and contrasted with those obtained by Rice and Cleary (1976). Although applications of these solutions are not explored here, Rudnicki (1986) has discussed the implications of the shear dislocation solutions for slip on an impermeable fault in the earth's crust. The dislocation solutions provide only crude models of sliding or opening cracks, but solutions for more realistic geometries can be constructed by superposition or by implementing the fundamental dislocation solutions in a numerical procedure.

This paper first concisely describes the governing equations and obtains the solution for the doubly transformed stresses and pore pressure. Then the boundary conditions for the different solutions are presented. The inversion of the transformed solution for the shear dislocation is discussed in detail, but inversion of the opening dislocation is similar and, consequently, is only outlined. Finally, the interrelations of these solutions with those obtained by Rice and Cleary (1976) are discussed.

Governing Equations

The governing equations for linear elastic, fluid-infiltrated solids were first derived by Biot (1941a), but the description here follows a convenient rearrangement of these equations by Rice and Cleary (1976). In this theory, the presence of the diffusing species is incorporated via two variables in addition to the usual ones of linear elasticity. Here, these are taken to be the pore fluid pressure p and the mass content of diffusing species per unit volume of porous solid m. For plane strain deformation in the xy plane (no displacement in the z direction) the displacements in the x and y directions, u_x and u_y, do not depend on z. The nonzero strains are

$$\epsilon_{\alpha\beta} = \frac{1}{2}(\partial u_\alpha/\partial x_\beta + \partial u_\beta/\partial x_\alpha) \tag{1}$$

where $(\alpha, \beta) = (x, y)$. These strains and the alteration of m from an ambient value m_o are related to the total stresses σ_{xx}, σ_{xy}, and σ_{yy} and to the pore fluid pressure p as follows:

$$2G\epsilon_{\alpha\beta} = \sigma_{\alpha\beta} - \nu(\sigma_{xx} + \sigma_{yy})\delta_{\alpha\beta} + [3(\nu_u - \nu)/B(1+\nu_u)]p\ \delta_{\alpha\beta} \tag{2}$$

$$m - m_o = \frac{3\rho_o(\nu_u - \nu)}{2GB(1+\nu_u)}[\sigma_{xx} + \sigma_{yy} + 3p/B(1+\nu_u)] \tag{3}$$

In equations (2) and (3) G is the shear modulus; ν and ν_u are Poisson's ratios governing drained (long-time) and undrained (short-time) response, respectively; B is Skempton's coefficient, the ratio of an increment of pore fluid pressure to an increment of mean normal compression during undrained response; ρ_o is the density of the homogeneous diffusing species; and $\delta_{\alpha\beta}$ is the Kronecker delta ($\delta_{\alpha\beta} = 1$, if $\alpha = \beta$ and $\delta_{\alpha\beta} = 0$, otherwise).

For deformation that is slow enough so that any alterations in pore fluid pressure are equilibrated by mass diffusion, the response is said to be drained and, since $p = 0$ in this case, equation (2) reduces to the usual elasticity relation. Deformation that is too rapid to allow time for diffusion is said to be undrained. In this case, $m = m_o$, and solving for p in equation (3) and substituting in equation (2) again yields the form of the usual elasticity relation with ν_u replacing ν.

The final constitutive equation is Darcy's law which, in the absence of body forces, states that the mass flow rate in the α direction per unit area, q_α, is proportional to the gradient of pore fluid pressure:

$$q_\alpha = -\rho_o \kappa\ \partial p/\partial x_\alpha \tag{4}$$

Here κ is a permeability often expressed as k/μ where k has the units of area and μ is the fluid viscosity.

For plane strain deformation, the governing field equations can be written as follows in terms of the stresses $\sigma_{\alpha\beta}$ and pore pressure p:

$$\partial\sigma_{xx}/\partial x + \partial\sigma_{xy}/\partial y = 0 \tag{5}$$

$$\partial\sigma_{xy}/\partial x + \partial\sigma_{yy}/\partial y = 0 \tag{6}$$

$$\nabla^2(\sigma_{xx} + \sigma_{yy} + 2\eta p) = 0 \tag{7}$$

$$(c\nabla^2 - \partial/\partial t)[\sigma_{xx} + \sigma_{yy} + (2\eta/\mu)p] = 0 \tag{8}$$

where $\nabla^2(\ldots) = [(\partial^2/\partial x^2) + (\partial^2/\partial y^2)](\ldots)$, c is a diffusivity, $\mu = (\nu_u - \nu)/(1 - \nu)$

and

$$\eta = 3(\nu_u - \nu)/2B(1+\nu_u)(1-\nu).$$

Equations (5) and (6) express equilibrium of total stresses in the absence of body forces and equation (7) expresses compatibility of strains. The diffusion equation (8) is the result of combining Darcy's law (4) with an equation of fluid mass conservation and using equation (7). Comparing equation (8) with (3) reveals that the quantity in square brackets in equation (8) is proportional to the alteration of fluid mass content. Hence, as emphasized by Rice and Cleary (1976), the fluid mass content m satisfies a homogeneous diffusion equation although the pore fluid pressure, in general, does not. Rice and Cleary (1976) have given a full discussion of these equations and have tabulated values of material parameters inferred from laboratory tests on rocks (also see Rudnicki, 1985) and Rice (1979b) and Rice and Rudnicki (1979) have given some estimates of ν and ν_u for conditions near faults in the earth's crust.

The equations (5)–(8), subject to boundary conditions to be discussed in succeeding subsections, will be solved using the Fourier transform on x and the Laplace transform on t. The Laplace transform of a function $f(x, t)$ is defined by

$$\tilde{f}(x,s) = \int_0^\infty \exp(-st) f(x,t)\, dt \tag{9}$$

and the inversion is denoted by

$$f(x,t) = L^{-1}\{\tilde{f}(x,s)\} = \frac{1}{2\pi\iota}\int_{\mathrm{Br}} \tilde{f}(x,s)\exp(st)\, ds \tag{10}$$

where $\iota = (-1)^{1/2}$ and Br denotes the Bromwich contour. The Fourier transform is defined by

$$\hat{f}(\kappa,s) = \int_{-\infty}^\infty \tilde{f}(x,s)\exp(-\iota\kappa x)\, dx \tag{11}$$

with inversion

$$\tilde{f}(x,s) = F^{-1}[\hat{f}(\kappa,s)] = \frac{1}{2\pi}\int_{-\infty}^\infty \hat{f}(\kappa,s)\exp(\iota\kappa x)\, d\kappa \tag{12}$$

Applying the Fourier and Laplace transforms to equations (5)–(8) yields the following results:

$$\iota\ \kappa\ \hat{\sigma}_{xx} + \frac{d\hat{\sigma}_{xy}}{dy} = 0 \tag{13}$$

$$\iota\ \kappa\ \hat{\sigma}_{xy} + \frac{d\hat{\sigma}_{yy}}{dy} = 0 \tag{14}$$

$$\left(-\kappa^2 + \frac{d^2}{dy^2}\right)(\hat{\sigma}_{xx} + \hat{\sigma}_{yy} + 2\eta\ \hat{p}) = 0 \tag{15}$$

$$\left\{\frac{d^2}{dy^2}-n^2(\kappa)\right\}[\hat{\sigma}_{xx}+\hat{\sigma}_{yy}+(2\eta/\mu)\hat{p}]=0 \tag{16}$$

where $n^2(\kappa) = \kappa^2 + s/c$. These equations are identical to those obtained by Rice and Simons (1976) except that $-\iota\kappa V/c$ in their expression for $n^2(k)$ is replaced by s/c. Consequently, the solution of these equations can be obtained directly from their results and is as follows:

$$\frac{1}{2}(\hat{\sigma}_{xx}+\hat{\sigma}_{yy}) = A\ e^{-m(\kappa)y}+Be^{-n(\kappa)y} \tag{17}$$

$$\eta\ \hat{p} = -\mu\ A\ e^{-m(\kappa)y}-B\ e^{-n(\kappa)y} \tag{18}$$

$$\frac{1}{2}(\hat{\sigma}_{yy}-\hat{\sigma}_{xx}) = [C+m(\kappa)y\ A]e^{-m(\kappa)y}$$

$$-\frac{B}{s/c}[\kappa^2+n^2(\kappa)]e^{-n(\kappa)y} \tag{19}$$

$$\hat{\sigma}_{xy} = -[\iota\kappa\ m^{-1}(\kappa)C+\iota\kappa\ y\ A]e^{-m(\kappa)y}$$

$$+\frac{B}{s/c}\ 2\iota\kappa n(\kappa)e^{-n(\kappa)y} \tag{20}$$

where A, B, and C are functions of κ and s to be determined, and $m^2(\kappa) = \kappa^2$. Note that the doubly transformed solution for the fluid mass content per unit volume m is proportional to $B\ e^{-n(\kappa)y}$. To insure convergence of the inversion integrals in $y > 0$, $m(\kappa)$ and $n(\kappa)$ are subject to the following restrictions:

$$\mathrm{Re}[m(\kappa)]\geq 0 \tag{21}$$

$$\mathrm{Re}[n(\kappa)]\geq 0 \tag{22}$$

where Re[. . .] stands for "the real part of [. . .]."

The functions A, B, and C can be determined from the boundary conditions which are discussed in the next subsection.

Boundary Conditions

The introduction of a shear (gliding edge) dislocation at the origin corresponds to cutting the negative x axis, displacing the top to the right and the bottom to the left by the same amount, then bonding the cut elastic plane back together. The resulting discontinuity in the x displacement is described as follows:

$$u_x(x,y=0^+,t)-u_x(x,y=0^-,t)$$
$$=[2\pi(1-\nu_u)b_x/G]H(-x)H(t) \tag{23}$$

where $H(. . .)$ denotes the unit step function and the notation $y = 0^{\pm}$ indicates that u_x is to be evaluated as the x axis is approached from above or below. The magnitude of the discontinuity is measured by b_x and the factor $2\pi(1 - \nu_u)/G$ has been introduced with a view to simplifying later expressions. Because the displacements are antisymmetric with respect to the plane $y = 0$, the problem can be formulated in the upper half-plane, $y \geq 0$, with equation (23) rewritten as

$$u_x(x,0^+,t) = [\pi(1-\nu_u)b_x/G]H(-x)H(t) \tag{24}$$

Because of antisymmetry and continuity of total tractions on $y = 0$, the normal stress on this plane σ_{yy} is zero:

$$\sigma_{yy}(x,0^+,t) = 0 \tag{25}$$

If the pore fluid pressure p is continuous, then antisymmetry requires that it be zero on the plane $y = 0$:

$$p(x,0^+,t) = 0 \tag{26}$$

This is the problem for which the solution has been given by Rice and Cleary (1976) (and earlier by Booker, 1974, for incompressible constituents corresponding to $B = 1$ and $\nu_u = 0.5$). Because $\partial p/\partial y$ is not zero on $y = 0$ in this case, flow across $y = 0$ occurs according to equation (4). Another possibility is, however, that the plane $y = 0$ is impermeable to the diffusing species. As discussed by Rudnicki (1986), this can occur for an earth fault because clay gouge or finely ground material is present in the fault zone. In this case no flow can occur across $y = 0$ and the boundary condition enforcing this constraint is the following:

$$\frac{\partial p}{\partial y}(x,0^+,t) = 0 \tag{27}$$

Because the solution to the field equations is written in terms of stresses, it is also convenient to express the boundary condition (24) in terms of the stresses. Differentiating (24) with respect to x yields

$$\frac{\partial u_x}{\partial x}(x,0^+,t) = -[\pi(1-\nu_u)b_x/G]\delta(x)H(t) \tag{28}$$

where $\delta(x)$ is the Dirac delta function. Because $\epsilon_{xx} = \partial u_x/\partial x$, equation (28) can be substituted into equation (2) and the result, after using equation (25), is

$$-2\pi(1-\mu)b_x\delta(x)H(t) = \sigma_{xx}(x,0,t)+2\eta\ p(x,0,t) \tag{29}$$

where μ and η are defined following equation (8). If the plane $y = 0$ is permeable and equation (26) is satisfied, the second term vanishes. In this case, the change in fluid mass content on $y = 0$ is proportional to $\sigma_{xx}(x, 0, t)$. Because m satisfies the homogeneous diffusion equation, the solution is that for a fluid mass dipole (Carslaw and Jaeger, 1959) given by Rice and Cleary (1976). If the fault plane is impermeable and equation (27) is appropriate, the resulting boundary condition on m is not so simple and this is a source of the additional complexity in this solution by comparison with that for the permeable plane.

The boundary condition for an opening (climbing edge) dislocation corresponds to introducing a discontinuity in the y displacement on the negative x axis. This problem can again be formulated in the upper half-plane, $y \geq 0$, by noting that the displacements are now symmetric about $y = 0$. The boundary conditions can be written as follows:

$$u_y(x,0^+,t) = [\pi(1-\nu_u)b_y/G]H(-x)H(t) \tag{30}$$

$$\sigma_{xy}(x,0^+,t) = 0 \tag{31}$$

where, again, the constant factor multiplying b_y has been introduced to simplify later expressions.

If the derivative of the pore pressure in the y direction is continuous, then the symmetry of the problem requires that it be zero on $y = 0$. Now, however, an alternative boundary condition is equation (26). In this case the fluid mass flux is discontinuous on $y = 0$. This boundary condition models a thin high permeability layer in which the easy flow of fluid maintains the pore fluid pressure at its ambient value. This boundary condition may also be appropriate when opening is accompanied by injection of fluid mass.

In the next section the solution for the shear dislocation with an impermeable boundary at $y = 0$ (equation (27)) will be completed. The following section treats the opening dislocation with a permeable boundary at $y = 0$ (equation (26)). Because the conversion of the boundary condition (30) to a condition on the stresses is accomplished more easily in terms of the transformed quantities, this task is deferred to this later section.

Shear Dislocation on an Impermeable Boundary

The boundary conditions for the shear dislocation are equa-

tions (25), (29), and, for an impermeable boundary at $y = 0$, equation (27). Taking the Fourier and Laplace transforms of these equations, then substituting equations (17)–(20) yields three equations for the functions A, B, and C. Solving these equations and substituting these expressions into equations (17)–(20) yields the doubly transformed stresses and pore fluid pressure. The expressions for the mean stress and pore pressure are as follows:

$$\hat{\sigma} = (-b_x\pi/s)\{e^{-m(\kappa)y} - \mu[m(\kappa)/n(\kappa)]e^{-n(\kappa)y}\} \tag{32}$$

$$\eta\,\hat{p} = (b_x\pi/s)\mu\{e^{-m(\kappa)y} - [m(\kappa)/n(\kappa)]e^{-n(\kappa)y}\} \tag{33}$$

where $\hat{\sigma} = (1/2)(\hat{\sigma}_{xx} + \hat{\sigma}_{yy})$. It will be convenient to combine equations (19) and (20) into the complex form:

$$\tau = \frac{1}{2}(\sigma_{yy} - \sigma_{xx}) + \iota\,\sigma_{xy} \tag{34}$$

The result for the double transform of τ, after substituting the expressions for A, B, and C, is as follows:

$$\hat{\tau} = (b_x\pi/s)[1 + \kappa/m(\kappa)][1 - m(\kappa)y]e^{-m(\kappa)y} - (b_x\mu\pi c/s^2)[m(\kappa)/n(\kappa)]\{[\kappa + n(\kappa)]^2 e^{-n(\kappa)y} - 2\kappa^2[1 + \kappa/m(\kappa)]e^{-m(\kappa)y}\} \tag{35}$$

The Laplace transform variable s appears only in $n(\kappa)$ and as a simple divisor. Terms without $n(\kappa)$ can be inverted immediately by noting that s^{-1} is the transform of the unit step function. These terms give the instantaneous undrained response and it can be anticipated that the spatial dependence, given by the inversion of the Fourier transforms in those terms, is identical to that of the usual elasticity solution. This can be verified by doing the following inverse transforms:

$$F^{-1}\{e^{-m(\kappa)y}\} = y/\pi r^2 \tag{36}$$

$$F^{-1}\{[1 + \kappa/m(\kappa)][1 - m(\kappa)y]\} = \iota x(x - \iota y)^2/\pi r^4 \tag{37}$$

where $r^2 = x^2 + y^2$.

The expression for the mean normal stress (32) and pore fluid pressure (33) can be written using equation (36) as follows:

$$\sigma = -b_x\{(y/r^2) - \mu\, I(x,y,t)\} \tag{38}$$

$$\eta p = \mu b_x\{(y/r^2) - I(x,y,t)\} \tag{39}$$

where the Laplace transform of I is given by

$$\tilde{I} = (2s)^{-1}\int_{-\infty}^{\infty}[m(\kappa)/n(\kappa)]\exp[\iota\kappa x - n(\kappa)y]d\kappa \tag{40}$$

and equations (38) and (39) are understood to apply for $t \geq 0$. The restrictions on $m(\kappa)$ and $n(\kappa)$ (equations (21), (22)) can be used to convert equation (40) to the following integral over positive values of κ:

$$\tilde{I} = s^{-1}\int_0^{\infty}\kappa(\kappa^2 + s/c)^{-1/2}\cos(\kappa x)\exp[-y(\kappa^2 + s/c)^{1/2}]d\kappa \tag{41}$$

where the expression for $n(\kappa)$ has been used. Substitution of equation (37) into (35) and use of equations (21), (22) leads to an expression for τ:

$$\tau = b_x[\iota\, x/(x + iy)^2] - \mu b_x\left\{\iota\frac{\partial^2}{\partial x\partial y}I^* - \frac{\partial^2}{\partial x^2}I^* - I^{\#} + I\right\} \tag{42}$$

where the Laplace transforms of I^* and $I^{\#}$ are given by

$$\tilde{I}^*(x,y,s) = (2c/s)\tilde{I}(x,y,s) \tag{43}$$

and

$$\tilde{I}^{\#}(x,y,s) = (2c/s^2)\int_0^{\infty}\kappa^3(\kappa^2 + s/c)^{-1/2}\exp[\iota\kappa(x + \iota y)]d\kappa \tag{44}$$

The inversion of $I(x, y, t)$ is described in the Appendix. The result can be written compactly as

$$I(x,y,t) = -\mathrm{Im}[W(x,y,t)/z] \tag{45}$$

where $z = x + \iota y$, Im [. . .] denotes "the imaginary part of [. . .]" and $W(x, y, t)$ is defined by

$$W(x,y,t) = \mathrm{erfc}[y/(4ct)^{1/2}] + \exp(-r^2/4ct)\mathrm{erf}[\iota\, x/(4ct)^{1/2}] \tag{46}$$

In equation (46) erf(ξ) is the error function defined by equation (7.1.2) of Abramowitz and Stegun (1964) (hereafter abbreviated AS):

$$\mathrm{erf}(\xi) = (2/\pi^{1/2})\int_0^{\xi}\exp(-\alpha^2)d\alpha \tag{47}$$

where ξ can be complex and the complementary error function is given by

$$\mathrm{erfc}(\xi) = 1 - \mathrm{erf}(\xi) \tag{48}$$

The task remaining is the inversion of the integrals $I^{\#}$ and I^*. Because of equation (43), I^* is given by

$$I^*(x,y,t) = 2c\int_0^t I(x,y,\lambda)d\lambda \tag{49}$$

Substituting equation (45) into (49) yields an expression for I^*:

$$I^*(x,y,t) = (8cty/r^2)i^2\mathrm{erfc}[y/(4ct)^{1/2}] - [2x^2(4ct)^{1/2}/r^2]i\mathrm{erfc}[y/(4ct)^{1/2}] + 2|x|\int_{y/r}^{1}(1 - \xi^2)^{1/2}\mathrm{erfc}[\xi r/(4ct)^{1/2}]d\xi \tag{50}$$

where $i^n\mathrm{erfc}(z)$ are repeated integrals of the complementary error function [AS, Section 7.2].

The details of the inversion of equation (44) for $I^{\#}$ are described in the Appendix and the result is

$$I^{\#}(x,y,t) = -2(4ct/\pi)^{1/2}z^{-2} - \iota[4ctz^{-3} + 2/z]w[z/(4ct)^{1/2}] \tag{51}$$

where (AS, 7.1.3)

$$w(\zeta) = \exp(-\zeta^2)\mathrm{erfc}(-\iota\zeta) \tag{52}$$

Substituting equations (45), (50), and (51) into equations (38), (39), and (42) and carrying out the differentiations in equation (42) yield the following expressions for the stresses and pore fluid pressure due to sudden introduction of a shear dislocation on an impermeable plane:

$$\sigma = b_x\mathrm{Im}\{[1 - \mu\, W(x,y,t)]/z\} \tag{53}$$

$$\eta p = -\mu\, b_x\mathrm{Im}\{[1 - W(x,y,t)]/z\} \tag{54}$$

$$\tau = \iota x b_x/z^2 - \mu b_x\Big\{\iota 4ctz^{-3}\{w[z/4ct)^{1/2}] - W(x,y,t)\} + 2(4ct/\pi)^{1/2}z^{-2}[1 - \exp(-y^2/4ct)] + z^{-2}\mathrm{Im}[zW(x,y,t)] + 2\iota z^{-1}w[z/4ct)^{1/2}]\Big\} \tag{55}$$

The first term in each expression gives the instantaneous response at $t = 0$. These terms are identical to the usual elasticity expressions with the undrained value of Poisson's ratio, ν_u. For $t \to \infty$, these expressions again reduce to those of classical elasticity with the drained value of Poisson's ratio.

Opening Dislocation on a Permeable Plane

The solution for an opening dislocation with a permeable boundary ($p = 0$) at $y = 0$ can be obtained in a manner similar to that of the last section and, hence, will be described concisely. As before, the solution to the governing equations (5)–(8) is given by equations (17)–(20) subject to equations (21) and (22), and the functions A, B, and C are to be determined by the boundary conditions (26), (30), and (31). Because the boundary condition (30) is not expressed in terms of stresses and pore pressure some manipulations are, however, required.

Differentiating equation (30) with respect to x and using equation (31) with equations (1) and (2) yields

$$\frac{\partial u_x}{\partial y}(x,0,t) = [\pi(1-\nu_u)b_y/G]\delta(x)H(t) \tag{56}$$

Doubly transforming gives

$$\frac{d\hat{u}_x}{dy}(\kappa,0,s) = \pi(1-\nu_u)b_y/Gs \tag{57}$$

This condition can be converted to one on stress and pore pressure by using equation (1) in (2) with $\alpha = \beta = x$, differentiating with respect to y, and doubly transforming. The result is

$$2G\iota\kappa\frac{d\hat{u}_x}{dy} = \frac{d\hat{\sigma}_{xx}}{dy}(1-\nu) - \nu\frac{d\hat{\sigma}_{yy}}{dy} + 2\eta(1-\nu)\frac{d\hat{p}}{dy} \tag{58}$$

Recognizing that equilibrium (6) and (31) require $d\hat{\sigma}_{yy}/dy$ to vanish on $y = 0$ and substituting from equation (57) yields the desired condition:

$$2\pi\ \iota\kappa b_y(1-\mu)/s = \frac{d\hat{\sigma}_{xx}}{dy}(\kappa,0,s) + 2\eta\frac{d\hat{p}}{dy}(\kappa,0,s) \tag{59}$$

The solutions for A, B, and C in equations (17)–(20) can be determined by doubly transforming (26), (31), and (59), then substituting equations (17)–(20). Solving the resulting three equations for A, B, and C, substituting into equations (17)–(20) and manipulating in the manner of the solution for the shear dislocation yields:

$$\sigma = b_y\{(x/r^2) - \mu\ K(x,y,t)\} \tag{60}$$

$$\eta p = -\mu\ b_y\{(x/r^2) - K(x,y,t)\} \tag{61}$$

$$\tau = b_y(\iota y/r^2) - \mu\ b_y\left\{K(x,y,t) - K^{\#}(x,y,t) - 2\iota\frac{\partial^2 K^*}{\partial x\partial y} - 2\frac{\partial^2 K^*}{\partial y^2}\right\} \tag{62}$$

where the Laplace transforms of K, $K^{\#}$, and K^* are as follows:

$$\tilde{K}(x,y,t) = s^{-1}\int_0^\infty \exp[-(\kappa^2 + s/c)^{1/2}y]\sin(\kappa x)d\kappa \tag{63}$$

$$\tilde{K}^{\#}(x,y,t) = (2\iota c/s^2)\int_0^\infty \kappa(\kappa^2 + s/c)^{1/2}\exp[\iota\kappa(x+\iota y)]d\kappa \tag{64}$$

$$\tilde{K}^*(x,y,t) = (c/s)\tilde{K}(x,y,t) \tag{65}$$

Inversion of these expressions proceeds along the same lines as inversion of the corresponding integrals in the shear dislocation solution. The results are as follows:

$$K(x,y,t) = \mathrm{Re}\{W(x,y,t)/z\} \tag{66}$$

$$K^*(x,y,t) = 2ct(x/r^2)\{\mathrm{erfc}[y/(4ct)^{1/2}] - 2i^2\mathrm{erfc}[y/(4ct)^{1/2}]\} - y\ \mathrm{sgn}(x)\int_{y/r}^1 (1-u^2)^{1/2}\mathrm{erfc}[ur/(4ct)^{1/2}]du \tag{67}$$

$$K^{\#}(x,y,t) = 4ct\ z^{-3}w[z/4ct)^{1/2}] - 2\iota(4ct/\pi)z^{-2} \tag{68}$$

where the notation is the same as that used earlier. The final expressions for the stress and pore pressure are obtained by substituting equations (66)–(68) into equations (60)–(62) and carrying out the differentiations in equation (62):

$$\sigma = b_y\mathrm{Re}\{[1-\mu\ W(x,y,t)]/z\} \tag{69}$$

$$\eta p = -\mu\ b_y\mathrm{Re}\{[1-W(x,y,t)]/z\} \tag{70}$$

$$\tau = \iota b_y(y/z^2) - \mu b_y\{4ctz^{-3}[W(x,y,t) - w(z/(4ct)^{1/2})] + 2\iota(4ct/\pi)^{1/2}z^{-2}[1-\exp(-y^2/4ct)] - z^{-2}\mathrm{Re}[z\ W(x,y,t)]\} \tag{71}$$

These expressions reduce to the usual ones from ordinary elasticity in the limits $t \to 0$ (undrained response) and $t \to \infty$ (drained response). In the latter limit, $p = 0$ and the drained Poisson's ratio ν enters; in the former $m = m_o$ and the undrained Poisson's ratio ν_u enters.

Discussion

The similarity between the solutions for the shear dislocation (53)–(55) and the opening dislocation (69)–(71) suggests that they can be combined advantageously in a form analogous to that of complex variable elasticity. To this end, define the complex Burger's vector as

$$b = b_x + \iota b_y \tag{72}$$

Then the two solutions can be written compactly as follows:

$$\sigma = \mathrm{Im}\{bz^{-1}[1-\mu W(x,y,t)]\} \tag{73}$$

$$\eta p = -\mu\ \mathrm{Im}\{bz^{-1}[1-W(x,y,t)]\} \tag{74}$$

$$\tau = \iota z^{-2}\mathrm{Re}(b\bar{z}) - \mu\{\iota b4ctz^{-3}[w(z/(4ct)^{1/2}) - W(x,y,t)] + 2b(4ct/\pi)^{1/2}z^{-2}[1-\exp(-y^2/4ct)] + z^{-2}\mathrm{Im}[\bar{b}z\ W(x,y,t)] + 2\iota b_x w[z/(4ct)^{1/2}]\} \tag{75}$$

where $\bar{b} = b_x - \iota b_y$. For comparison, the solution for a shear dislocation on a permeable boundary and an opening dislocation on an impermeable boundary (Rice and Cleary, 1976) can be written in the same form:

$$\sigma = \mathrm{Im}\{bz^{-1}[1-\mu\ \exp(-r^2/4ct)]\} \tag{76}$$

$$\eta p = -\mu\ Im\{bz^{-1}[1-\exp(-r^2/4ct)]\} \tag{77}$$

$$\tau = \iota z^{-2}\mathrm{Re}(b\bar{z}) - \mu\{\iota b4ctz^{-3}[1-\exp(-r^2/4ct)] + z^{-2}\mathrm{Im}[b\bar{z}\ \exp(-r^2/4ct)]\} \tag{78}$$

(Rice and Cleary, 1976, display only the solution for the shear dislocation in polar coordinates, but the solution for the opening dislocation is extracted from their results for the complex stress functions).

It is of interest to compare the stresses induced by the dislocations on $y = 0$ for the various cases. For the shear dislocation on the impermeable plane and the opening dislocation on the permeable plane the tractions on $y = 0$ are as follows:

$$\sigma_{yy} + \iota\sigma_{xy} = (\iota b_x/x)\{1+\mu(4ct/x^2)[1-e^{-x^2/4ct}] - 2\mu\ e^{-x^2/4ct}\} + (b_y/x)\{1-\mu(4ct/x^2)[1-e^{-x^2/4ct}]\} \tag{79}$$

For comparison the tractions obtained from the Rice and Cleary (1976) solution are

$$\sigma_{yy} + \iota\sigma_{xy} = (\iota b_x/x)\{1-\mu(4ct/x^2)[1-e^{-x^2/4ct}]\} + (b_y/x)\{1+\mu(4ct/x^2)[1-e^{-x^2/4ct}] - 2\mu\ e^{-x^2/4ct}\} \tag{80}$$

Note that the spatial dependence of the tractions is the same for the opening dislocation and for the shear dislocation if $y = 0$ is permeable or if $y = 0$ is impermeable. The time dependence of the tractions does, however, depend significantly on whether $y = 0$ is permeable or impermeable. In both cases, the traction decays from a short time limit, corresponding to the usual elasticity expression based on the undrained value of Poisson's ratio, to a long time limit that is smaller by the factor $1 - \mu = (1 - \nu_u)/(1 - \nu)$. The time dependence at intermediate times is shown in Fig. 1, which plots

$$\frac{\sigma_{yy}(x,0,t) - \sigma_{yy}(x,0,\infty)}{\sigma_{yy}(x,0,0) - \sigma_{yy}(x,0,\infty)} \tag{81}$$

against $4ct/x^2$ for the opening dislocation on permeable and impermeable boundaries. As shown, the induced stress for the permeable boundary decays monotonically from the short-

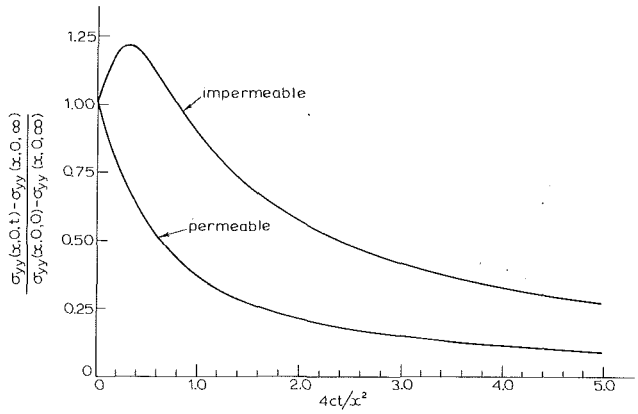


Fig. 1 Time-dependence of the normal traction σ_{yy} on $y = 0$ ahead ($x > 0$) of an opening dislocation at the origin. Results are shown for $y = 0$ permeable and impermeable to the diffusing species. The plot for the shear traction σ_{xy} ahead of a shear dislocation would be identical.

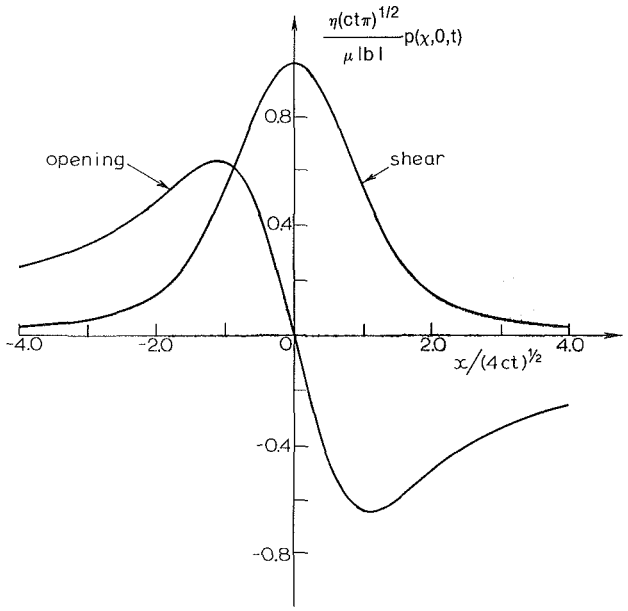


Fig. 2 Nondimensional pore pressure induced on an impermeable plane $y = 0$ by shear and opening dislocations at the origin. The plot is for a fixed time not equal to zero. The pore pressure for the shear dislocation is shown for $y = 0^+$; values for $y = 0^-$ are the negative of those shown.

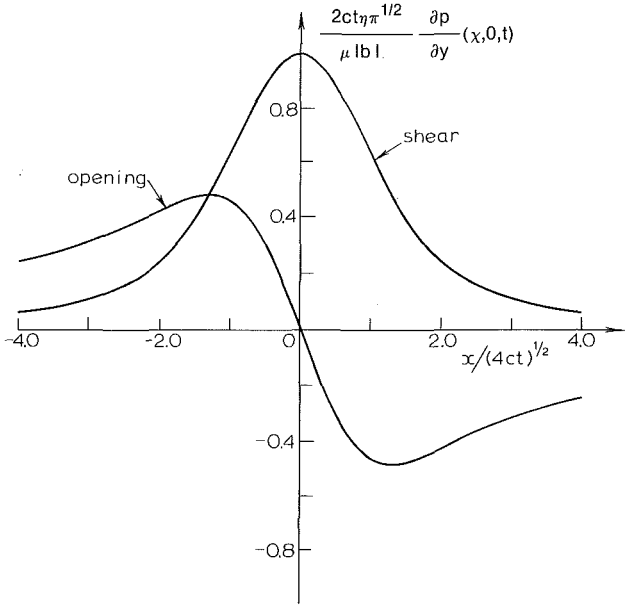


Fig. 3 Nondimensional gradient of pore fluid pressure (proportional to the negative of the fluid mass flux) on a permeable plane $y = 0$ due to shear and opening dislocations at the origin. The plot is for a fixed time not equal to zero. Values shown for the opening dislocation are for $y = 0^+$; those for $y = 0^-$ are the negative of those shown.

time undrained value to the long-time drained value. In contrast, the stress on the impermeable boundary first rises to a maximum that exceeds the undrained value by approximately 20 percent of the difference between the undrained and drained values. This maximum occurs at $4ct/x^2 \simeq 0.3$. A plot of σ_{xy} on $y = 0$ for the shear dislocation would be identical to Fig. 1.

As discussed by Rudnicki (1986), the increase of the shear stress predicted for the impermeable fault suggests that the effect of coupling between diffusion and deformation is initially destabilizing for sudden seismically emplaced slip. Also the differences in the time scale of shear stress decay for permeable and impermeable faults suggest differences in the effects of coupling on the reloading of faults, which has been proposed as a mechanism for aftershocks, and on processes preceding earthquakes.

Figure 2 shows the pore pressure in nondimensional form $\eta p(ct\pi)^{1/2}/\mu\,|b|$ induced on $y = 0$ by a shear dislocation and an opening dislocation on an impermeable plane. For the shear dislocation, the pore pressure is antisymmetric about $y = 0$. Consequently, the pore pressure is discontinuous on $y = 0$ and the values on $y = 0^-$ are the negative of those shown in Fig. 2. For the opening dislocation, the pore pressure is symmetric about $y = 0$ and, consequently, continuous on $y = 0$.

Figure 3 plots the gradient of pore pressure in nondimensional form $(2ct\eta\pi^{1/2}/\mu|b|)\partial p/\partial y$ on $y = 0$ induced by dislocations on a permeable plane. This quantity is proportional to the negative of the fluid mass flux across $y = 0$(4). As noted earlier, $\partial p/\partial y$ is antisymmetric about $y = 0$ for the opening dislocation and, hence, is discontinuous on $y = 0$. As shown in Fig. 3, $\partial p/\partial y$ is negative for $x > 0$ and positive for $x < 0$. Consequently, there is a net gain of fluid mass on $y = 0$ for $x < 0$ and a net loss on for $x > 0$. For the shear dislocation, $\partial p/\partial y$ on $y = 0$ is positive and symmetric about $x = 0$. Hence, fluid flows from the upper half-plane to the lower. The nature of the solutions and differences and similarities among them are further illustrated in Figs. 4–9. These figures plot contours of the pore pressure, mean stress, and the magnitude of τ in nondimensional form for the various solutions. These plots are all for a fixed time not equal to zero. Contours for the solutions due to Rice and Cleary (1976), that is, the shear dislocation on a permeable plane and the opening dislocation on an impermeable plane, are shown dashed.

Figures 4 and 5 plot contours of the nondimensional pore pressure $\eta p(4ct)^{1/2}/\mu\ |b|$ in the upper half plane. Figure 4 shows the contours for the shear dislocation ($b = b_x$) on permeable (dashed lines) and impermeable planes. The values in the lower half-plane are the negative of those shown. The contours coincide for large y, but differ near $y = 0$ because of the different boundary conditions there. As shown, the contours for the shear dislocations on an impermeable plane meet $y = 0$ at right angles as required by the boundary condition. Also, note that the maximum pore pressure change for the impermeable plane occurs at the origin whereas that for the permeable plane occurs at a finite value of y that increases with increasing time. Figure 5 shows the contours of nondimensional pore pressure for the opening dislocation on permeable and impermeable (dashed lines) planes. The pore pressure induced by an opening dislocation on an impermeable plane is identical to that for the shear dislocation on a permeable plane rotated 90 deg counterclockwise. As in the ordinary elasticity solution for shear and opening dislocations this feature applies to the entire stress and pore pressure

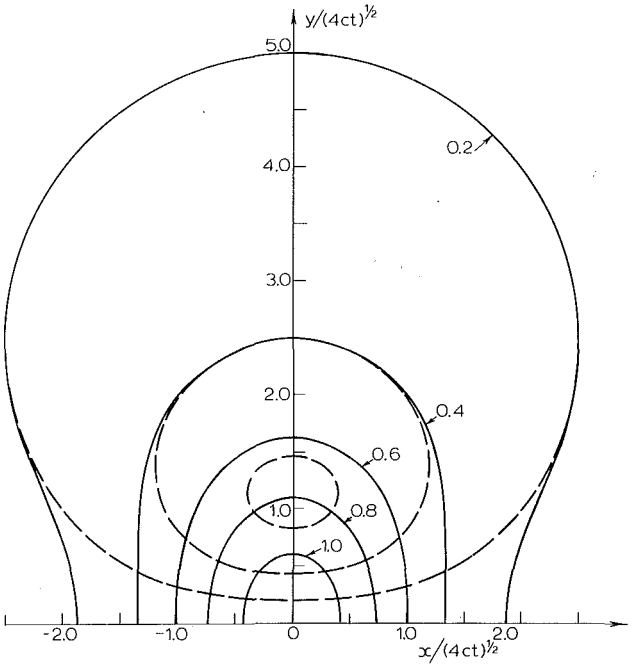


Fig. 4 Contours of nondimensional pore pressure $\eta\, p(4ct)^{1/2}/\mu\,|b|$ induced in $y \geq 0$ by a shear dislocation at the origin for the plane $y = 0$ impermeable and permeable (dashed)

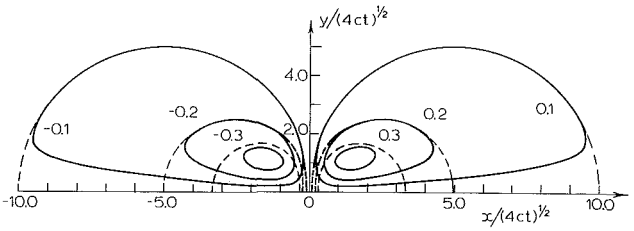


Fig. 5 Same as Fig. 4 for an opening dislocation. Dashed lines indicate the solution when $y = 0$ is impermeable.

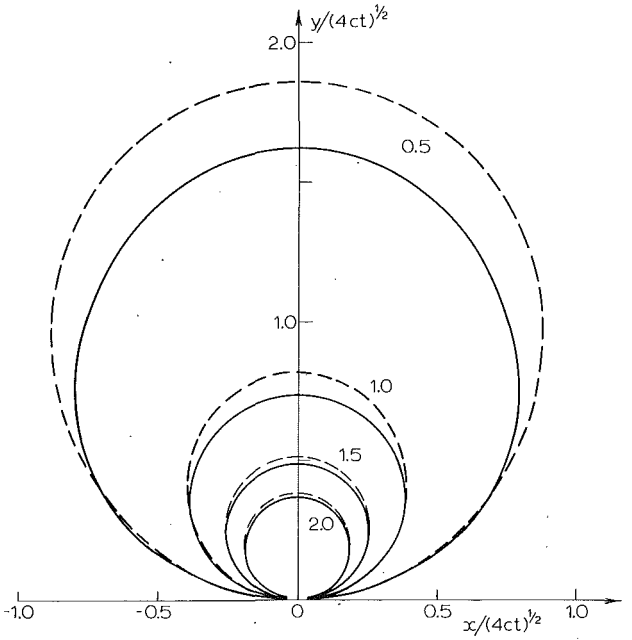


Fig. 6 Contours of nondimensional mean stress $\sigma(4ct)^{1/2}/|b|$ induced in $y \geq 0$ by a shear dislocation at the origin for the plane $y = 0$ impermeable and permeable (dashed)

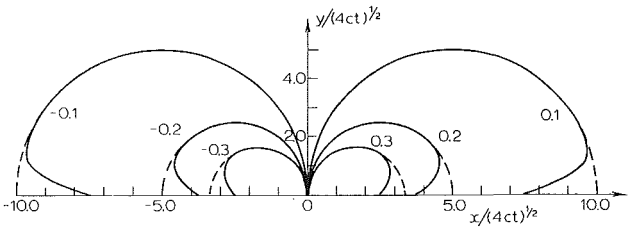


Fig. 7 Same as Fig. 6 for an opening dislocation. Dashed lines indicate the solution when $y = 0$ is impermeable.

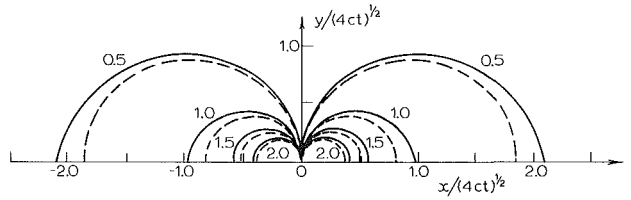


Fig. 8 Contours of nondimensional shear $|\tau|\,(4ct)^{1/2}/|b|$ induced in $y \geq 0$ by a shear dislocation at the origin for the plane $y = 0$ impermeable and permeable (dashed)

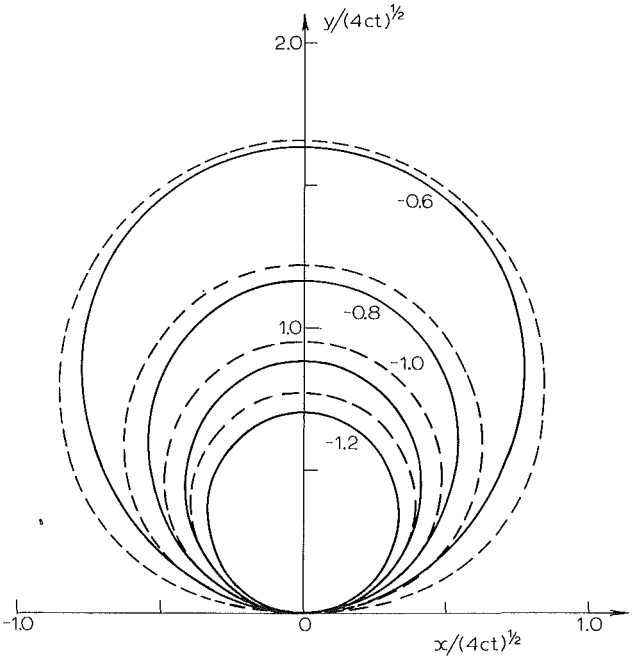


Fig. 9 Same as Fig. 8 for an opening dislocation. Dashed lines indicate the solution when $y = 0$ is impermeable.

fields: those for the opening dislocation on an impermeable plane can be obtained from the shear dislocation on a permeable plane by 90 deg counterclockwise rotation. As is evident from Figs. 4 and 5 the solutions with discontinuous pore pressure or fluid mass flux on $y = 0$ do not possess this property.

Contours of the mean stress, in the nondimensional form $\sigma(4ct)^{1/2}/|b|$ are shown in Fig. 6 for a shear dislocation ($b_x = 1.0$, $b_y = 0$) and in Fig. 7 for an opening dislocation ($b_x = 0$, $b_y = 1.0$) on permeable and impermeable planes. Figures 8 and 9 show contours of the magnitude of the shear stress in the nondimensional form $|\tau|(4ct)^{1/2}/|b|$ for shear (Fig. 8) and opening (Fig. 9) dislocations. In each plot, the two solutions shown approach the undrained solution ($t = 0$) far from the origin and the drained solution ($t \to \infty$) near the origin. The approach to these limits need not, however, be the same for the two solutions. This is the cause of the different positions

of the contours near the origin in Fig. 6 and for large values of y in Fig. 9.

Acknowledgment

I am grateful to J. Rice for suggesting sometime ago that I look at this problem and to E. Roeloffs for many helpful discussions.

This work was supported by the US Geological Survey and National Science Foundation Geophysics Program.

APPENDIX

This appendix describes some details of the inversion of the integrals $I(x, y, t)$ and $I^{\#}(x, y, t)$. The Laplace transforms of I and $I^{\#}$ are given by equations (41) and (44), respectively.

First consider the inversion of $I(x, y, t)$. Interchanging the order of the Laplace and Fourier inversions yields

$$I(x,y,t) = \int_0^\infty \kappa \cos(\kappa x) L^{-1}\left\{\frac{\exp[-(\kappa^2+s/c)^{1/2}y]}{s(\kappa^2+s/c)^{1/2}}\right\} d\kappa \qquad (A1)$$

Formulae (29.2.14) and (29.3.84) of Abramowitz and Stegun (1964) (hereafter abbreviated AS) yield the following result

$$L^{-1}\left\{\frac{\exp[-(\kappa^2+s/c)^{1/2}y]}{(\kappa^2+s/c)^{1/2}}\right\} = (c/\pi t)^{1/2}\exp[-\kappa^2 ct - y^2/4ct] \qquad (A2)$$

Formula (29.2.6) of AS can then be used to express the Laplace transform in equation ($A1$) as the following integral:

$$L^{-1}\left\{\frac{\exp[-(\kappa^2+s/c)^{1/2}y]}{s(\kappa^2+s/c)^{1/2}}\right\} = \int_0^t (c/\pi\lambda)^{1/2}\exp[-\kappa^2 c\lambda - y^2/4c\lambda] d\lambda \qquad (A3)$$

The integration can be accomplished by the change of variable $\lambda = \beta^2$ and the use of AS (7.4.33). The result is

$$L^{-1}\left\{\frac{\exp[-(\kappa^2+s/c)^{1/2}y]}{s(\kappa^2+s/c)^{1/2}}\right\} = (2\kappa)^{-1}\exp(-\kappa y)\{1 + \operatorname{erf}[\kappa(ct)^{1/2} - y/(4ct)^{1/2}] - e^{\kappa y}\operatorname{erfc}[\kappa(ct)^{1/2} + y/(4ct)^{1/2}]\} \qquad (A4)$$

Substituting into equation ($A1$), writing $\cos(\kappa x)$ in exponential form, changing variables, and using AS (7.4.36) then yields the final result, given by equation (45).

The inversion of $I^{\#}$ is lengthier, but proceeds along the same lines. Again interchange the order of the inversions. The Laplace transform can be inverted by using equation ($A3$) with $y = 0$ and (29.2.6) of AS. The result is

$$L^{-1}\{(c/s^2)/(\kappa^2+s/c)^{1/2}\} = [ct\kappa^{-1} - (2\kappa^3)^{-1}]\operatorname{erf}[\kappa(ct)^{1/2}] + \kappa^{-2}(ct/\pi)^{1/2}\exp(-\kappa^2 ct) \qquad (A5)$$

Now, $I^{\#}$ can be written as follows:

$$I^{\#}(x,y,t) = -\left\{\left(2ct\frac{\partial^2}{\partial x^2}+1\right)\int_0^\infty \exp(\iota\kappa z)\operatorname{erf}[\kappa(ct)^{1/2}]d\kappa + \iota(4ct/\pi)^{1/2}\frac{\partial}{\partial x}\int_0^\infty \exp(-\kappa^2 ct + \iota\kappa z)d\kappa\right\} \qquad (A6)$$

where $z = x + \iota y$. The remaining integrals can be done using (7.4.17) and (7.4.2) of AS:

$$\int_0^\infty \exp(-\kappa^2 t + \iota\kappa z)d\kappa = (\pi/4ct)^{1/2}\exp(-z^2/4ct)\operatorname{erfc}[-\iota z/(4ct)^{1/2}] \qquad (A7)$$

$$\int_0^\infty \exp(\iota\kappa z)\operatorname{erf}[\kappa(ct)^{1/2}]d\kappa = \iota z^{-1}\exp(-z^2/4ct)\operatorname{erfc}[-\iota z/(4ct)^{1/2}] \qquad (A8)$$

The final expression for $I^{\#}$ is given by equation (51).

References

Abramowitz, M., and Stegun, I. A., eds., 1964, *Handbook of Mathematical Functions, Appl. Math. Ser. 55*, National Bureau of Standards, Washington, D.C.

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

Biot, M. A., 1941b, "Consolidation Settlement under a Rectangular Load Distribution," *J. Appl. Phys.*, Vol. 12, pp. 426–430.

Biot, M. A., and Clingan, F. M., 1941, "Consolidation Settlement of a Soil with an Impervious Top Surface," *J. Appl. Phys.*, Vol. 12, pp. 578–581.

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

Booker, J. R., 1974, "Time-Dependent Strain Following Faulting of a Porous Medium," *J. Geophys. Res.*, Vol. 79, pp. 2037–2044.

Carslaw, H. S., and Jaeger, J. C., 1959, *Conduction of Heat in Solids*, 2nd Ed., Oxford University Press, Oxford.

Kuei, S., 1977, "Rheological Modeling of Synovial Fluid and Application of the Mixture Theory to Articular Cartilage," Ph.D. Thesis, Rensselaer Polytechnic Institute.

Mow, V. C., and Lai, W. M., 1980, "Recent Developments in Synovial Joint Biomechanics," SIAM Review, Vol. 22, pp. 275–317.

McNamee, J., and Gibson, R. E., 1960a, "Displacement Functions and Linear Transforms Applied to Diffusion Through Porous Elastic Media," *Quart. J. Mech. Appl. Math.*, Vol. 13, pp. 98–111.

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

Rice, J. R., 1979a, "The Mechanics of Quasistatic Crack Growth," in *Proceeding of the 8th U.S. National Congress of Applied Mechanics*, Kelley, R. E., ed., Western Periodicals, North Hollywood, CA, pp. 191–216.

Rice, J. R., 1979b, "Theory of Precursory Processes in the Inception of Earthquake Rupture," *Gerlands Beitr. Geophys.*, pp. 91–127.

Rice, J. R., and Cleary, M. P., 1976, "Some Basic Stress Diffusion Solutions for Fluid-Saturated Elastic Porous Media With Compressible Constituents," *Rev. Geophys. Space Phys.*, Vol. 14, pp. 227–241.

Rice, J. R., and Simons, D. A., 1976, "The Stabilization of Spreading Shear Faults by Coupled Deformation-Diffusion Effects in Fluid-Infiltrated Porous Materials," *J. Geophys. Res.*, Vol. 81, pp. 5322–5334.

Rice, J. R., and Rudnicki, J. W., 1979, "Earthquake Precursory Effects Due to Pore Fluid Stabilization of a Weakening Fault Zone," *J. Geophys. Res.*, Vol. 84, p. 2177–2193.

Rudnicki, J. W., 1985, "Effects of Pore Fluid Diffusion on Deformation and Failure of Rock," *Mechanics of Geomaterials, Proceedings of the IUTAM William Prager Symposium on Mechanics of Geomaterials: Rocks, Concrete, Soils*, Bazant, Z. P., ed., Wiley, New York, Chapter 15, pp. 315–347.

Rudnicki, J. W., 1986, "Slip on an Impermeable Fault in a Fluid-Saturated Rock Mass," *Earthquake Source Mechanics*, Das, J., Boatwright, J., and Scholz, C. H., eds., American Geophysical Union, Geophysical Monograph 37, pp. 81–89.

The Stability of a Dislocation Threading a Strained Layer on a Substrate

L. B. Freund
Division of Engineering,
Brown University,
Providence, RI 02912
Fellow ASME

The continuum theory of elastic dislocations is applied to estimate the critical thickness of a strained layer bonded to a substrate for a given mismatch strain. The formation of strained epitaxial layers is of interest due to their special electronic or optical properties, and critical thickness is understood to be the smallest thickness at which interface dislocations can form "spontaneously." The criterion invoked here is based on the work done by the layer stress in driving a threading dislocation to lay down a misfit dislocation along the layer-substrate interface, and it is applied in a way that leads to a result that is independent of the deflected shape of the threading dislocation. The general form of the dependence of critical layer thickness on mismatch strain is similar to that based on equilibrium dislocation analysis.

1 Introduction

Unique performance characteristics of electronic devices may be obtained by fabricating a composite semiconductor consisting of a thin layer or layers epitaxially grown onto a substrate. Because the materials are selected for reasons other than perfect match of their lattice spacing, some lattice mismatch must be accommodated at the layer-substrate interface. In simplest terms, the mismatch can be accommodated in either of two ways. One possibility is that the layer and substrate each retain their natural stress free crystalline structure except for sites within a few lattice spacings of the interface where an array of misfit dislocations exists to permit bonding. The other possibility is that the layer grows with a homogeneous strain of the magnitude necessary to bring the layer structure into perfect register with the substrate. The latter option is preferable in some applications in order to avoid undesirable electronic or other functional properties of the interface with misfit dislocations. Thus, the understanding of crystalline defects in strained coherent layers is of technological significance, as well as fundamental interest.

The existence of a critical layer thickness for epitaxial growth of a coherent layer on a substrate is well-established. That is, if the natural misfit between the substrate material and the layer material is sufficiently small, the first atomic layers to be deposited will be strained to match the lattice spacing of the substrate. As the layer becomes thicker, however, a point is reached at which alignment between the layer and the substrate is lost, presumably due to misfit dislocations introduced at the interface. The existence of the critical layer thickness was first proposed by Frank and van der Merwe (1949) on the basis of their analysis of a one-dimensional mechanical model and it was subsequently confirmed experimentally.

Theoretical studies of interface dislocations in strained epitaxial layers based on continuum mechanics have followed one of two types of conceptual approaches. These will be called simply the energy approach and the mechanistic approach. It should be noted at the outset that the two points of view are not independent nor are they in conflict. Indeed, a purpose of this paper is to discuss the physical phenomenon in a way that provides some unification of the two points of view.

The main idea in the energy approach is to consider the total (mechanical) potential energy of two possible configurations of the layer-substrate system, typically one with misfit dislocations and one without. Energy is usually expressed as "areal energy density" or average energy per unit area of interface. In the absence of misfit dislocations, the areal energy density is exactly the elastic strain energy stored in the layer due to homogeneous strain per unit area of interface. The comparison state is typically a configuration with one or more long straight dislocations lying at the interface. The configurational energy of the dislocations due to their proximity to the traction free surface of the layer, and possibly due to each other, is estimated and "averaged" over the interface. It is then argued on the basis of the minimum potential energy principle that the preferred configuration between the two is the one with the lower potential energy. The critical layer thickness is defined as that thickness at which the homogeneously strained layer configuration gives way to the configuration with misfit dislocations as the lower energy state. In this approach, there is no concern for the way in which the physical system transforms from one configuration to the other, or even if the required transformation is possible.

On the other hand, the main idea in the mechanistic approach is to identify a particular dislocation configuration and to calculate the driving force on the dislocation due to both

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 22, 1987; final revision, April 2, 1987.

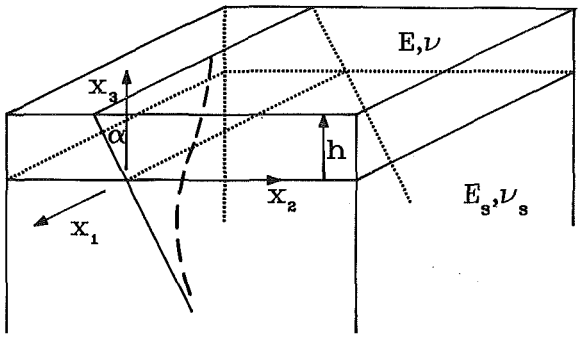


Fig. 1 Configuration of the layer-substrate system, showing a dislocation on a glide plane inclined at an angle α to the surface normal that is bowed out due to the layer strain

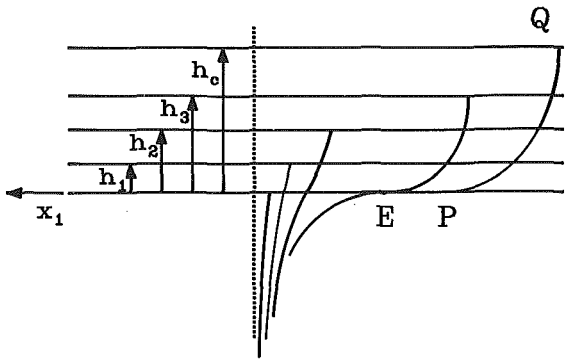


Fig. 2 Schematic of bowed dislocation configurations for increasing layer thickness

strain in the layer and image effects due to boundaries. For a given mismatch strain, the driving force depends on the layer thickness. The critical layer thickness is that thickness at which the driving force becomes large enough to lay down a line of misfit dislocation along the interface. The dislocation lines are typically curved in the postulated mechanistic models. Consequently, the stress field is a full three-dimensional field and the configurational forces on a dislocation can only be estimated on the basis of approximations.

The phenomenon may be approached on the basis of a framework other than continuum dislocation theory, of course. A notable example is the recent analysis reported by Dodson and Taylor (1986) of a discrete or atomistic model of strained layer epitaxy involving mismatched silicon-like materials. Through application of a Monte Carlo technique and a stability criterion, they estimated the critical layer thickness for a coherently strained structure. For small thicknesses, they reported *nonmonotonic* dependence of critical thickness on mismatch strain.

The purpose here is to re-examine the matter from the continuum dislocation point of view. The points to be made are (*i*) a work criterion for formation of interface dislocations is proposed that does not depend on the detailed shape of the bowed dislocation in the layer, and (*ii*) the resulting criterion is similar in general form to the more familiar equilibrium approach summarized by Matthews (1975). In several recent articles on the critical thickness phenomenon, the results of analysis leading to the dependence of critical thickness on mismatch strain were compared to an expression given by Matthews and Blakeslee (1974); see, for example, People and Bean (1985), and Dodson and Taylor (1986). However, the analysis of Matthews and Blakeslee was based on a strained layer superlattice, whereas the critical thickness expression cited by these later authors applies for an individual layer in the superlattice. As noted by Matthews and Blakeslee (1974, p. 124), the critical thickness for a layer in a superlattice (with layers of equal thickness) is four times the critical thickness for a single layer growing on a substrate. A factor of two arises from the fact that the mismatch strain between adjacent layers is shared equally, and a second factor of two arises because the dislocations must be bowed from both interfaces in the superlattice whereas they must be bowed only from the interface for a single layer.

2 A Representative Model

The formation of a misfit dislocation is considered here within the framework of the elastic continuum theory of dislocations. Strains are small enough so that the material response is adequately described by Hooke's law. A rectangular coordinate system x_1, x_2, x_3 is introduced. The substrate is an elastic half space with Young's modulus E_s and Poisson's ratio ν_s occupying $x_3 \leq 0$. The substrate is assumed to contain a dislocation. As a specific case, suppose that the glide plane of the dislocation contains the x_1 axis and it is inclined to the x_1, x_3 plane at an angle α measured positive with respect to rotation about the x_1 axis (see Fig. 1). Far from the substrate surface, the dislocation line is straight and in the x_2, x_3 plane, but it has some curvature near the free surface, in general. For the time being, suppose that the Burgers displacement vector has components $(0, -b \sin \alpha, b \cos \alpha)$ where b is the magnitude of the Burgers vector. Thus, far from the substrate surface, the dislocation is a pure screw dislocation. The substrate is stress free, except for the stress field induced by the dislocation.

Suppose that a strained layer with Young's modulus E and Poisson's ratio ν begins to grow on the surface of the substrate. For points on the interface far from where the dislocation line meets the interface, the layer strain is a uniform isotropic extension in the plane of the interface. Denote the extensional strain required for coherency by ϵ_o. This imposed strain induces a stress field with components

$$\sigma_{11} = \sigma_{22} = \frac{E}{1-\nu}\,\epsilon_o, \quad \sigma_{33} = 0. \tag{1}$$

Near where the dislocation meets the interface, the layer grows so as to extend the substrate dislocation into the layer. The layer thickness is assumed to be spatially uniform (see Fig. 1).

As the thickness of the layer increases, the shear traction on the glide plane in the layer due to the internal stress induces a configuration glide force on the dislocation. In general, a necessary condition for this to be so is that the inner product of the Burgers vector with the shear traction on the glide plane is nonzero. If it is positive (negative) the dislocation tends to advance (recede) along the glide plane. If it is zero, the dislocation is unaffected by the shear traction.

Consider the shape of the dislocation line in the glide plane, as shown schematically in Fig. 2 for several layer thicknesses. Even for a very thin layer (thickness h_1), the dislocation will deflect due to the layer stress further than it did due to the free surface alone. In the substrate, however, there is no driving force other than the force due to the curvature of the dislocation line in the layer, and the force due to the curvature tends to straighten the dislocation line. The substrate produces a retarding effect on the dislocation in the layer. As the layer becomes thicker (thickness h_2), the dislocation line deflects further to the right. It does so because, with $h_2 > h_1$, it can achieve a larger deflection without significantly increasing the curvature of the dislocation line at any point. While the dislocation is of pure screw type deep in the substrate, it is of mixed screw and edge type along the curved portion.

As the thickness is increased further to h_3, the deflection becomes large enough without the energetically unfavorable high curvature anywhere along its length so that it is tangent to the interface (at point E in Fig. 2). At point E, the dislocation is of pure edge type. The traction due to the stress in the layer holds it against the interface. Because there is no such stress in

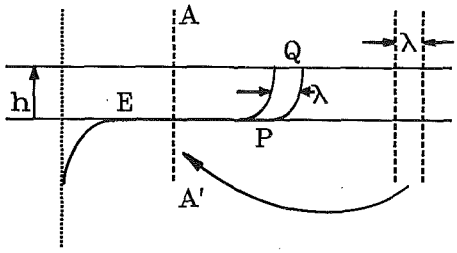


Fig. 3 Conceptual process for extending the length of misfit dislocation on the interface by a length λ

the substrate, however, the edge dislocation has no tendency to glide into the substrate. As the thickness is increased further, the edge segment along the interface is expected to become longer. Eventually, the thickness will become large enough so that the curved portion of the dislocation line (PQ in Fig. 2) in the layer can translate self-similarly, leaving a segment of misfit edge dislocation in its wake. The thickness at which this becomes possible is the critical thickness h_c, and a simple argument on a way to calculate h_c is given next. The actual calculation is carried out in Section 3.

Suppose the layer thickness is just large enough to advance the dislocation line along the layer, leaving a line of misfit edge dislocation behind. With reference to Fig. 3, consider the process of advancing the point P a distance λ to the right through self-similar translation of the segment PQ. Conceptually, the final state can be achieved by the following steps. First, a slab of thickness λ is cut out from the body far ahead of the point Q where the layer strain is essentially uniform. The faces of the slab are parallel to the plane $x_1 = 0$. The body is then cut along the plane AA' and the cut is opened uniformly to a gap distance of λ, thereby closing the gap far ahead of the dislocation without introducing dicontinuities. The uniform slab that has been cut out first is then dislocated to match the condition of the material in the interval EP and it is inserted into the remaining gap in that interval. The final state is just the state that would result from advancing PQ to the right a distance λ. However, the process of introducing slabs provides a basis for stating a condition on whether or not the dislocation produces a long segment of interface misfit dislocation at all.

The process of dislocating the slab that was cut out far ahead of the dislocation so that it fits into the interval EP requires that a certain amount of work be done on the slab. This work is the work of formation of a through-the-thickness glide edge dislocation on a plane inclined at an angle α ot the x_3 direction with the dislocation line at a distance h from the surface of the layer. Evidently, if this work is negative the dislocation will advance spontaneously along the glide plane, laying down a misfit dislocation in its path. If the work is positive, on the other hand, the dislocation will recede on its glide plane. The case when the work is zero is the critical case. Because the layer thickness is the only variable system parameter, the criterion of zero net work in forming the edge dislocation yields a condition for the critical layer thickness.

3 Calculation of the Critical Thickness

Let $E_s = E$ and $\nu_s = \nu$ for the time being. If the elastic moduli of the layer and substrate are indeed similar, then the effect of the free surface will be far greater than the effect of the interface on the dislocation. Suppose that the strained layer carries a self equilibrating isotropic tensile stress σ_o. Then the work to be computed is the work required to introduce the plane strain edge dislocation with Burgers displacement b shown in Fig. 4. The dislocation is introduced by first cutting the layer along the glide plane and applying tractions on the faces of the cut to hold the two faces together without slip. The normal traction

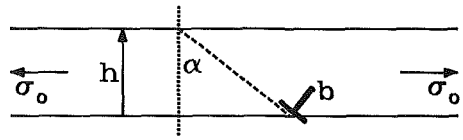


Fig. 4 Configuration on which the calculation of work of formation of the dislocation leading to equations (2) and (7) is based

and shear traction are uniform over the faces and are of magnitude $\sigma_o \cos^2 \alpha$ and $\sigma_o \sin \alpha \cos \alpha$, respectively. Next, *additional* tractions are applied to the faces of the cut that produce no normal offset of the faces, but that produce a unifrom shear offset of the faces equal to b. In the process, the normal tractions on the faces of the cut do no net work, the additional shear tractions do positive work on the body, and the tractions due to the initial strain in the layer do negative work.

The second of the two nonzero work contributions is calculated first. The uniform shear traction $-\sigma_o \sin \alpha \cos \alpha$ acts through the displacement b over the slip plane length $h \sec \alpha$. Thus, the work is

$$W_{\text{layer}} = -\sigma_o b h \sin\alpha \tag{2}$$

per unit thickness of the slab.

The work of the additional shear traction in forming the dislocation is computed next. Suppose that a dislocation with Burgers displacement b' is introduced on the glide plane as shown in Fig. 4 in an otherwise stress free half plane, and that the resulting shear stress on the glide plane (shown dashed in Fig. 4) is

$$\frac{b'}{h}\mu T(\xi/h,\nu), \quad 0 < \xi < h\sec\alpha \tag{3}$$

where $\mu = E/2(1+\nu)$ is the elastic shear modulus. The form of the stress distribution follows from dimensional considerations, linearity of the problem, and the fact that h is the only characteristic length in the model. Then the work that must be done by the additional shear traction on the faces of the cut to produce the offset b is

$$W_{\text{disl}} = \int_0^{h\sec\alpha - r_o} \int_0^b \frac{b'}{h}\mu T(\xi/h,\nu) db' d\xi$$

$$- \int_{-\pi}^{\pi} \frac{1}{2}(\sigma_{rr}^{\infty} u_r^{\infty} + \sigma_{r\theta}^{\infty} u_\theta^{\infty})_{r=r_o} r_o d\theta$$

$$= \frac{\mu b^2}{2}\int_0^{\sec\alpha - r_o/h} T(\eta,\nu)d\eta - \int_{-\pi}^{\pi} \frac{1}{2}(\sigma_r^{\infty} u_r^{\infty} + \sigma_{r\theta}^{\infty} u_\theta^{\infty})_{r=r_o} r_o d\theta \tag{4}$$

where r_o is the formal cut-off radius for the dislocation core. The second integral in equation (4) is the contribution due to introduction of the cut-off radius. It is computed by replacing the core of the dislocation with a cylindrical hole of radius r_o. The surface of the cylindrical hole is subjected to the appropriate tractions for formation of the dislocation with the cut at $\theta = \pi$. These tractions are, in fact, the tractions for an edge dislocation in an infinite medium along with the corresponding displacements. The contribution of the second integral in equation (4) to the total work is thus independent of the presence of the free boundary. The nondimensional function $T(\eta,\nu)$ is given by Freund and Barnett (1976) as

$$T(\eta,\nu) = \frac{1}{2\pi(1-\nu)}$$

$$\left\{ \frac{1}{\sec\alpha - \eta} + \frac{\sum_{n=0}^{5} c_n \eta^{5-n} \sec^n \alpha}{(\eta^2 + \sec^2\alpha + 2\eta\sec\alpha\cos 2\alpha)^3} \right\} \tag{5}$$

where the parameters c_n are given by

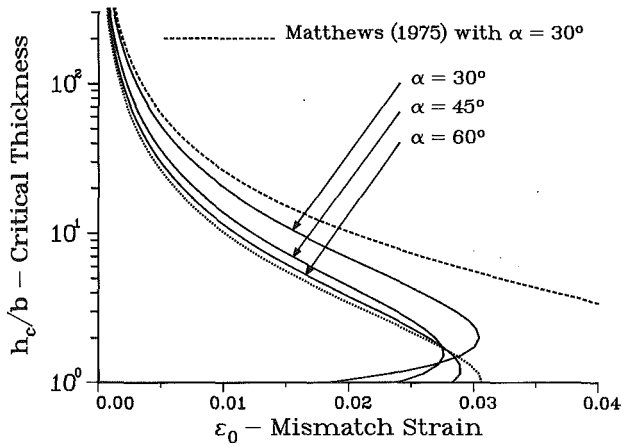


Fig. 5 Plot of calculated critical thickness versus mismatch strain for formation of a mismatch interface dislocation in edge orientation, with $\nu = 0.3$, $\alpha = 30$ deg, 45 deg, 60 deg, and $r_o = b$. The dashed curve is obtained from equation (13) with $\alpha = 30$ deg, and the dotted curve is obtained from equation (14).

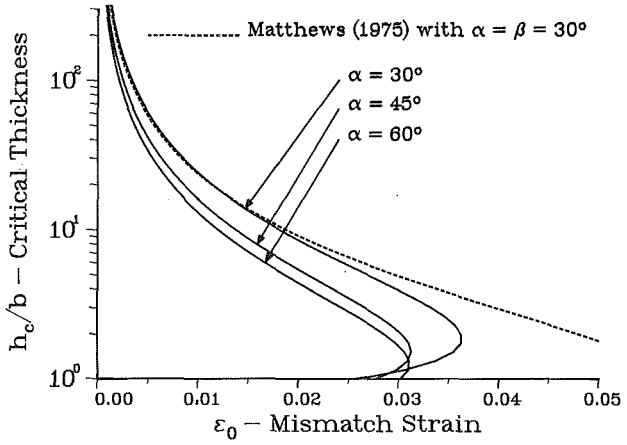


Fig. 6 Plot of calculated critical thickness versus mismatch strain for formation of a mismatch interface dislocation in mixed screw and edge orientation, with $\nu = 0.3$, $\beta = 30$ deg, $\alpha = 30$ deg, 45 deg, 60 deg and $r_o = b$. The dashed curve is obtained from equation (12) with $\alpha = 30$ deg.

$$c_0 = 1, \quad c_1 = -1 + 6\cos 2\alpha + 2\cos^2 2\alpha$$
$$c_2 = 6 + 2\cos 2\alpha + 6\cos^2\alpha, \quad c_3 = 6\cos 2\alpha + 4\cos^3 2\alpha \tag{6}$$
$$c_4 = -3 - 2\cos 2\alpha + 6\cos^2 2\alpha, \quad c_5 = 1 - 2\cos^2 2\alpha.$$

Evaluation of the integral in equation (4) leads to

$$W_{\text{disl}} = \frac{\mu b^2}{4\pi(1-\nu)}\left\{\ln\left(\frac{2h}{r_o}\right) - \frac{1}{2}\cos 2\alpha - \frac{(1-2\nu)}{4(1-\nu)}\right\}. \tag{7}$$

The first integral in equation (4) was evaluated numerically, and the simple analytic expression in equation (7) was evident from the result. This analytic expression was subsequently verified independently by means of a symbolic manipulation computer program. The second integral may be evaluated in terms of elementary functions, and its contribution is only the last term in equation (7). The contribution of this integral is discussed in a more general context by Gavazza and Barnett (1976). The configurational energy of the dislocation (the potential for force on the dislocation as a function of position) is discussed by Hirth and Lothe (1982) who note that the independence of the force of α is a feature of a more general result on dislocation image forces. In the present case, the total work of formation of the dislocation is of interest so that the contributions that are independent of h but dependent on α must be retained, even though they have no influence on the configurational force.

The condition that the total work $W_{\text{layer}} + W_{\text{disl}}$ is zero provides an equation for the critical layer thickness h_c, namely,

$$\frac{b}{8\pi(1+\nu)h_c\sin\alpha}\left\{\ln\left(\frac{2h_c}{r_o}\right) - \frac{1}{2}\cos 2\alpha - \frac{(1-2\nu)}{4(1-\nu)}\right\} = \epsilon_o. \tag{8}$$

For any given set of system parameters, this nonlinear equation may be solved for the critical layer thickness in the nondimensional form h_c/b. The result of calculations carried out with $\nu = 0.3$, $r_o/b = 1$, and $\alpha = 30$ deg, 45 deg or 60 deg are shown in Fig. 5. This figure shows graphs of h_c/b necessary to satisfy equation (8) as a function of misfit strain ϵ_o for the three values of α considered.

This analysis can be modified in a straightforward way in order to account for restrictions on the Burgers vector imposed by the crystallography. For a face centered cubic crystal structure with the interface being the [001] direction, for example, a candidate slip system for introduction of interface dislocations is slip on the (111) plane in the $[1\bar{1}0]$ direction or the $[10\bar{1}]$ direction. In this case, the inclination of the glide plane to the interface is specified by $\alpha = 30$ deg. Furthermore, the interface misfit dislocation that is formed has both edge and screw components. The edge and screw components of the Burgers vector are then $b_e = \pm b\cos\beta$ and $b_s = \pm b\sin\beta$, respectively, where $\beta = 30$ deg. The algebraic signs are determined by the sign of the mismatch strain and geometrical factors.

In this case, the work per unit length done by the stress in the film during formation of a misfit dislocation is again given by equation (2), except that the total Burgers displacement b is replaced by b_e. The film stress does no work as displacements in the screw direction occur. Thus,

$$W_{\text{layer}} = -\sigma_o b_e h\sin\alpha. \tag{9}$$

Work is done by tractions on the glide plane as the dislocation is formed, however, and the result equivalent to equation (7) above is

$$W_{\text{disl}} = \frac{\mu b_e^2}{4\pi(1-\nu)}\left\{\ln\left(\frac{2h}{r_o}\right) - \frac{1}{2}\cos 2\alpha - \frac{(1-2\nu)}{4(1-\nu)}\right\} + \frac{\mu b_s^2}{4\pi}\ln\left(\frac{2h}{r_o}\right). \tag{10}$$

As before, the critical condition for spontaneous formation of a line of misfit dislocation is $W_{\text{layer}} + W_{\text{disl}} = 0$ which takes the form

$$\frac{b\cos\beta}{8\pi(1+\nu)h_c\sin\alpha}\left\{\ln\left(\frac{2h_c}{r_o}\right) - \frac{1}{2}\cos 2\alpha - \frac{(1-2\nu)}{4(1-\nu)}\right\} + \frac{b(1-\nu)\sin\beta\tan\beta}{8\pi(1+\nu)h_c\sin\alpha}\ln\left(\frac{2h_c}{r_o}\right) = \epsilon_o \tag{11}$$

for the case at hand. The equivalent result obtained by Matthews (1975, p. 585) on the basis of dislocation equilibrium arguments is

$$\frac{b(1-\nu\sin^2\beta)}{8\pi(1+\nu)h_c\sin\alpha}\left\{\ln\left(\frac{h_c}{b}\right) + 1\right\} = \epsilon_o \tag{12}$$

in the present notation. The coefficient of the logarithmic term in equation (12) differs from that in equation (11) only by a factor $\cos\beta$ in the denominator. The critical thickness implied by equation (11) is shown in Fig. 6 for three values of α. In addition, the variation implied by equation (12) is plotted for $\alpha = \beta = 30$ deg.

4 Some Observations

The qualitative dependence of the critical layer thickness on misfit strain is as expected. The variation is quite similar to corresponding results obtained earlier by other methods. In-

deed, the equation for the critical layer thickness deduced here by the stability arguments is similar to the corresponding result obtained by Matthews (1975) on the basis of equilibrium methods. For the notation used here, his expression analogous to equation (8) above is

$$\frac{b}{8\pi(1+\nu)h_c\sin\alpha}\left\{\ln\left(\frac{h_c}{r_o}\right)+1\right\}=\epsilon_o. \tag{13}$$

The more commonly cited special case of this expression is

$$\frac{b}{8\pi(1+\nu)h_c}\left\{\ln\left(\frac{h_c}{b}\right)+1\right\}=\epsilon_o. \tag{14}$$

One essential difference between the present result and the result of the equilibrium calculation is in the nonlogarithmic term on the right of equation (7). This term is unimportant for relatively thick layers, but it is of the same order of magnitude as the logarithmic term for thin layers. Of course, the continuum dislocation theory is of limited applicability for very thin layers. If the force equilibrium criterion and the minimum energy criterion could both be applied on the basis of *exact* mathematical solutions of the elasticity boundary value problems, then the results would be identical. In the absence of such exact solutions, the stability approach has the advantage that the result does not depend on the *actual deflected shape* of the dislocations, but only on their ability to lay down a line of misfit dislocation along the interface. Furthermore, the development of the model on the basis of stability of a dislocation introduced from the substrate is not an essential feature of the result. As noted by others, for example, the stability of any dislocation loop expanding in the layer can be analyzed in the same way.

The results reported here differ significantly from the results of a theoretical study described by People and Bean (1985). They applied a certain energy criterion to the physical system in order to obtain a mathematical estimate of the critical layer thickness. They predict a dependence of h_c/b on the misfit strain ϵ_o that is far stronger than suggested by equation (8) or by the corresponding result due to Matthews and Blakeslee (1974). Their representation of the Matthews and Blakeslee result is incorrect, as has already been noted (People and Bean, 1986). Furthermore, their energy criterion has no apparent basis in the minimum potential energy principle of mechanics. The energy criterion that is introduced does not compare two actual or realizable energy states in order to determine the preferred state on the basis of the minimum principle. Instead, the two states considered are (*i*) the energy density of the uniformly strained layer and (*ii*) the energy of a dislocated but otherwise unstressed state. This comparison does not include the important interaction of the pre-existing stress field due to the layer strain with the forming dislocation. The work of this stress as dislocations are formed is, in fact, an essential element in the difference in energies of the two states.

In a recent article, Dodson and Taylor (1986) report a study based on the application of Monte Carlo methods to an atomistic model of a coherently strained layer. The model was intended to simulate the GeSi/Si structures. By invoking a stability criterion for determining the critical layer thickness, they showed that the dependence of the critical thickness on mismatch strain is very similar in form to the dependence found by the dislocation equilibrium or that given in equation (8) above for critical thicknesses greater than about 5 angstroms and mismatch strains less than about 4 percent.

A comparison between the critical thickness prediction based on this kind of analysis and experimental observations has been discussed by Matthews (1975). For metal films, the agreement is quite good. For other materials, including semiconductor materials, however, interface coherency persists up to thicknesses that exceed the predicted critical thickness by as much as an order of magnitude. The reasons for this discrepancy are not clear. In covalently bonded materials such as silicon or germanium, the resistance to dislocation glide may be sufficiently great so as to preclude the dislocation distortions presumed in the above analysis until the layer becomes much thicker than suggested by equation (8). On the other hand, the density of threading dislocations is often too low to account for the amount of strain relaxation observed upon loss of coherence. Consequently, a dislocation nucleation process may be required in modelling. In either case, the effects mentioned would tend to increase the estimate of critical thickness.

Acknowledgment

It is a pleasure to acknowledge helpful comments from Professor D. M. Barnett of Stanford University on this work, particularly on the way to handle the core region in deriving equation (7). The research support of the IBM Corporation and of the NSF Materials Research Laboratory at Brown University is gratefully acknowledged.

References

Dodson, B. W., and Taylor, P. A., 1986, "Atomistic Monte Carlo Calculation of Critical Layer Thickness for Coherently Strained Silicon-Like Structures," *Applied Physics Letters*, Vol. 49, pp. 642–644.

Frank, F. C., and van der Merwe, J. H., 1949, "One-Dimensional Dislocations. I. Static Theory," *Proceedings of the Royal Society*, Vol. A198, pp. 205–216; "II. Misfitting Monolayers and Oriented Growth," ibid, pp. 216–225.

Freund, L. B., and Barnett, D. M., 1976, "A Two-Dimensional Analysis of Surface Deformation Due to Dip-Slip Faulting," *Bulletin of the Seismological Society of America*, Vol. 66, pp. 667–675; ibid, pp. 2083–2084.

Gavazza, S. D., and Barnett, D. M., 1976, "The Self-Force on a Planar Dislocation Loop in an Anisotropic Linear-Elastic Medium," *Journal of the Mechanics and Physics of Solids*, Vol. 24, pp. 171–185.

Hirth, J. P., and Lothe, J., 1982, *Theory of Dislocations*, Wiley-Interscience, pp. 86–91, pp. 152–155.

Matthews, J. W., 1975, "Coherent Interfaces and Misfit Dislocations," *Epitaxial Growth*, Part B, Matthews, J. W., ed., Academic Press, pp. 559–609.

Matthews, J. W., and Blakeslee, A. E., 1974, "Defects in Epitaxial Multilayers," *Journal of Crystal Growth*, Vol. 27, pp. 118–125.

People, R., and Bean, J. C., 1985, "Calculation of Critical Layer Thickness Versus Lattice Mismatch for Ge_xSi_{1-x}/Si Strained Layer Heterostructures," *Applied Physics Letters*, Vol. 47, pp. 322–324; ibid, 49, 1986, p. 229.

I. Sheinman
Assoc. Prof.

M. Adan
Graduate Student.

Faculty of Civil Engineering,
Technion-Israel Institute of Technology,
Haifa, Israel 32 000

The Effect of Shear Deformation on Post-Buckling Behavior of Laminated Beams

A geometrical nonlinear theory of composite laminated beams is derived with the effect of transverse shear deformation taken into account. The theory is based on a high-order kinematic model, with the nonlinear differential equations solved by Newton's method and a special finite-difference scheme. A parametric study of the shear effect involving several kinematic approaches was carried out for isotropic and anisotropic beams.

I Introduction

Most of the previous research on composite structures is confined to linear problems and largely based on the classical thin plate theory which disregards the transversed shear deformation effect. The classical laminated beam theory, based on the Kirchhoff hypothesis (see Reissner and Stavski, 1961) has been shown to be quite adequate for thin laminates with a high span-to-thickness ratio. Due to the low transverse shear modulus relative to the inplane modulus of elasticity, the effect of shear deformation should be taken into account even for moderate span-to-thickness ratios. In some "linear" research works (e.g., Whitney and Pagano, 1970) the shear deformation effect is allowed for by means of the Mindlin kinematic model (Mindlin, 1951). Others like Chen and Sun (1985), Sirakumaran and Chia (1985), and Reddy and Chao (1985), extend the same Mindlin model to goemetrically nonlinear cases. High order models for the linear case, incorporating higher powers of the thickness, were developed by Nelson and Lorch (1974), Reissner (1975), and Lo et al. (1977). Reddy (1984a) used a modified higher-order model with the same number of unknowns as the Mindlin model, assuming no shear-strain coupling between γ_{yz} and γ_{xz}; more recently Phan and Reddy (1985), Reddy (1984b), and Putcha and Reddy (1986) applied the model in the nonlinear context, but his analysis concerned bending rather than postbuckling behavior. A completely different approach was adapted by Stein and Jagley (1985), who added trigonometric terms to the first terms of the power series.

The present work uses a higher order kinematic model in the post buckling context, with the shear deformation effect taken into account. The model is so constructed as to allow a variety of alternatives. The equilibrium equations and approximate boundary conditions are derived by applying the variational principle on the potential energy. The solution procedure is based on reduction of the nonlinear differential equations to a linear sequence, by a modification of Newton's method, and conversion to an algebraic one by a special finite-difference scheme eliminating the "locking" phenomenon which may occur in cases where the shear deformation effect is insignificant. A parametric study of the shear deformation effect and the accuracy of the kinematic model was carried out by applying the procedure to isotropic and anisotropic beams.

II Governing Equations

Kinematics. Consider a composite beam consisting of homogeneous orthotropic layers, of arbitrary orientation, with total thickness h. Let (x, z) be a rectangular coordinate system in the axial and thickness directions, respectively. The displacement field is assumed to be a cubic function of z:

$$u(x,z) = u^0(x) + \psi(x)z + \delta_2\xi(x)z^2 + \delta_3\phi(x)z^3$$
$$w(x,z) = w^0(x) \tag{1}$$

where u and w are the displacement functions in the x and z directions, respectively; u^0 is the displacement of the reference surface $z=0$ (not necessarily the midplane); $\psi(x)$ is the rotation about the normal to the $z=0$ plane; ξ and ϕ are additional functions of x which violate the assumption of planeness of the cross section and enable the transverse shear strain to be a parabolic function of the thickness coordinate; the parameter δ_2 and δ_3 are introduced for the purpose of investigating various alternatives. For example, with $\delta_2 = \delta_3 = 0$ we have the Mindlin (1951) model, which assumes that predeformation planes remain plane; here the transverse shear strain is assumed constant over the thickness, and a shear correction factor has to be used. Similarly, the model of Nelson and Lorch (1974) is obtained by setting $\delta_2 = 1$, $\delta_3 = 0$, that of Reissner (1975) $\delta_2 = 0$, $\delta_3 = 1$, and that of Lo et al. (1977) $\delta_2 = \delta_3 = 1$. It should be noted that with $\delta_3 = 1$ the shear strain is a parabolic function over the thickness and no correction factor

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13-18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted untol two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 6, 1986.

Paper No. 87-WA/APM-7.

is needed. There are no constraints on the shear stress (as in Reddy, 1984) at the beam top and bottom; the shear stress is expected to vanish automatically in application of the variational principle on the potential energy.

The kinematic relations can be written as:

$$\epsilon_{xx} = \epsilon^0_{xx} + K^1_{xx} z + K^2_{xx} z^2 + K^3_{xx} z^3$$
$$\gamma_{xz} = \gamma^0_{xz} + K^1_{xz} z + K^2_{xz} z^2 \tag{2}$$

where ϵ^0_{xx} and γ^0_{xz} denote the strain of the reference surface, K^i_{xx} and K^j_{xz} ($i = 1, 2, 3, j = 1, 2$) are the changes of curvature under deformation.

The Von-Karman strain of the reference surface and the change of curvature, associated with the displacement field (equations (1)) and imperfection function $Dw(x)$, can be written as:

$$\epsilon^0_{xx} = u^0_{,x} + \tfrac{1}{2}(w^0_{,x})^2 + w^0_{,x} Dw_{,x}$$
$$\gamma^0_{xz} = \psi + w^0_{,x}$$
$$K^1_{xx} = \psi_{,x}$$
$$K^2_{xx} = \delta_2 \xi_{,x}$$
$$K^3_{xx} = \delta_3 \phi_{,x}$$
$$K^1_{xz} = 2\delta_2 \xi$$
$$K^2_{xz} = 3\delta_3 \phi \tag{3}$$

$(\ )_{,x}$ denote the derivitive with respect to x.

Constitutive Equations. Under the classical laminate theory (i.e., for a single anisotropic equivalent layer) the force strain relations can be written as:

$$\begin{Bmatrix} N_{xx} \\ M^1_{xx} \\ M^2_{xx} \\ M^3_{xx} \end{Bmatrix} = \begin{bmatrix} A_{11} & B_{11} & C_{11} & D_{11} \\ B_{11} & C_{11} & D_{11} & E_{11} \\ C_{11} & D_{11} & E_{11} & F_{11} \\ D_{11} & E_{11} & F_{11} & G_{11} \end{bmatrix} \begin{Bmatrix} \epsilon^0_{xx} \\ K^1_{xx} \\ K^2_{xx} \\ K^3_{xx} \end{Bmatrix} \tag{4}$$

$$\begin{Bmatrix} Q_{xz} \\ M^1_{xz} \\ M^2_{xz} \end{Bmatrix} = \begin{bmatrix} A_{44} & B_{44} & C_{44} \\ B_{44} & C_{44} & D_{44} \\ C_{44} & D_{44} & E_{44} \end{bmatrix} \begin{Bmatrix} \gamma^0_{xz} \\ K^1_{xz} \\ K^2_{xz} \end{Bmatrix} \tag{5}$$

where

$$(N_{xx}, M^1_{xx}, M^2_{xx}, M^3_{xx}) = \int_A S_{xx}(1, z, z^2, z^3) dA$$
$$(Q_{xz}, M^1_{xz}, M^2_{xz}) = \int_A S_{xz}(1, z, z^2) dA \tag{6}$$

S_{ij} is the Kirchhoff stress tensor in the undeformed system. A_{ii}, B_{ii}, etc. ($i = 1, 4$) are the elasticity coefficients defined by

$$(A_{11}, B_{11}, C_{11}, D_{11}, E_{11}, F_{11}, G_{11}) = b \int_{h_B}^{h_T} \bar{Q}_{11}(1, z, z^2, z^3, z^4, z^5, z^6) dz$$
$$(A_{44}, B_{44}, C_{44}, D_{44}, E_{44}) = b \int_{h_B}^{h_T} \bar{Q}_{44}(1, z, z^2, z^3, z^4) dz \tag{7}$$

b is the beam width, $\bar{Q}_{11}$ and $\bar{Q}_{44}$ are the elastic stiffnesses transformed to the x directions. For a single layer with α orientation with respect to the x axis:

$$\bar{Q}_{11} = Q_{11}\cos^4\alpha + 2(Q_{12} + 2G_{12})\sin^2\alpha\cos^2\alpha + Q_{22}\sin^4\alpha$$
$$\bar{Q}_{44} = G_{13}\cos^2\alpha + G_{23}\sin^2\alpha$$

where

$$Q_{11} = E_{11}/(1 - \nu_{12}\nu_{21})$$
$$Q_{12} = Q_{11}\nu_{21}$$
$$Q_{22} = E_{22}/(1 - \nu_{12}\nu_{21})$$

Equilibrium Equations. The equilibrium equations and the appropriate boundary conditions are derived by applying the following variational principle:

$$\delta\pi = \int_x \{N_{xx}\delta\epsilon^0_{xx} + M^1_{xx}\delta K^1_{xx} + M^2_{xx}\delta K^2_{xx} + M^3_{xx}\delta K^3_{xx} + k(Q_{xz}\delta\gamma^0_{xz} + M^1_{xz}\delta K^1_{xz} + M^2_{xz}\delta K^2_{xz})\} dx - \int_x (n_{xx}(x)\delta u + q(x)\delta w + m(x)\delta\psi) dx = 0 \tag{8}$$

where n_{xx}, q and m are the external axial, transverse, and moment loading. k is the shear correction factor assigned to $k = 1$ when $\delta_3 = 1$.

Substituting equations (1), (4), and (5) in equation (8), integrating the latter by parts, we obtain the equilibrium equations:

$$N_{xx,x} = -n_{xx}$$
$$[N_{xx}(w^0_{,x} + Dw_{,x})]_{,x} + kQ_{xz,x} = -q$$
$$M^1_{xx,x} - kQ_{xz} = -m \tag{9}$$
$$\delta_2(M^2_{xx,x} - 2kM^1_{xz}) = 0$$
$$\delta_3(M^3_{xx,x} - 3kM^2_{xz}) = 0$$

with the following boundary conditions:

$$u = u^* \text{ or } N_{xx} = N^*$$
$$w = w^* \text{ or } N_{xx}(w_{,x} + Dw_{,x}) + kQ_{xz} = Q^*$$
$$\psi = \psi^* \text{ or } M^1_{xx} = M^{1*}_{xx} \tag{10}$$
$$\xi = \xi^* \text{ or } M^2_{xx} = M^{2*}_{xx}$$
$$\phi = \phi^* \text{ or } M^3_{xx} = M^{3*}_{xx}$$

where $(\ )^*$ denotes the given displacement and/or forces at the boundaries.

III Solution Procedure

A modification of Newton's method (Thurston, 1965), applicable to differential equations, is employed for reducing the nonlinear equilibrium equations (equations (9)), and the boundary conditions (equations (10)), to a linear sequence. Under this approach, the iteration equations are derived by applying to an approximate solution (initially taken as linear) a small correction obtained through solution of the linearized differential equations.

Taking the displacements and their derivatives with respect to the x coordinate as unknown dependent variables,[1] the unknown vector reads:

$$\{z\}^T = \{u, w, \psi, \xi, \phi, \bar{u}, \bar{w}, \bar{\psi}, \bar{\xi}, \bar{\phi}\} \tag{11}$$

where $(\bar{\ })$ denotes the first derivative of $(\ )$ with respect to x. By this means, the sequence is reduced to first order but the number of equations increases to ten.

In order to eliminate the "locking" phenomenon (see Introduction) a special "half-station" finite-difference scheme is adopted. This scheme consists of two interlaced distinct nets as shown in Fig. 1. All equations (equations (9))

[1] This choice was made because of the solution scheme, which is based on a finite-difference procedure.

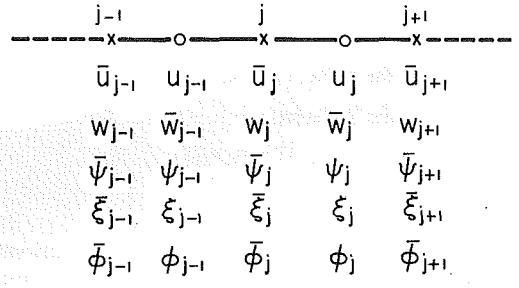


Fig. 1 Interlaced nets and corresponding unknowns

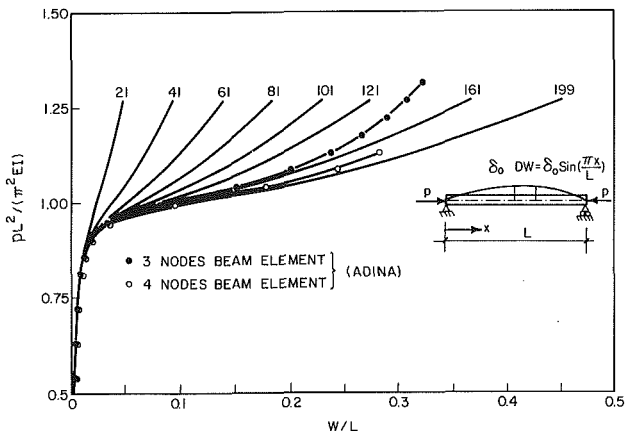


Fig. 2 Load-deflection convergence of w at midpoint ($\delta_2 = \delta_3 = 0$)

are written between the mesh points—except the second equation, which is written at the point itself, thereby achieving a high degree of accuracy even with relatively sparse nets. Differentiation at the boundaries is derived with the aid of fictitious points on either exterior side of the beam.

The differential equations are converted into an algebraic sequence which can be written in matrix form as follows:

Equilibrium equation:

$$C_j Z_{j-1} + B_j Z_j + A_j Z_{j+1} = G_j \quad j = 1,2, \ldots, N \tag{12}$$

Boundary conditions at the first point:

$$R_0^1 Z_0 + R_0^2 Z_1 + R_0^3 Z_2 = R_0^4 \tag{13}$$

Boundary conditions at the last point:

$$R_{n+1}^1 Z_{N-1} + R_{N+1}^2 Z_N + R_{N+1}^3 Z_{N+1} = R_{N+1}^4 \tag{14}$$

where N is the number of finite-difference points, Z_0 and Z_{n+1} are the unknown vectors at the fictitious end points. Equations (12) and (14) are an aglebraic sequence which is solved by a modification of Potter's method (Sheinman and Simitses, 1984).

IV Numerical Results and Discussion

For the procedure outlined above, a general computer program NABS (Nonliner Analysis of Beams with Shear Deformation) was written, covering nonlinear behavior of any laminated composite beam under arbitrary external loading and boundary conditions, as well as any geometrical initial imperfection. This program is especially suitable for parametric study of the effect of shear deformation and for investigating the accuracy of the given kinematic approach. Two examples (worked out on a VAX-750 digital computer) were used for illustrating the above methodology; (*a*) an isotropic beam and (*b*) an anisotropic carbon/epoxy laminated beam.

(*a*) Isotropic Beam. This example is reproduced from

Table 1 VAX-750 CPU-time as function of number of mesh points

No. of points	21	41	61	81	101	121	161	199	ADINA 4 Node ele.
CPU time (sec)	4	8	12	17	23	42	46	54	240
Max No. of iterations	5	6	6	8	12	15	19	24	40

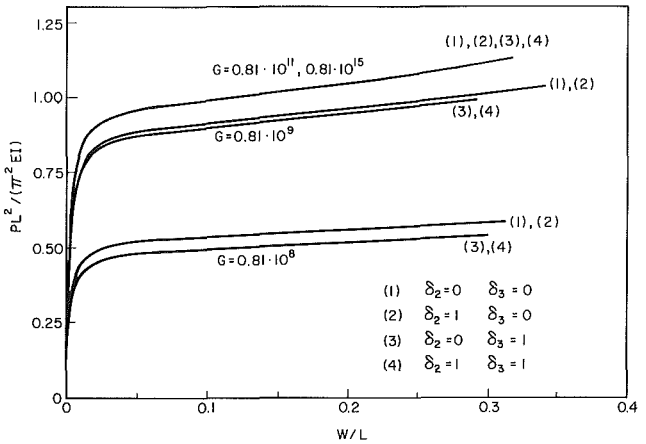


Fig. 3 Load-deflection curves for isotropic beam with different shear moduli

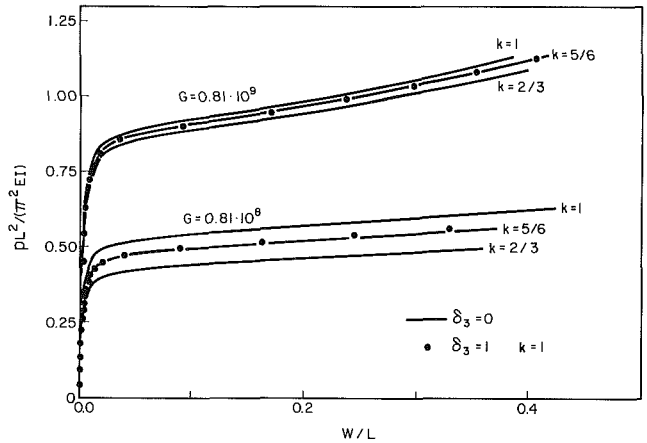


Fig. 4 Load-deflection curves for isotropic beam with different shear correction factors (k)

Sheinman (1982) and demonstrates convergence procedure as well as the accuracy of the various kinematic approaches. The data for the example are: Length $L = 4$ m; cross section area $A = 0.0032$ m^2; moment of inertia $I = 170.7 \cdot 10^{-8}$ m^4; modulus of elasticity is $E = 2.1 \cdot 10^{11}$ N/m^2; the initial imperfection is taken as $Dw(x) = \delta_0 \sin(\pi x/L)$ with $\delta_0/h = 0.1$.

In Fig. 2, the convergence of the solution with respect to the number of finite-difference mesh points, is shown for the kinematic approach of Mindlin's model ($\delta_2 = \delta_3 = 0$), with shear modulus $G = 0.81 \cdot 10^{11}$ N/m^2 and shear correction factor $k = 5/6$. From this figure it is clear that the convergence is a function of the load level. Up to $P/P_E = 0.85$, which is still within the linear region, convergence is achieved with 21 points; the higher the load-level, the larger the number of points needed for convergence. For very high levels, at which the lateral nondimensional midpoint displacement exceeds 0.15, 200 points are needed. This example was run also with 20 isoparametric 3 and 4-node beam elements using the ADINA code (Bathe, 1981). Convergence here also depends on the load level; the 3-node element is up to about $P/P_E = 1$, and

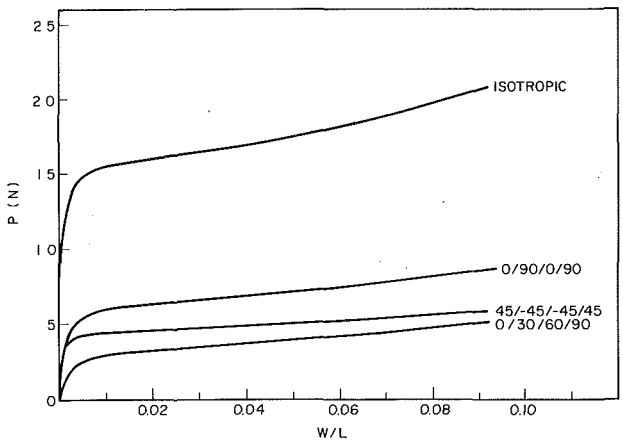


Fig. 5 Load-deflection curves for different stacking combinations, L/h = 200

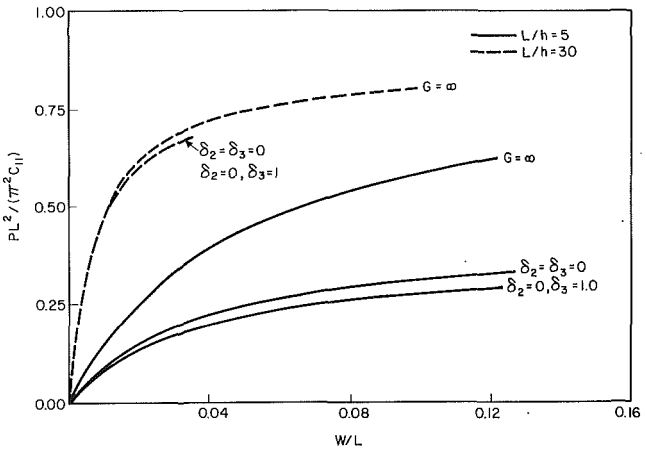


Fig. 6 Effect of kinematic approach on midpoint deflection w for carbon/epoxy laminated (0/90/0/90) beam

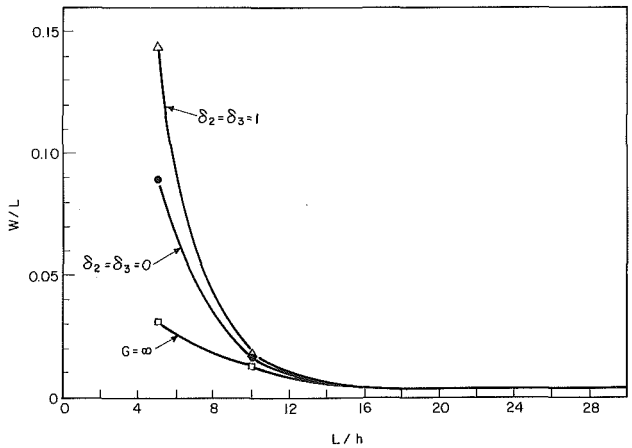


Fig. 7 Effect on L/h ratio on midpoint deflection w for carbon/epoxy (0/90/0/90) laminated beam subject to load level ($pL^2/(\pi^2 C_{11})$ = 0.3, under different kinematic approaches

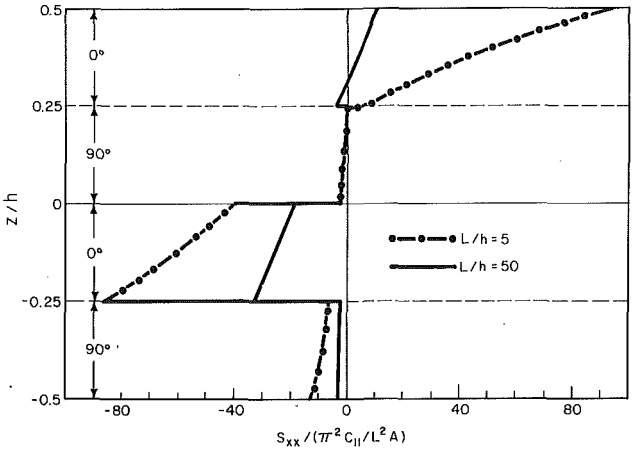


Fig. 8 Longitudinal stress at $x = l/2$ for ($pL^2A/(\pi^2 C_{11})$ = 0.32

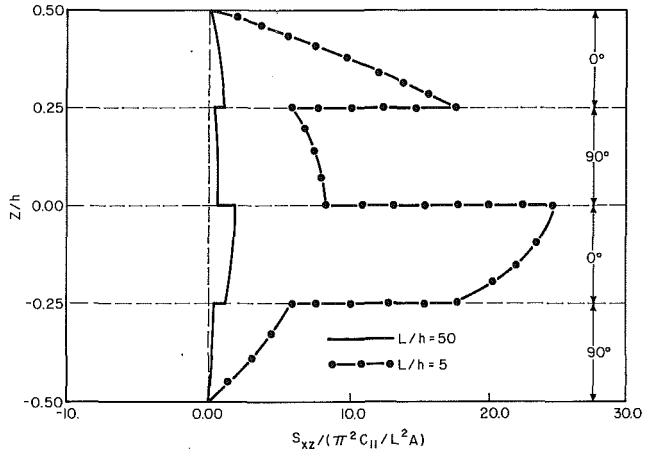


Fig. 9 Shear stress at x = 0 for ($pL^2A/(\pi^2 C_{11})$ = 0.32

the 4-node element up to P/P_E = 1.14 (see Fig. 2); at the level above 1.14 the 4-node element fails to converge with 40 iterations. The ADINA results are in good agreement with ours, but the CPU computer time is much longer (see Table 1).

In Fig. 3 load-deflection curves for different given shear moduli are plotted; it is seen that at low moduli there are significant differences between the various kinematic approaches, and an accurate approach is called for. Load-deflection curves for different shear correction factors are plotted in Fig. 4. It should be noted that while there is no need for correction (k = 1), under the kinematic approach δ_3 = 1 (since the shear stress vanishes at the top and bottom of the beam) correction is necessary with the Mindlin model ($\delta_1 = \delta_3 = 0$); the most accurate factor is seen to be k = 5/6.

(b) **Anisotropic Carbon/Epoxy Beam.** The data for this example are: 4-ply laminate with h_{ply} = 0.000125 m; h_{total} = 0.0005 m; $E_{11} = 1.4 \cdot 10^{11}$ N/m^2; $E_{22} = 0.1 \cdot 10^{11}$ N/m^2; $G_{12} = 0.1 \cdot 10^{10}$ N/m^2, ν_{12} = 0.34 and it assumed that: $G_{13} = G_{23} = 0.2 \cdot 10^{10}$ N/m^2. The initial imperfection is again $Dw(x) = \delta_0 \sin(\pi x/L)$ with δ_0/h = 0.1, and width of b = 0.01 m. In Fig. 5, load-deflection curves are plotted for L/h_{total} = 200 at different stacking combinations. The isotropic curve was obtained for $E_{22} = E_{11} = 1.4 \cdot 10^{11}$ N/m^2. The fact that indentical results were obtained for the Kirchhoff-Love ($G = \infty$), Mindlin ($\delta_2 = \delta_3 = 0$) and Lo [10] ($\delta_2 = \delta_3 = 1$) approaches at this span-to-thickness ratio, indicates that the effects of shear deformation is insignificant for this ratio. The effect of shear deformation and the accuracy of the kinematic approach as a function of L/h ratio were checked, for the stacking combination 0/90/0/90, and Fig. 6 shows that while the effect is still small for L/h = 30, it is very large for L/h = 5; the entire range of L/h ratios is covered in Fig. 7, with the conclusion that for this stacking combination, the effect of shear deformation is significant up to $L/h < 15$ (for isotropic materials, up to 5).

Significant differences were also seen between kinematic approaches, widening as the load level increases. In Fig. 8 and 9, the longitudinal stress at $x = L/2$, and shear stress at x = 0, respectively, are plotted for load level 0.32 for L/h = 50 and

$L/h = 5$. The calculated shear stress at the top and bottom of the beam is not zero as in Fig. 9, but still very small compared with the maximum stress.

V Conclusion

A nonlinear higher-order shear deformation theory and a solution procedure are presented for beams of arbitrary rectangular cross section. The nonlinear equations are written in a special form, whereby different kinematic approches can be considered. The differential equations are reduced to a linear sequence by a modification of Newton's method and converted into an algebraic sequence by applying a special finite-difference scheme, in which the "locking" phenomenon is eliminated.

The theory and solution procedure are general and suitable for investigating the effect of shear deformation. The procedure was applied for an isotropic and an anisotropic beam with a view to comparing the kinematic approaches. Of the principal findings, the following should be emphasized:

1. The convergence of the solution with respect to the number of finite-difference points, depends on the load level. The higher the latter, the larger the number of points needed.
2. Application of a regular central finite-difference scheme showed that the special scheme actually eliminates the "locking" phenomenon.
3. For the isotropic case with a low-order kinematic model, the shear correction factor of $k = 5/6$ yields the most accurate results.
4. The shear deformation is affected by the length-to-thickness ratio, both in the laminated and in an isotropic beam. For a laminated beam, where E/G ratio is relatively high, the L/h ratio at which the shear deformation is significant increases. For low L/h ratios, a higher-order kinematic model is necessary for accurate results.
5. With a more accurate kinematic model, the beam is characterized by higher flexibility.

Acknowledgments

The authors acknowledge the financial suport of the Loewengart Research Fund, and are indebted to Ing. E. Goldberg for editorial assistance.

References

Bathe, K. J., 1981, "ADINA: A Finite Element Program for Automatic Dynamic Incremental Nonlinear Analysis," Rep. AE 81-1, Sweden.

Chen, J. K., and Sun, C. T., 1985, "Analysis of Impact Response of Buckled Composite Laminates," *J. Composite Structures*, Vol. 3, pp. 97–118.

Lo, K. H., Christensen, R. M., and Wu, E. M., 1977, "A High-Order Theory of Plate Deformation—Part 1: Homogeneous Plates; Part 2: Laminated Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 99, pp. 663–676.

Mindlin, R. D., 1951, "Influence of Rotary Inertia and Shear on Flexural Motions of Isotropic Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 73, pp. 31–38.

Nelson, R. B., and Lorch, D. R., 1974, "A Refined Theory of Laminated Orthotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 96, pp. 177–183.

Phan, N. D., and Reddy, J. N., 1985, "Analysis of Laminated Composite Plates Using a Higher-Order Shear Deformation Theory," *Int. J. for Numerical Methods in Engineering*, Vol. 21, pp. 2201–2219.

Putcha, N .S. and Reddy, J. N., 1986, "A Refined Mixed Shear Flexible Finite Element for the Nonlinear Analysis of Laminated Plates," *Computers and Structures*, Vol. 22, No. 4, pp. 529–538.

Reddy, J. N., 1984a, "A Simple Higher-Order Theory of Laminated Composite Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 745–752.

Reddy, J. N., 1984b, "A Refined Nonlinear Theory of Plates with Transverse Shear Deformation," *Int. J. Solids and Structures*, Vol. 20, No. 9, pp. 881–896.

Reddy, J. N., and Chao, W., 1981, "Large-Deflection and Large-Amplitude Free Vibration of Laminated Composite Material Plates," *Computers and Structures*, Vol. 13, pp. 341–347.

Reissner, E., 1975, "On Transverse Bending of Plates, Including the Effects of Transverse Shear Deformation," *Int. J. of Solids and Structures*, Vol. 11, pp. 569–573.

Reissner, E., and Stavsky, Y., 1961, "Bending and Streching of Certain Types of Heterogeneous Aeolotropic Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, p. 402.

Stein, M., and Jagley, D. C., 1985, "Effect of Transverse Shearing on Cylindrical Bending Vibration, and Buckling of Laminated Plates," AIAA paper No. 85-0744-CP.

Sirakumaran, K. S., and Chia, C. Y., 1985, 'Large-Amplitude Oscillations of Unsymmetrically Laminated Anisotropic Rectangular Plates Including Shear Deformation, Rotary Inertia and Transverse Normal Stress," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 536–542.

Sheinman, I., 1982, "Large Deflection of Curved Beam with Shear Deformation," ASCE, Vol. 108, EM4, pp. 636–647.

Sheinman, I., and Simitses, G. J., 1984, "A Modification of Potter's Methods for Diagonal Matrices with Common Unknown," *Computers and Structures*, Vol. 18, No. 2, pp. 273–275.

Thurston, G. A., 1965, "Newton's Method Applied to Problems in Nonlinear Mechanics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 87, pp. 383–388.

Whitney, J. M., 1972, "Stress Analysis of Thick Laminated Composite and Sandwich Plates," *J. Composite Materials*, Vol. 6, pp. 426–440.

Whitney, J. M., and Pagano, N. J., 1970, "Shear Deformation in Heterogeneous Anisotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 1031–1036.

C. T. Tsai

O. W. Dillon, Jr.
Mem. ASME

Department of Engineering Mechanics,
University of Kentucky,
Lexington, KY 40506

Viscoplastic Buckling of Silicon Ribbon

Silicon ribbon grown by the dendritic web process passes through a rapidly changing thermal profile in the growth direction. This rapidly changing profile induces stresses which cause buckling. Based on a viscoplastic material response function (Haasen-Sumino model), the creep buckling behavior of the silicon ribbon is investigated. The lateral deflection speeds describing the viscoplastic buckling behavior are calculated. It is found that the deflections of some modes increase with time while others die out. The role of the residual stresses in viscoplastic buckling is examined.

1 Introduction

Silicon ribbon is being considered for use in terrestrial photovoltaic applications (Ciszek, 1985). There is considerable interest in increasing the (area) rate of productivity of sheet silicon in order to reduce the cost of photovoltaic power. One promising method for doing this is known as the dendritic web ribbon growing process. However, the thermal stresses caused by the combination of the rapidly changing thermal profile and wide ribbon induces buckling. This buckling is actually the major limitation to growing wider silicon ribbon (thin plate) in all existing industrial processes.

The thermal elastic buckling analysis of an initially flat silicon ribbon has been previously discussed (Dillon and De Angelis, 1984; Duncan et al., 1982; Seidensticker and Hopkins, 1980; Seidensticker and Schruben, 1984). The dislocation density in the solidified silicon ribbon is very low (i.e., below 10^5 cm^{-2}) so that the yield stress is also small. The thermal stresses generated from the rapidly changing thermal profile are, therefore, appreciably higher than the local yield stress, and then viscoplastic flow occurs. Industrially grown ribbon product frequently has permanent lateral ripples (Seidensticker and Hopkins, 1980; Seidensticker and Schruben, 1984), especially when growth of wide ribbon is attempted. This permanent deflection must be the result of the viscoplastic flow. Deviations from flatness clearly cause difficulty in later steps in the manufacturing sequence involved in changing ribbon into useable photovoltaic cells. The analysis of buckling of a thermal viscoplastic (i.e., silicon) cantilevered plate is carried out here in order to more accurately evaluate the importance of the viscoplastic effect in the web growing process and to provide guidance on how to grow flatter ribbon by preventing buckling.

Since the temperature in the ribbon changes rapidly and nonuniformly from the melting point to room temperature during the growth process, a nonuniform stress field and an inhomogeneous strain rate field are generated. The viscoplastic material has inplane residual stresses which cannot be considered in the elastic plate. These residual stresses can themselves cause buckling.

This paper is the adaptation of the methodology contained in Tsai (1985), and Tsai and Dillon (1987), to the more complex thermal profile associated with a specific industrial process. The second spatial derivative of the temperature distribution for this profile (and all others known to us) is much greater than the generic cases previously investigated. This results in higher stresses near the solid-melt interface and, therefore, increases the potential for viscoplastic effects to be larger.

A three-dimensional constitutive model of silicon material is given in Section 2. The governing equations and the boundary condition for viscoplastic buckling of the thin plate are discussed in Section 3. Solution procedures and numerical results are contained in Section 4. A discussion is given in Section 5, while Section 6 summarizes the results.

2 The Viscoplastic Material Model (Haasen-Sumino Model)

Haasen (1962, 1967) in Germany, and Sumino et al. (1978, 1979) in Japan developed an accurate one-dimensional material response function for silicon. This model matches the stress-strain and dislocation density-strain data obtained during one-dimensional tensile tests of silicon over a wide temperature range. It is found that the viscoplastic behavior of silicon is temperature sensitive and that data at different temperatures are correlated by the term, $\exp(-Q/kT')$, where Q, k are material constants and T' is the absolute temperature (Tsai, 1985; Dillon et al., 1986).

A three-dimensional response is assumed by generalizing the Haasen-Sumino one-dimensional material model and assuming that silicon is isotropic in both its elastic and plastic responses (Tsai, 1985; Dillon et al., 1986). Therefore, we assume that the material model for silicon is such that (Dillon et al., 1986)

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N. Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 5, 1986; final revision, March 2, 1987.
Paper No. 87-WA/APM-16.

$$\dot{\epsilon}_{ij}=\frac{1+\nu}{E}\,\dot{\sigma}_{ij}-\frac{\nu}{E}\,\dot{\sigma}_{kk}\delta_{ij}+\alpha\dot{T}\delta_{ij}+\dot{\epsilon}^{p1}_{ij}, \tag{1}$$

where

$$\dot{\epsilon}^{p1}_{ij}=fS_{ij}, \tag{2}$$

and where T is the temperature (°C), ν is Poisson's ratio, α is the coefficient of thermal expansion, and $\dot{\epsilon}_{ij}$, $\dot{\sigma}_{ij}$, δ_{ij} and $\dot{\epsilon}^{p1}_{ij}$ are the components of the total strain rate, stress rate, Kronecker delta function, and plastic strain rate tensor, respectively. In equation (1), E is the Young's modulus, S_{ij} are the components of the deviatoric stress tensor which are $S_{ij}=\sigma_{ij}-\sigma_{kk}\delta_{ij}/3$, and the viscosity f is (Dillon et al., 1986)

$$f=\frac{bk_oN_me^{-Q/kT'}\,(\sqrt{J_2}-D\sqrt{N_m})^m}{\sqrt{J_2}}, \tag{3}$$

where N_m is the mobile dislocation density; J_2 is the second invariant of the deviatoric stress tensor defined by $J_2=S_{ij}S_{ij}/2$. Haasen and Sumino demonstrated the need to consider the changes in the dislocation density in order to model silicon behavior. This is especially true at low values of N_m. The rate of generation of the dislocation density is (Dillon et al., 1986)

$$\dot{N}_m=\phi Kk_oN_me^{-Q/kT'}\,(\sqrt{J_2}-D\sqrt{N_m})^{m+\lambda}, \tag{4}$$

where $\dot{\epsilon}^{p1}_{ij}=\dot{N}_m=0$ if $\sqrt{J_2}-D\sqrt{N_m}<0$.
The values of b (3.8×10^{-10} m), K (3.4×10^{-4} m/Newton), k_o (8.58×10^{-4} $m^{3.2}$/s-Newton$^{1.1}$), m (1.1), D (7.84 Newton/m) and λ (1.0) are material constants and ϕ is a "shape factor" which is taken to be 0.1 for thin silicon sheets (Dillon et al., 1987). The value of Q and k are taken to be 2.17 eV and 8.617×10^{-5} eV/K here, and the Young's modulus for silicon is $E(T')=1.7^{11}-2.771\times10^4\times(T')^2$ Pascals (Hartzell, 1984). We assume that this model is applicable all the way to the melting temperature (1412°C) although it has only been experimentally verified as being valid to 1200°C. This implies that the dislocation mechanisms do not change in the temperature range between 1200°C and 1412°C. The use of equation (4) couples the thermal stress field and the dislocation density. As recently discussed in Dillon et al. (1986), one cannot neglect changes in N_m in photovoltaic materials. In turn this results in a more complex "rate effect" in the material response.

3 Analysis

The lateral deflection (buckling shape) grows with time due to the effect of viscoplasticity. We are only interested in buckling as the plate starts to deflect and do not treat any "postbuckling" problem here. The analysis is applicable to very thin plates and is limited to the development of threshold buckling predictions. The material properties are assumed to be those represented by equations (1)–(4) where the "viscoplastic flow" of silicon is spatially inhomogeneous due to the spatial variations of the temperature and the dislocation density. Based on the ideas in Tvergaard's work (1979) on the creep buckling of simply supported plates subjected to a constant inplane stress in one direction, the governing equation for the thermal creep buckling of a plate of the Haasen-Sumino material was derived in detail in Tsai (1985), Tsai and Dillon (1987), and is summarized in Appendix B to be

$$\frac{D_e}{h}\,\nabla^4\dot{w}^e=\sigma_{xx}\frac{\partial^2\dot{w}^e}{\partial x^2}+2\sigma_{xy}\frac{\partial^2\dot{w}^e}{\partial x\partial y}+\sigma_{yy}\frac{\partial^2\dot{w}^e}{\partial y^2}$$

$$+\frac{2fE}{3}\left(\sigma_{xx}\frac{\partial^2w^e}{\partial x^2}+2\sigma_{xy}\frac{\partial^2w^e}{\partial x\partial y}+\sigma_{yy}\frac{\partial^2w^e}{\partial y^2}\right), \tag{5}$$

where $D_e=\dfrac{Eh^3}{12(1-\nu^2)}$;

the x axis is taken to be along the growth direction and the y axis is in the width direction of the ribbon as shown in Fig. 1. The parameter w^e is the elastic part of the lateral deflection, σ_{xx}, σ_{xy}, and σ_{yy} are the inplane stresses which are entirely due to the thermal field and are obtained from the analysis of the prebuckling state as described in (Dillon et al., 1986), while h is the ribbon thickness. This equation involves only the elastic deflection w^e as the dependent variable as a matter of convenience. However, this does not mean that the plate is being considered as an elastic one.

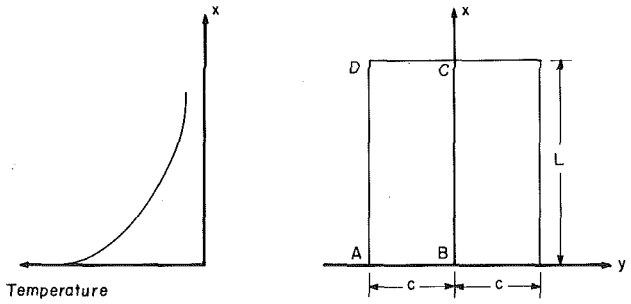


Fig. 1 Dimension of the ribbon and the schematic temperature variation along the growth direction (x) of the ribbon

In deriving equation (5) it was assumed that f is independent of the thickness coordinate. In addition to this, the usual assumptions of thin plate theory are used. The last group of terms in equation (5) represent the major effect of the viscoplastic material while the σ_{ij} are also different than the elastic ones. We consider $E(x)$ to depend on space as obtained from combining the expression given above for $E(T')$ and the thermal profile. We do not include spatial derivatives of $E(x)$, because earlier numerical work where they were retained, gave results similar to those obtained from equation (5).

The industrial process being modeled uses a take-up reel of large diameter to hold the finished ribbon. All other edges of the ribbon are free, except for surface tension at the solid-melt interface which we neglect.

The problem to be solved is, therefore, the thermal viscoplastic buckling (time dependent growth of an initially deformed) cantilever plate that is governed by equation (5).

The inplane stresses are due to the spatial variation of the temperature field and, therefore, are always self-equilibrated. They are evaluated from the equations given in detail in Dillon et al. (1987) and listed in Appendix A. The boundary conditions for the inplane stresses considered in Tsai and Dillon (1987) are different than those used here. The relations between the moments, shearing force and derivatives of w^e are precisely the same as those of elastic plate theory as can be seen in Appendix B.

The inplane stresses that exist near the solid-melt interface are very close to the elastic ones in numerical value in that region (near $x=0$) and are large and are due to the low number of the dislocations. However, due to the material being viscoplastic, a σ_{xx} residual stress field of significant magnitude exists farther away from the solid-melt interface. Since the stresses are self-equilibrated, this potentially changes the nature of the buckling from the purely elastic case.

The solution of equation (5) is assumed separable in the form

$$w^e(x,y,t)=g(t)\,W(x,y), \tag{6}$$

where $W(x, y)$ is the deflected shape of the plate and $g(t)$ is its amplitude at time t. By substituting equation (6) into equation (5) and then using the separation of variables approach, we obtain equations for the time dependence of g and the spatial dependence of the deflection shape W. They are

$$\dot{g}(t)-\varphi\,g(t)=0 \tag{7}$$

and

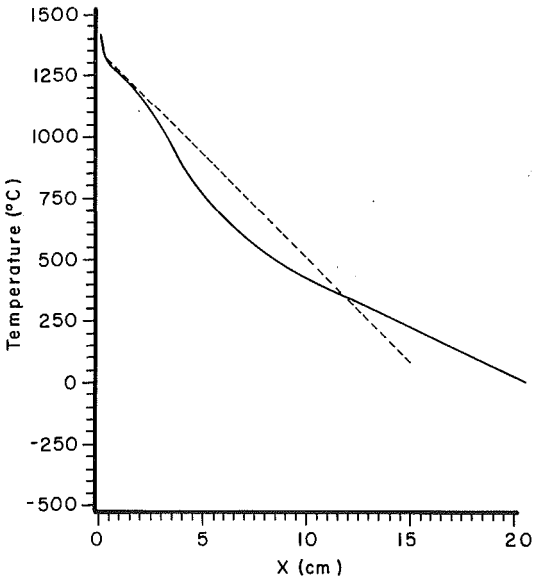


Fig. 2 The approximate web thermal profile given by equation (13) (solid curve) and equation (14) (dashed curve)

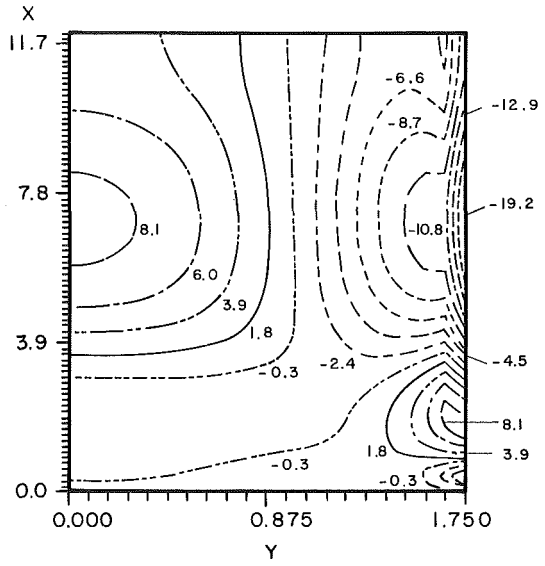


Fig. 3 The σ_{xx} stress (MPa) induced in a 3.5 cm × 12 cm ribbon by the profile given in equation (13), the end x = 12 cm is subjected to the residual stress boundary conditions; initial dislocation density = 13 cm^{-2}

$$D_e \nabla^4 W = \left(1 + \frac{2fE}{3\varphi}\right) h \times$$

$$\left(\sigma_{xx} \frac{\partial^2 W}{\partial x^2} + 2\sigma_{xy} \frac{\partial^2 W}{\partial x \partial y} + \sigma_{yy} \frac{\partial^2 w}{\partial y^2}\right), \tag{8}$$

where φ is the separation parameter. The solution of the time dependent equation (7) is (Tsai, 1985; Tsai and Dillon, 1987)

$$g(t) = g^o e^{\varphi t}, \tag{9}$$

where g^o is the magnitude of the increment in the deflection of the plate due to the thermal stresses being applied to an elastic material.

The inelastic behavior results in viscoplastic buckling which is governed by the same spatial equation as classical elastic thin plates, but where the inplane stresses are replaced by $(1 + 2fE/3\varphi)\sigma_{ij}$ as shown in equation (8). For the special case where σ_{xx} and σ_{yy} are constant, the temperature is uniform, f is constant and the plate simply supported, equation (8) is readily solved for φ (Tsai and Dillon, 1987; Trevgaard, 1979). In the present context $f(x, y)$ is a complex function of space, the term $2fE/3\varphi$ can vary in magnitude from 10^6 to 10^{-16} in the same problem. Hence the term $(1 + 2fE/3\varphi)$ changes sign several times which then has obvious implications for buckling.

As will be developed below, higher modes are needed to capture the significant physical phenomena. Due to the complex nature of the spatial variation of the term $(1 + 2fE/3\varphi)\sigma_{ij}$ mentioned above, it is unlikely that shapes $W(x, y)$ can be found which are consistent with φ being constant. Hence a Galerkin procedure is first used on equation (8) in order to evaluate φ, where the $W(x, y)$ that is used is an assumed deflection shape.

The shape functions $W_m(x, y)$ used here are the mth mode shapes of the buckled elastic cantilevered plate that are associated with

$$D_e \nabla^4 W = h\sigma_{ij} \frac{\partial^2 W}{\partial x_i \partial x_j}. \tag{10}$$

The mth eigen-parameter in equation (10) is the thickness h_m^2, and W_m are the corresponding eigenvectors of this equation. A particular $W_m(x, y)$ is substituted into equation (8) and the result multiplied by W_m. These products are then integrated over the entire plate. The result is rearranged as

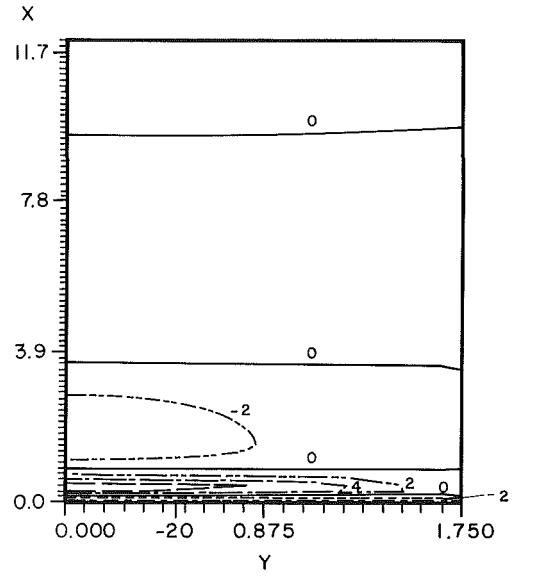


Fig. 4 The σ_{yy} stress (MPa) induced in a 3.5 cm × 12 cm ribbon by the profile given in equation (13), the end x = 12 cm is subjected to the residual stress boundary conditions; initial dislocation density = 13 cm^{-2}

$$\varphi = \frac{h_m^2}{h^2 - h_m^2} \frac{S1}{S2}, \tag{11}$$

where h_m is the critical thickness of the mth "pseudo" elastic mode from equation (10) and h is the actual plate thickness, and

$$S1 = \int_o^a \int_o^b \frac{2fE}{3} (\nabla^4 W_m) W_m dxdy, \tag{12a}$$

$$S_2 = \int_o^a \int_o^b E(\nabla^4 W_m) W_m dxdy, \tag{12b}$$

For numerical work described below, h is taken as $1.1h_1$.

4 Numerical Results

The thermal profile for the web growth process is approximately expressed as (Seidensticker, 1984)

Table 1 Typical results for residual stress boundary conditions with ribbon length = 12 cm

initial dislocation density (cm^{-2})		13	6	1
critical ribbon width (cm)		3.5	3.9	4.6
maximum σ_{xx} (MPa)	com.	-20.7	-23.8	-27.5
	ten.	8.66	9.88	11.7
maximum σ_{yy} (MPa)	com.	-22.5	-22.0	-21.5
	ten.	8.12	8.99	10.5
max. residual σ_{xx} (MPa)	com.	-11.5	-12.8	-13.2
	ten.	4.48	4.03	2.91
final dislocation density (cm^{-2})		2310	2730	3150

Table 2 Creep buckling results for a 3.5 cm × 12 cm ribbon (residual stress boundary conditions)

mode	critical thickness (mm)	φ (sec^{-1})
1	0.233	-0.694
2	0.181	-0.00280
3	0.155	-0.00821
4	0.141	-0.00000629
5	0.130	0.00225
6	0.125	-0.0000399
7	0.110	-0.0134
8	0.105	-0.0000565
9	0.102	0.0316
10	0.0957	-0.00350
11	0.0944	0.000193
12	0.0834	0.00335

$$T(x) = 1600e^{-0.0827x} + 85e^{-5x}\cos(\pi x) + 75\sin(\pi x) - 273, \text{ if } 0 \le x < 4;$$

$$T(x) = 1600e^{-0.0827x} + 85e^{-5e}\cos(\pi x) - 35\sin\frac{\pi(x-4)}{6} - 273, \text{ if } 4 \le x < 10;$$

$$T(x) = -39.545x + 822.229, \text{ if } 10 \le x < 20.1;$$

$$T(x) = 27, \text{ if } x \ge 20.1, \quad (13)$$

where x (cm) is the position along a ribbon as measured from the solid-melt interface, $x = 0$. The temperature $T(0) = 1412$ °C is the melting point of silicon. Although the industrial profile also varies along the y direction, its variation is small so that only the x dependent thermal profile is considered here. The temperature distribution in the ribbon given by equation (13) is shown as the solid curve in Fig. 2.

The inplane stresses used in the buckling analysis are obtained from a prebuckling analysis. Typical σ_{xx} and σ_{yy} stress distributions obtained from solving the equations in Appendix A are illustrated in Figs. 3 and 4 for a 12 cm long and 3.5 cm wide ribbon with an initial dislocation density of 13 cm^{-2}. This plate is subjected to the thermal profile of equation (13) and the residual stress boundary condition at $x = 12$ cm. The stresses existing at $x = 12$ cm are the residual stresses for this case. As can be seen from Figs. 3 and 4, the σ_{xx} stresses are large while the value of σ_{yy} are nearly zero along $x = 12$ cm.

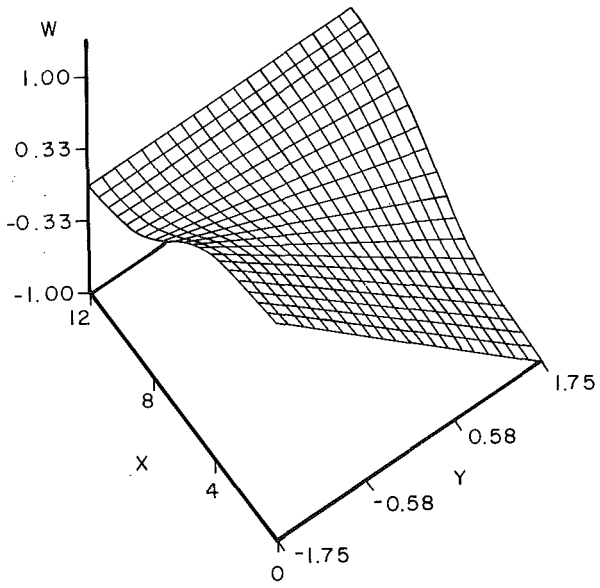


Fig. 5 The first buckling mode shape induced in a 3.5 cm × 12 cm ribbon by the profile given in equation (13), the end $x = 12$ cm is subjected to the residual stress boundary conditions; initial dislocation density = 13 cm^{-2}; the critical thickness = 0.233 mm

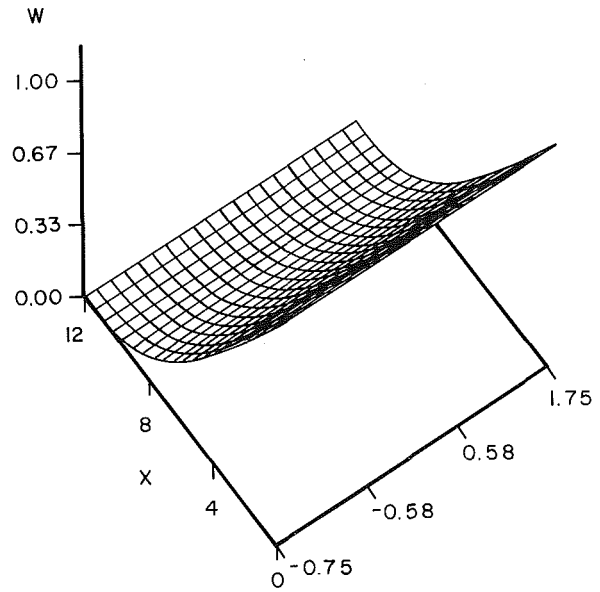


Fig. 6 The second buckling mode shape induced in a 3.5 cm × 12 cm ribbon by the profile given in equation (13), the end $x = 12$ cm is subjected to the residual stress boundary conditions; initial dislocation density = 13 cm^{-2}; the critical thickness = 0.181 mm

Typical inplane stress and final dislocation density values obtained from using the residual stress boundary conditions for differential initial dislocation densities and ribbon widths are listed in Table 1.

Equation (10) is solved by the finite element method, using a sixteen degree-of-freedom Hermitian-conforming rectangular element. The general procedure was discussed in (Tsai, 1985), where a computer code was developed for a ribbon divided into 20 × 20 elements for the calculations associated with buckling. Once the thickness and the corresponding deflection shape W_m are obtained from the solution of equation (10), the value of φ can be calculated from equation (11).

For a 12 cm long and 3.5 cm wide ribbon subjected to the thermal profile of equation (13) and the initial dislocation den-

Table 3 Creep buckling results for a 3.9 cm × 12 cm ribbon (residual stress boundary conditions)

mode	critical thickness (mm)	φ (sec^{-1})
1	0.255	-0.0357
2	0.235	-0.00971
3	0.167	-0.00406
4	0.155	-0.0000235
5	0.139	0.00398
6	0.138	-0.000185
7	0.123	-0.00375
8	0.112	-0.00634
9	0.111	0.00000473
10	0.100	0.0000590
11	0.0995	-0.00236
12	0.0863	0.000155

Table 4 Creep buckling results for a 4.6 cm × 12 cm ribbon (residual stress boundary conditions)

mode	critical thickness (mm)	φ (sec^{-1})
1	0.279	-0.000813
2	0.259	-0.00317
3	0.185	-0.00689
4	0.173	-0.0000116
5	0.158	0.0108
6	0.1484	-0.00776
7	0.1483	-0.00042
8	0.120	0.00193
9	0.119	0.0000863
10	0.107	0.000624
11	0.106	-0.00122
12	0.0984	0.000663

sity of 13 cm^{-2}, thickness results and the associated value of φ are given in Table 2 for several modes. The thickness of the first mode (twisting) is 0.233 mm, which is seen to have a negative value of φ. Its mode shape is shown in Fig. 5. The second mode (bending) also has a negative value of φ. Its shape is shown in Fig. 6. The ninth mode (bending and curling) has the maximum positive value of φ which makes this mode the most likely one to be seen in the experimental situation. Its shape is shown in Fig. 7. The thickness of the first mode is 0.328 mm when the ribbon length increases to 18 cm, and the associated thickness is 0.201 mm when the length decreases to 9 cm. Other results for the cases of 3.9 cm and 4.6 cm wide ribbons are listed in Tables 3 and 4.

In the calculations described above the spatial shape functions used to calculate φ, via the Galerkin procedure, were the mode shapes of the "pseudo" elastic problem, equation (10).

To obtain better estimates of the deflection, we now specify w^e as

$$w^e = g^o e^{\varphi t} W^*(x,y),$$

where φ is known from the previous step. Substitution of this deflection into equation (5) again produces equation (8) but with a "known" value for φ. New mode shapes are then calculated from equation (8). In turn these new mode shapes

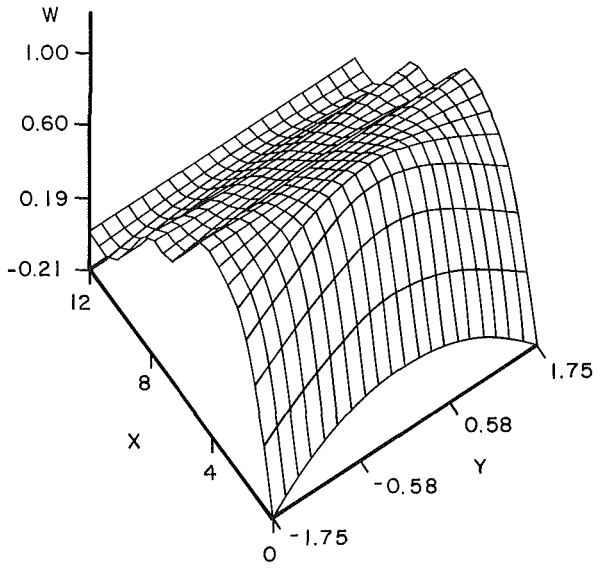


Fig. 7 The ninth buckling mode shape induced in a 3.5 cm × 12 cm ribbon by the profile given in equation (13), the end x = 12 cm is subjected to the residual boundary conditions; initial dislocation density = 13 cm^{-2}; the critical thickness for this mode = 0.102 mm

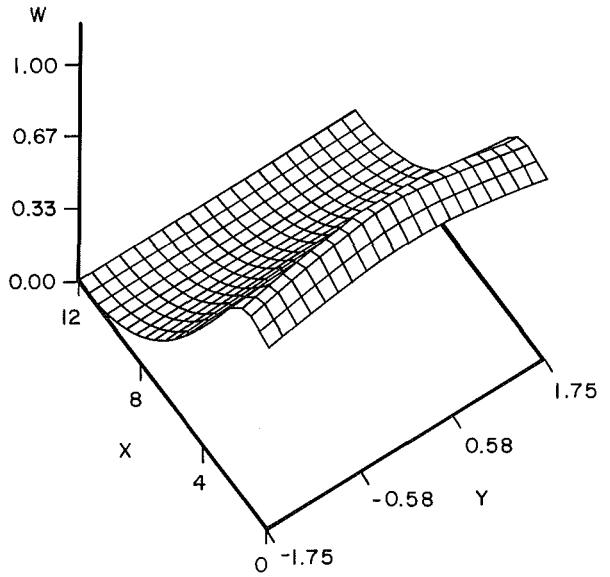


Fig. 8 The third "improved" buckling mode shape induced in a 3.5 cm × 12 cm ribbon by the profile given in equation (13) with φ = 0.00319. *Other conditions are as in Fig. 6.*

W_m^* can be used as "improved" shape functions in equation (12) to yield improved values of φ.

Many of the improved mode shapes $W_m^*(x, y)$ obtained from equation (8) are identical to those obtained from equation (10). However it is found that all modes are not identical. In particular if φ is small the term $(1 + 2fE/3\varphi)$ appreciably differs from unity and makes the new modes important. Typical of the "correction" to the deflected shapes is that shown in Fig. 8 in which the region near $x = 0$ is drastically bent with respect to that shown in Fig. 6. Clearly this reflects the viscoplastic nature of the material. If φ is large, most of the improved modes obtained from equation (8) are the same as those of equation (10).

What we are doing is a typical creep buckling analysis in which an initially deflected plate continues to deflect in time. We presume that the initial deflection is in a mode shape and

Table 5 Creep buckling results for a 4.6 cm × 12 cm ribbon (traction free boundary conditions)

mode	critical thickness (mm)	φ (sec^{-1})
1	0.289	-0.0145
2	0.235	0.00342
3	0.174	-0.00724
4	0.162	-0.00000218
5	0.158	0.00492
6	0.143	-0.000115
7	0.132	-0.00252
8	0.107	0.000679
9	0.092	0.00160

then the magnitude of the w grows with time. Due to f being assumed to be independent of w, we find the deflections grow as $g^o e^{\varphi t}$. Since most of the interest is in the fastest growing modes, and this implies large φ, the improved modes are primarily of academic interest.

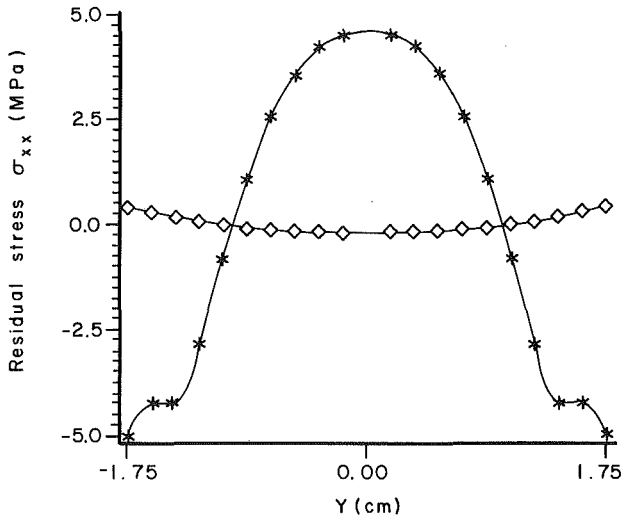


Fig. 9 Residual σ_{xx} stress distribution along the width direction (y) for a 3.5 cm × 12 cm ribbon with initial dislocation density = 13 cm^{-2} subjected to the profile of equation (13) (curve with stars) and equation (14) (curve with diamonds)

5 Discussions

The main effect of the residual stresses can be illustrated by considering a plate with different boundary conditions for the inplane stresses at $x = L$. The results for such a plate having the traction free boundary condition at $x = L$, but otherwise of the same geometry as in Table 4, are given in Table 5. The thickness needed to prevent elastic buckling is not much different in the two cases. However, the values of φ, and hence which modes will grow, are considerably different.

In order to further study the contribution of the residual stresses themselves to the buckling, we considered two hypothetical plates in which the inplane stresses are the residual stresses throughout the entire length of the plate (i.e., inplane stresses are uniform in the x direction). The residual stresses developed by the profile of equation (13) are shown in Fig. 9. For a plate with 18 cm in length and 3.5 cm in width, we calculated a critical thickness equal to 0.457 mm. This is greater than the 0.328 mm calculated for the same geometry but the stress field is similar to that shown in Figs. 3 and 4. If the residual stresses are multiplied by minus one, thus putting tension on the outside of the plate, the critical thickness decreases to 0.138 mm. This illustrates that a thermal profile which generates tensile residual σ_{xx} stresses on the outside of the plate has a great advantage in preventing buckling over one with edge compression.

Consider now a second profile which has been used in a thermal modeling effort for improving the web growing process by changing the furnace design (Seidensticker, 1986). This profile results in small tensile residual stresses σ_{xx} on the edges. This new profile is

$$T_{\text{new}}(x) = 60e^{-8.306x} - 85x + 1352, \qquad (14)$$

Equation (14) is shown as a dashed curve in Fig. 2. For this profile and the case of a 12 cm long and 3.5 cm wide ribbon, the residual stresses σ_{xx} variation along the ribbon width is shown in Fig. 9 (curve with diamonds). The profile of equation (14) results in a much smaller magnitude of the residual stresses than those (see the curve with stars in Fig. 9) obtained from equation (13). The critical thickness is found to be 0.045 mm when only these new residual stresses are applied throughout the entire ribbon. This new profile probably represents a nearly optimum one in so far as residual stresses are concerned.

The value of φ controls the speed of the lateral deflection of plates. When the value of φ is equal to infinity, elastic buckling occurs. When φ is positive, the amplitude of the buckling shape will grow with time. Larger values of φ will cause the lateral deflection of plates to grow faster. When φ is negative, the amplitude of the imperfection in the plate will decrease with time.

This ambivalence about whether a specific mode will grow or damp out (go back to flat) is probably due to the stresses being self-equilibrated. It certainly is extremely sensitive to the specific inplane stress distribution that is used to calculate φ. To the authors knowledge this is a new type behavior.

The results in Tables 2–4 show that the value of φ is small even though we used $h = 1.1h_1$. Based on equation (12), a greater ribbon thickness h will further reduce the value of φ, and hence the buckling speed. That is, the threshold for the lateral deflection of silicon ribbon can be moved to higher stress (i.e., wider ribbon) by growing ribbon thicker than h_1.

Once the value of φ is reduced so that the growth behavior can be neglected, the buckling of the final product of thin ribbon seems to be primarily due to the residual stresses existing in regions of the ribbon where there are small plastic strain rates. There are many ways to decrease the critical ribbon thickness by adjusting the thermal profile. A profile which reduces the magnitude of the residual stresses is one way. Selecting one that causes tensile residual stresses to occur on the edge region and compression near the center is a second way. Of course, there are other ways (for example, putting reinforcements on the outer edges) than adjusting the thermal profile that can also reduce the critical thickness.

6 Summary

The inplane stresses obtained from using the residual stress boundary conditions in the prebuckling state are used to calculate buckled shapes and their lateral deflection speeds. The lateral deflection speed can either be positive or negative. Hence the lowest mode may not be the one that is likely to be observed in an experiment.

Ribbon that is moderately thicker than the eigen-thickness can reduce the buckling speed. However, the use of viscoplasticity in the inplane stress calculation cannot be neglected because it is responsible for the residual stresses as well as some modes growing in time.

A thermal profile which generates smaller tensile residual stresses on the edge regions and smaller compression on the

center regions can be used to produce ribbon that is significantly thinner than one which has large compression on the outside edges.

Residual stresses are important in evaluating buckling of silicon ribbon when the ribbon is long.

Acknowledgments

The writers would like to thank numerous colleagues in the JPL-DOE Solar Energy Flat Plate Project for their contributions to our program. We especially want to thank our colleague, Professor R. J. de Angelis, for help in understanding the material model that we used. We also want to thank Dr. M. Leipold (JPL) for his initial help, as well as Professor T. Gross (UK) and Tim O'Donnell (JPL) for numerous technical conversations on silicon. Drs. R. G. Seidensticker (Westinghouse), J. Spitznagel (Westinghouse), J. Kalejs (Mobil Solar), and R. Hartzell (Texas Instruments) have been of tremendous help with reports on the real world observations. This research was financed by the Department of Energy through subcontracts from JPL Contract No. 956571 and SERI subcontract XL-7-06123-1.

References

Ciszek, T. F., 1985, "Silicon in Solar Cells," *Crystal Growth of Electronic Materials*, Kaldis, E., ed., Elsevier Science Publishers B.V., pp. 185–210.

Dillon, O. W., Jr., and De Angelis, R. J., 1984, "On Improved Analysis of the Thermal Buckling of Silicon Sheet," *Proc. of the Flat-Plate Solar Array Project Research Forum on High-Speed Growth and Characterization of Crystals for Solar Cells*, Jet Propulsion Laboratory Publication 84-23, Pasadena, CA, pp. 331–348.

Dillon, O. W., Jr., De Angelis, R. J., and Tsai, C. T., 1986, "Dislocation Dynamics During the Growth of Silicon Ribbon," *J. Appl. Phys.*, Vol. 60, No. 5, pp. 1784–1792.

Dillon, O. W., Jr., Tsai, C. T., and De Angelis, R. J., 1987, "Dislocation Dynamics of Web Type Silicon Ribbon," *Journal of Crystal Growth*, in press.

Duncan, C. S., Seidensticker, R. G., Mchugh, J. P., and Schruben, J., 1982, "Advanced Dendritic Web Growth Development," annual report, October 23, 1981, to October 22, 1982, Westinghouse R and D Center, Pittsburgh, PA.

Haasen, P., 1962, "Zur Plastischen Verformuny Von Germanium und In-Sb," *Z. Phys.*, Vol. 167, pp. 461–467.

Haasen, P., 1967, "Dislocation Dynamics in the Diamond Structure," *Dislocation Dynamics*, Rosenfield, A. R., Hahn, G. T., Bement, A. L., Jr., and Jaffee, I., eds., Battelle Institute Materials Science Colloquia, pp. 701–739.

Hartzell, R., 1984, personal communication, Texas Instruments Corporation, Dallas, Texas.

Seidensticker, R. G., and Hopkins, R. H., 1980, "Silicon Ribbon Growth by the Dendritic Web Process," *Journal of Crystal Growth*, Vol. 50, pp. 221–235.

Seidensticker, R. G., and Schruben, J. S., 1984, "Control of Thermal Stress in Dendritic Web Growth," *Proc. of the Flat-Plate Solar Array Project Research Forum on High-Speed Growth and Characterization of Crystals for Solar Cells*, Jet Propulsion Laboratory Publication 84-23, Pasadena, CA, pp. 385–397.

Seidensticker, R. G., 1984, personal communication, Westinghouse R and D Center, Pittsburgh, PA.

Suezava, M., Sumino, K., and Yonenaga, I., 1979, "Dislocation Dynamics in the Plastic Deformation of Silicon Crystals II. Theoretical Analysis of Experimental Results," *Phys. Stat. Sol.* (*a*), Vol. 51, pp. 217–226.

Tsai, C. T., 1985, "Thermal Visco-Plastic Stress and Buckling Analysis of Silicon Ribbon," Ph.D. Dissertation, Department of Engineering Mechanics, University of Kentucky, Lexington, KY.

Tsai, C. T., and Dillon, O. W., Jr., 1987, "Thermal Viscoplastic Buckling During the Growth of Silicon Ribbon," *Int. J. Solids Structures*, in press.

Tvergaard, V., 1979, "Creep Buckling of Rectangular Plates Under Axial Compression," *Int. J. Solids Structures*, Vol. 15, pp. 441–456.

Yonenaga, I., and Sumino, K., 1978, "Dislocation Dynamics in Plastic Deformation of Silicon Crystals, I. Experiments," *Phys. Stat. Sol.* (*a*), Vol. 50, pp. 685–693.

APPENDIX A

The inplane stresses σ_{ij} are obtained as the solution of the equilibrium equations expressed as

$$\frac{\partial^2 \sigma_{xx}}{\partial x^2} - \frac{\partial^2 \sigma_{yy}}{\partial y^2} = 0 \tag{A-1}$$

and the compatability equation in the form

$$\nabla^2(\sigma_{xx} + \sigma_{yy} + \alpha\, E\, T) = E\, F_P \tag{A-2}$$

where

$$F_P = \frac{\partial^2 \epsilon_{xx}^{PL}}{\partial y^2} + \frac{\sigma^2 \epsilon_{yy}^{PL}}{\partial x^2} - 2\,\frac{\partial^2 \epsilon_{xy}^{PL}}{\partial x \partial y} \tag{A-3}$$

and $\dot{N}$ given by equation (4). The plastic "strains" in equation (*A*-3) are

$$v\, \epsilon_{ij}^{PL} = \int_o^x \epsilon_{ij}^{\cdot PL} du$$

where the $\epsilon_{ij}^{\cdot PL}$ are given by equation (2) in terms of the inplane stresses.

During the growth of the silicon ribbon, the solid-melt interface ($x = 0$) and two outer edges ($y = \pm C$) are assumed free of tractions and moments in all directions. Since the value of $\partial^2 T/\partial x^2$ in growing ribbon usually drops to zero after a short distance (say L^*) from the solid-melt interface, the ribbon grows under the steady-state when the length of growing ribbon is larger than L^*. This implies that the stresses in the region beyond $x = L^*$ are constant so that the stress gradients in the x direction are zero in this region. The use of zero stress gradients along $x = L^*$ permits the residual in-plane stresses to be calculated in the prebuckling analysis. When the in-plane stresses σ_{xx}, σ_{xy}, and σ_{yy} are being calculated, the boundary conditions mentioned above are called the residual stress boundary condition. These stresses along $x = L^*$ are the residual stresses. The σ_{xy} and σ_{yy} stresses are found to be very small when compared to σ_{xx} (Dillon et al., 1986).

Hence the boundary conditions are

$$\sigma_{xx}(0,y) = \sigma_{xy}(0,y) = 0 = \sigma_{yy}(x, \pm C) = \sigma_{xy}(x, \pm C)$$

and the residual stress ones

$$\partial \sigma_{ij}/\partial x_j (L^*, y) = 0. \tag{A-4}$$

This system is expressed in their finite difference equivalent and the results solved iteratively for σ_{xx} and σ_{yy} on the digital computer. In the right-hand side of equation (*A*-2) the function F_P is evaluated using the stresses obtained in the previous iteration. The shearing stresses are obtained from one of the equilibrium equations or by "fitting" a stress function to the data on σ_{xx} and σ_{yy} and then differentiating the result to obtain σ_{xy}.

APPENDIX B

The transverse displacement $w(x, y)$ at a generic point in the plate is written as

$$w = w^o + w^e + w^{VP} \tag{B-1}$$

where w^o is the deflection at zero stresses, w^e is the increment in displacement when the stresses σ_{ij} are applied considering the material to be elastic, and w^{VP} is that part of w due to the stresses σ_{ij} when considering silicon to be viscoplastic (i.e., without elasticity). The strains are written as

$$\epsilon_{ij} = \epsilon_{ij}^{oe} + \epsilon_{ij}^{oVP} - z\,\frac{\partial^2 (w^e + w^{VP})}{\partial x_i \partial x_j} \tag{B-2}$$

using the usual assumptions of small deflection plate theory; therefore, the strains are written as

$$\epsilon_{ij} = \epsilon_{ij}^{e} + \epsilon_{ij}^{VP}$$

In equation (*B*-2) the ϵ_{ij}^{oe} and ϵ_{ij}^{VP} are independent of z. The moment intensities are given by the basic definition

$$M_{ij} = \int_{-h/2}^{h/2} z\, \sigma_{ij} dz \tag{B-3}$$

Hence we have

$$M_{xx} = -D_e\left(\frac{\partial^2 w^e}{\partial x^2} + \nu\,\frac{\partial^2 w^e}{\partial y^2}\right) \tag{B-4}$$

and also

$$M_{xx} = -D_{VP}\left(\frac{\partial^2 \dot{w}^{VP}}{\partial x^2} + \frac{\partial^2 \dot{w}^{VP}}{2\partial y^2}\right)^2 \qquad (B\text{-}5)$$

where $D_{VP} = h^3/12f$.

If we assume $\nu = \frac{1}{2}$, equations (B-4) and (B-5) imply

$$\frac{\partial^2 \dot{w}^{VP}}{\partial x^2} = \frac{2fE}{3}\frac{\partial^2 w^e}{\partial x^2} \qquad (B\text{-}6)$$

and corresponding expressions for $\partial^2 \dot{w}^{VP}/\partial y^2$ and $\partial^2 \dot{w}^{VP}/\partial x \partial y$. This procedure is analogous to Trevguard (1979).

The general equilibrium equation for a plate subjected only to inplane stresses σ_{ij} is

$$\partial^2 M_{ij}/\partial x_i \partial x_j = -h\sigma_{ij}\partial^2(w^o + w^e + w^{VP})/\partial x_i \partial x_j \qquad (B\text{-}7)$$

Using the expressions of the type given by equation (B-4) in the left-hand side of equation (B-7), one obtains

$$D_e \nabla\ w^e = h\sigma_{ij}\partial^2(w^o + w^e + w^{VP})/\partial x_i \partial x_j \qquad (B\text{-}8)$$

Equation (B-8) is then differentiated with respect to time and w^o thereby eliminated, except from the associated initial conditions.

Using the relations of the type (B-6), one eliminates $\dot{w}^{VP}$ in favor of w^e and thereby obtains equations (5) of the text. Relations similar to equation (B-4) can also be used to express the boundary conditions in terms of w^e, precisely as in elastic plate theory.

Thus the problem to be solved is equation (5) with

$$M_{xx}(o,y) = M_{yy}(x, \pm C) = M_{xy}(o,y) = o$$

$$M_{xy}(x, \pm C) = o = Q_x(o,y) = Q_y(x, \pm C)$$

We do not know the initial deflected shape so that we watch any one of the "pseudo" elastic mode shapes grow in time. That is

$$w^e(x,y,o) = g^o W_m(x,y).$$

Nonlinear Forced Response of Infinitely Long Circular Cylindrical Shells

A. H. Nayfeh

R. A. Raouf

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

A combination of the Galerkin procedure and the method of multiple scales is used to analyze the nonlinear forced response of infinitely long circular cylindrical shells (or circular rings) in the presence of internal (autoparametric) resonances. If ω_f and a_f denote the frequency and amplitude of a flexural mode and ω_b and a_b denote the frequency and amplitude of the breathing mode, the steady-state response is found to exhibit a saturation phenomenon when $\omega_b \approx 2\omega_f$ if the shell is excited by a harmonic load having a frequency Ω near ω_b. As the amplitude f of the excitation increases from zero, a_b increases linearly with f until a threshold value f_c of f is reached. Beyond f_c, a_b remains constant and the extra energy spills over into the flexural resonant mode whose amplitude grows nonlinearly. Results of numerical investigations, guided by the perturbation analysis, show that the long-time response exhibits a Hopf bifurcation, yielding amplitude and phase-modulated motions. The amplitudes and phases experience a cascade of period-doubling bifurcations ending up with chaos. The bifurcation values are finely tuned.

1 Introduction

Recently, the problem of the nonlinear vibration of shells has received considerable attention. The sources of the nonlinearities in the governing equations may be geometric, inertial, material, or any combination. These nonlinearities appear in the governing partial-differential equations and may appear in the boundary conditions. However, most of the existing studies of other than composite shells deal with geometric nonlinearities.

The methods of solution of the nonlinear partial differential equations governing shell motion can be broadly classified into three approaches: purely numerical methods, perturbation methods, and a combination of the Galerkin procedure with either perturbation or numerical methods.

The last is the most commonly used approach. It consists of expanding the dependent variables in terms of a linear combination of shape functions with time-varying coefficients. These temporal coefficients are treated as generalized coordinates. The Galerkin procedure is used to derive a set of nonlinear ordinary-differential equations. These equations are solved using numerical or perturbation techniques. Examples of this approach include the work of Atluri (1972), who employed the method of multiple scales, and the works of Reissner (1955), Chen and Babcock (1975), and Hui (1983), who employed the Lindstedt-Poincaré technique.

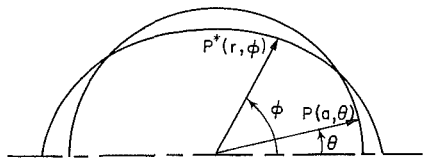


Fig. 1 Polar coordinates of a point of the shell which was initially at *P* and at *P at time *t****

Since the problem is governed by partial-differential equations, the response, in general, consists of many modes. In fact, using the Galerkin procedure one obtains an infinite set of nonlinear coupled equations describing the time variation of the amplitudes of the infinitely many modes. All existing studies truncate the infinite set of equations to a finite number and many of them keep only one mode.

The first studies of modal interactions in the response of shells were initiated by McIvor (1962, 1966), Goodier and McIvor (1964), McIvor and Sonstegard (1966), and McIvor and Lovell (1968). They analyzed the response of infinitely long cylindrical and spherical shells to radial and nearly radial impulses, taking into account the coupling of breathing and flexural modes when their frequencies are in the ratio of two-to-one. By considering the linearized equation of the breathing mode, they obtained a Mathieu-type equation for the flexural vibrations and used it to study the stability of the shell. Other examples of modal interaction studies include the works of Bieniek et al. (1966), who used Donell's equations to obtain a Mathieu-type equation to study the stability of the axisymmetric modes, Atluri (1972), who used the method of multiple scales to analyze free oscillations of shells in the absence of in-

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 17, 1986; final revision April 20, 1987.

Paper No. 87-WA/APM-15.

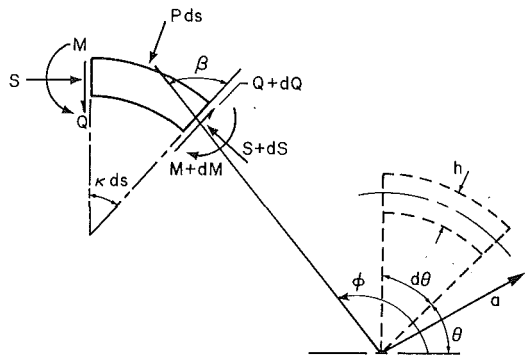


Fig. 2 An edge of an element of the deformed and undeformed shell

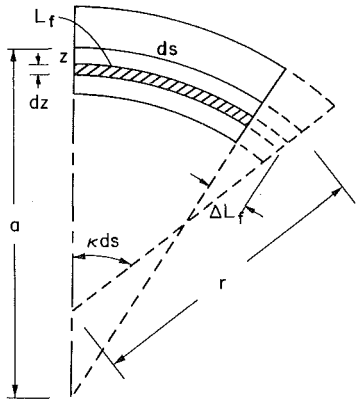


Fig. 3 Deformed shell element

ternal resonances, Mente (1973), who numerically solved a set of n nonlinear equations arising from the Galerkin procedure, and Chen (1972) and Chen and Babcock (1975), who used the Lindstedt-Poincaré technique to study the interaction of a flexural mode with its companion mode (i.e., one to one resonance) for a simply-supported cylindrical shell.

In the present paper, we analyze the nonlinear response of an infinitely long cylindrical shell (or a ring) to a harmonic excitation when the frequency of the breathing mode is approximately twice the frequency of a flexural mode. We use the method of multiple scales to fully account for the nonlinear interaction, including the influence of the flexural mode on the breathing mode. We demonstrate the saturation phenomenon. This phenomenon was found by Nayfeh et al. (1973) in the response of internally-resonant ships. We also show the existence of a Hopf bifurcation and numerically demonstrate the occurrence of chaotic motion. Systems with two-to-one internal resonances were studied by Mettler and Weidenhammer (1962), Miles (1984, 1985), Sethna (1965), Nayfeh et al. (1973), Mook et al. (1974), Yamamoto and Yasuda (1977), Hatwal et al. (1982), Haddow et al. (1984), among others.

2 Equations of Motion

Following McIvor (1962) and Goodier and McIvor (1964), we consider the case of plane strain in which the strain parallel to the generators of the shell is everywhere zero. Thus, the deformation of the shell is identical in every plane perpendicular to the shell axis, and the shell can be considered as being in plane motion (or simply as an elastic ring). In such a plane, we consider a point P on the undeformed shell midsurface with the polar coordinate (a, θ), which after a time t^* moves to P^* with the polar coordinates r and ϕ, as shown in Fig. 1. Figure 2 shows an edge of an element of unit width of the shell in both the deformed and undeformed configurations. Let the coordinates of an element of the undeformed midsurface be (a, θ) and $(a, \theta + d\theta)$ and those of the deformed element be (r, ϕ) and $(r + \delta r, \phi + \delta\phi)$. Then, the extensional strain is given by

$$\epsilon_o = \frac{ds^* - ds}{ds} = a^{-1}(r'^2 + r^2\phi'^2)^{1/2} - 1 \tag{1}$$

where the prime denotes the partial derivative with respect to θ. Thus

$$ds^* = ad\theta(1 + \epsilon_o) \tag{2}$$

It follows from differential geometry that the curvature is given by

$$\kappa = [\phi'(r^2\phi'^2 - rr'' + 2r'^2) + \phi'' r' r][r'^2 + r^2\phi'^2]^{-3/2} \tag{3}$$

The usual assumptions of thin shell theory are used here. Straight lines normal to the midsurface before deformation stay straight and normal to the midsurface after deformation, implying that the shear deformations are negligible. The thickness h of the shell is unchanged, and the normal stress is negligible. The ratio (h^2/a^2), where a is the initial radius of the shell, is small. Using the above assumptions and assuming plane strain, we find from Fig. 3 that

$$\epsilon_\phi = \epsilon_o - z(1 + \epsilon_o)\left(\kappa - \frac{1}{a}\right)\left(1 + \frac{z}{a} + \frac{z^2}{a^2} + \ldots\right) \tag{4}$$

We introduce the dimensionless displacement w and time t defined by

$$w = \frac{a - r}{a}, \quad t = \frac{ct^*}{a} \tag{5}$$

where t^* is the dimensional time, $c^2 = E/\rho(1 - \nu^2)$, E is Young's modulus, ν is Poisson's ratio, and ρ is the density of the shell per unit width. Moreover, we let

$$\psi = \phi - \theta \tag{6}$$

Using Hamilton's principle and calculus of variation, we obtain the equations of motion

$$\ddot{w} + \alpha^2(w^{iv} + 2w'' + w) - \psi' + w = w''(\psi' - w) - \dot{\psi}^2 + \psi'^2 - 2w\psi' + w'\psi'' - \frac{1}{2}w'^2 + \frac{a(1 - \nu^2)}{Eh}P(1 + \psi' - w) \tag{7}$$

and

$$\ddot{\psi} - \psi'' + w' = w'w'' - 2w'\psi' + 2\dot{w}\dot{\psi} + \frac{a(1 - \nu^2)}{Eh}w'P \tag{8}$$

in agreement with those obtained by McIvor (1962) and Goodier and McIvor (1964). These equations have the same mathematical form as those of a ring. The ring equations can be obtained from equations (7) and (8) by replacing $a(1 - \nu^2)/Eh$ by $ah^2/(12EAk^2)$, where A and k are, respectively, the area and radius of gyration of the cross section of the ring. Simmonds (1979) showed that the cubic nonlinearities in an elastic ring have a significant effect on the behavior of low modes of vibration. The present study will consider higher flexural modes and thus cubic nonlinearities can be ignored.

3 Inextensional Oscillations

Goodier and McIvor restricted their analysis to impulses with durations much less than the period of the uniform radial (i.e., breathing) mode of vibration. Such a restriction made it possible to convert the problem into that of free vibration. Under the assumption of inextensionality and by simplifying the expression for the kinetic energy through neglecting a $(2w)$ term, McIvor (1962, 1966) and Goodier and McIvor (1964) produced a numerical solution of the approximate equations of motion. In the present analysis, such assumptions concerning the energy are not used.

Since the shell is closed, w and ψ must be periodic in θ with

period 2π. Consequently, they can be expanded in Fourier series as

$$w=\eta_o(\tau)+\sum_{n=1}^{\infty}[\eta_n(\tau)\cos n\theta+\zeta_n(\tau)\sin n\theta] \tag{9a}$$

and

$$\psi=\sum_{n=1}^{\infty}[c_n(\tau)\cos n\theta+d_n(\tau)\sin n\theta] \tag{9b}$$

The basic response of the shell is the radial (breathing), purely extensional mode $\eta_o(\tau)$, which is assumed to be perturbed by the inextensional modes. Thus, we follow Goodier and McIvor (1964) and assume that the deviation from the breathing circular mode satisfies the inextensionality condition

$$\psi'-w=0$$

Hence, equations (9) become

$$w=\eta_o(\tau)+\sum_{n=2}^{\infty}[\eta_n(\tau)\cos n\theta+\zeta_n(\tau)\sin n\theta] \tag{10a}$$

and

$$\psi=\sum_{n=2}^{\infty}\left[-\frac{\zeta_n}{n}\cos n\theta+\frac{\eta_n}{n}\sin n\theta\right] \tag{10b}$$

The first harmonic is omitted in equations (10) because it corresponds to a rigid body translation in the inextensional model. The damping is assumed to be given by the dissipation function

$$D_f=\frac{1}{2}E_1ha\pi\left[2\mu_o\dot\eta_o^2+\sum_{n=2}^{\infty}\mu_n\frac{n^2+1}{n^2}(\dot\eta_n^2+\dot\zeta_n^2)\right] \tag{11}$$

so that the resulting damping is modal. The quantities η_o, η_n, and ζ_n can be considered as generalized coordinates resulting in the following equations of motion:

$$\ddot\eta_o+\omega_o^2\eta_o+2\mu_o\dot\eta_o+\sum_{n=2}^{\infty}\left[\frac{1}{2n^2}(\dot\eta_n^2+\dot\zeta_n^2)-\frac{1}{4}(n^2-2)(\eta_n^2+\zeta_n^2)\right]=P_o(t) \tag{12}$$

$$\ddot\eta_n+\omega_n^2\eta_n+2\mu_n\dot\eta_n-\frac{2}{n^2+1}\eta_o\ddot\eta_n-\frac{2}{n^2+1}\dot\eta_o\dot\eta_n-\frac{n^2(n^2-2)}{n^2+1}\eta_o\eta_n+\text{higher-order terms}=P_n(t) \tag{13}$$

$$\ddot\zeta_n+\omega_n^2\zeta_n+2\mu_n\dot\zeta_n-\frac{n^2(n^2-2)}{n^2+1}\eta_o\zeta_n-\frac{2}{n^2+1}\eta_o\ddot\zeta_n-\frac{2}{n^2+1}\dot\eta_o\dot\zeta_n+\text{higher-order terms}=Q_n(t) \tag{14}$$

where the P_n and Q_n are generalized forces, the μ_m are damping coefficients, and

$$\omega_o^2=1+\alpha^2,\quad \omega_n^2=\frac{n^2(n^2-1)^2}{(n^2+1)}\alpha^2 \tag{15}$$

The full form and details of the derivation of equations (12)–(14) are given by Nayfeh and Raouf (1986). The higher order terms are not listed because they do not appear in the second-order approximation presented in this paper.

4 Perturbation Analysis

In this section, the method of multiple scales (Nayfeh, 1973, 1981) is used to derive an asymptotically valid closed-form solution for equations (12)–(14) in the case of a two-to-one internal (autoparametric) resonance between the breathing mode η_o and a flexural mode η_s and ζ_s. Thus, we consider the case $\omega_o\approx 2\omega_s$. Moreover, we consider the case of a harmonic excitation of the breathing mode near primary resonance; that is, we let $P_n=S_n=0$ and

$$P_o=2F\cos\Omega t \tag{16}$$

where $\Omega\approx\omega_o$.

Following the method of multiple scales, we seek a uniformly valid expansion of the variables in the form

$$\eta_o(t;\epsilon)=\epsilon\eta_{o^1}(T_o,T_1)+\epsilon^2\eta_{o^2}(T_o,T_1)+\ldots \tag{17}$$

$$\eta_n(t;\epsilon)=\epsilon\eta_{n^1}(T_o,T_1)+\epsilon^2\eta_{n^2}(T_o,T_1)+\ldots \tag{18}$$

$$\zeta_n(t;\epsilon)=\epsilon\zeta_{n^1}(T_o,T_1)+\epsilon^2\zeta_{n^2}(T_o,T_1)+\ldots \tag{19}$$

for $n=2, 3, 4, \ldots$, where ϵ is a small dimensionless parameter that is used as a bookkeeping device, $T_o=t$ is a fast scale, and $T_1=\epsilon t$ is a slow scale. Moreover, we order the amplitude of the excitation and the damping coefficients so that

$$F\to\epsilon\omega_o f\quad\text{and}\quad\mu_n\to\epsilon\mu_n \tag{20}$$

Substituting equations (17)–(20) into equations (12)–(14) and equating coefficients of like powers of ϵ on both sides, we obtain equations describing the η_{nm} and ζ_{mn} (Raouf, 1985, Nayfeh and Raouf, 1986). The solution of the first-order problem can be expressed as

$$\eta_{o^1}=A_o(T_1)e^{i\omega_oT_o}+c.c. \tag{21}$$

$$\eta_{n^1}=A_n(T_1)e^{i\omega_nT_o}+c.c. \tag{22}$$

$$\zeta_{n^1}=B_n(T_1)e^{i\omega_nT_o}+c.c. \tag{23}$$

where *c.c.* stands for the complex conjugate of the preceding terms. The functions A_o, A_n, and B_n are arbitrary at this order; they are determined by imposing the solvability conditions at the next level of approximation.

Next, equations (21)–(23) are substituted into the second-order problem and detuning parameters σ_1 and σ_2 are defined as

$$\Omega=\omega_o+\epsilon\sigma_1\quad\text{and}\quad 2\omega_s=\omega_o+\epsilon\sigma_2 \tag{24}$$

Then, eliminating the terms that produce secular terms from the inhomogeneous equations governing the η_{m^2} and ζ_{m^2}, we obtain

$$2A_o'+2\mu_oA_o+4i\Lambda_1(A_s^2+B_s^2)e^{i\sigma_2T_1}+\text{if}\,e^{i\sigma_1T_1}=0 \tag{25}$$

$$2A_s'+2\mu_sA_s+4i\Lambda_2A_o\bar A_se^{-i\sigma_2T_1}=0 \tag{26}$$

$$2B_s'+2\mu_sB_s+4i\Lambda_2A_o\bar B_se^{-i\sigma_2T_1}=0 \tag{27}$$

where

$$4\omega_o\Lambda_1=\frac{1}{4}(s^2-2)+\frac{\omega_s^2}{2s^2} \tag{28}$$

$$4\omega_s\Lambda_2=\frac{1}{s^2+1}[s^2(s^2-2)-2(\omega_s^2+\omega_s\omega_o)] \tag{29}$$

Letting

$$A_o=\frac{1}{2}a_o(T_1)e^{i\beta_o(T_1)},\quad B_s=\frac{1}{2}b_s(T_1)e^{i\nu_s(T_1)},$$

$$A_s=\frac{1}{2}a_s(T_1)e^{i\beta_s(T_1)} \tag{30}$$

in equations (25)–(27) and separating real and imaginary parts, we obtain

$$a_o'+\mu_oa_o+\Lambda_1a_s^2\sin\gamma_2+\Lambda_1b_s^2\sin\gamma_3-f\sin\gamma_1=0 \tag{31}$$

$$a_s'+\mu_sa_s-\Lambda_2a_oa_s\sin\gamma_2=0 \tag{32}$$

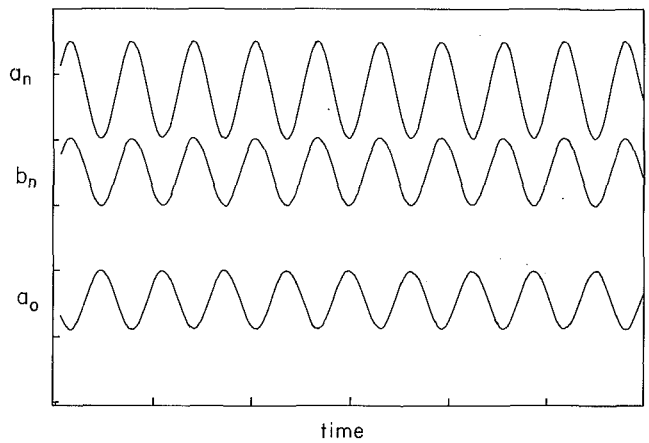


Fig. 4 A typical time history of the free-oscillation response

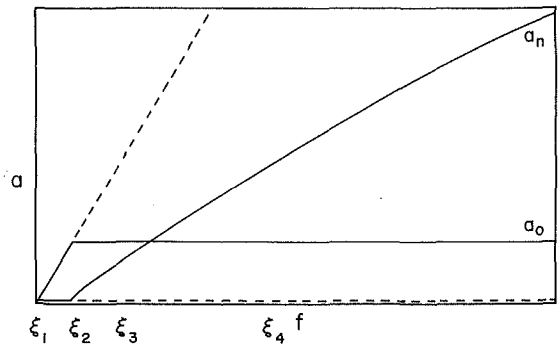


Fig. 5 Modal response amplitudes as functions of the amplitude of the excitation when $\Gamma < 0$

$$b_s' + \mu_s b_s - \Lambda_2 a_o b_s \sin\gamma_3 = 0 \tag{33}$$

$$a_o\beta_o' + \Lambda_1 a_s^2 \cos\gamma_2 + \Lambda_1 b_s^2 \cos\gamma_3 + f\cos\gamma_1 = 0 \tag{34}$$

$$a_s\beta_s' + \Lambda_2 a_o a_s \cos\gamma_2 = 0 \tag{35}$$

$$b_s\nu_s' + \Lambda_2 a_o b_s \cos\gamma_3 = 0 \tag{36}$$

where

$$\gamma_1 = \sigma_1 T_1 - \beta_o, \quad \gamma_2 = \beta_o - 2\beta_s - \sigma_2 T_1$$

$$\text{and } \gamma_3 = \beta_o - 2\nu_s - \sigma_2 T_1 \tag{37}$$

In the following section, we discuss the undamped case of free oscillations, and in Section 6, we discuss the case of forced damped oscillations.

5 Undamped Free Oscillations

In this case, $f = 0$ and $\mu_o = \mu_s = 0$. Under these conditions, eliminating γ_2 and γ_3 from equations (31)–(36) yields

$$\Lambda_2 a_o a_o' + \Lambda_1 a_s a_s' + \Lambda_1 b_s b_s' = 0 \tag{38}$$

Integrating equation (38) yields

$$(\Lambda_2/\Lambda_1)a_o^2 + a_s^2 + b_s^2 = E \tag{39}$$

where E is a constant of integration. Since all the terms on the left-hand side of equation (39) are positive, E must also be positive, which means that the motion described is a bounded one. When $b_s = 0$, the solutions of equations (31)–(34) and (37) can be expressed in terms of Jacobi elliptic functions (Nayfeh and Mook, 1979).

Next, we present numerical results for the case $\alpha^2 = 2.0918 \times 1.0^{-4}$, $(h/a \approx 1/20)$, which leads to $\omega_o = 1.0001$ and $\omega_6 = 0.4993$ so that $\omega_o \approx 2\omega_6$. In this case, $\Lambda_1 = 2.1259$ and $\Lambda_2 = 16.5203$. Figure 4 shows the time-history of the amplitudes of the free-oscillation response obtained by integrating equations (31)–(37). It clearly shows the continual exchange of energy between the breathing and flexural modes. The maximum amplitude of the flexural mode is about a factor of 3 larger than the maximum amplitude of the breathing mode.

6 Forced Damped Oscillations

Fixed points and hence steady-state periodic solutions of equations (31)–(37) correspond to $a_o' = a_s' = b_s' = 0$ and $\gamma_s' = 0$. It follows from equations (37) that $\beta_o' = \sigma_1$ and $\beta_s' = \nu_s' = \frac{1}{2}(\sigma_1 - \sigma_2)$. Hence, steady-state periodic solutions have four possibilities. First,

$$a_s = b_s = 0 \quad \text{and} \quad a_o = f(\mu_o^2 + \sigma_1^2)^{-1/2} \tag{40}$$

which is essentially the linear solution. Second, $a_s = 0$ and $b_s \neq 0$. Third, $a_s \neq 0$ and $b_s = 0$. Fourth, $a_s \neq 0$ and $b_s \neq 0$. The last solution includes the second and third solutions as special cases. Then, it follows from equations (32), (33), (35), and (36) that

$$a_o = a_o^* = \Lambda_2^{-1}\left[\mu_s^2 + \frac{1}{4}(\sigma_1 - \sigma_2)^2\right]^{1/2} \tag{41}$$

$$\tan\gamma_2 = \tan\gamma_3 = [2\mu_s/(\sigma_2 - \sigma_1)] \tag{42}$$

Then, it follows from equations (31) and (34) that

$$a_s^2 + b_s^2 = (\Lambda_1\Lambda_2)^{-1}\left\{\frac{1}{2}\sigma_1(\sigma_1 - \sigma_2) - \mu_o\mu_s \pm \left[f^2\Lambda_2^2 - \left(\sigma_1\mu_o + \frac{1}{2}\mu_s(\sigma_1 - \sigma_2)^2\right)^2\right]^{1/2}\right\} \tag{43}$$

Equation (41) shows that the amplitude a_o of the directly excited breathing mode is independent of the amplitude f of the excitation. It depends only on the damping of the flexural modes and the detuning parameters σ_1 and σ_2. On the other hand, the amplitudes a_s and b_s of the flexural mode are strongly dependent on the excitation amplitude f.

To determine the stability of the steady-state periodic solutions, we let

$$A_o = \frac{1}{2}(p_1 - iq_1)e^{i\nu_1 T_1}$$

$$A_s = \frac{1}{2}(p_2 - iq_2)e^{i\nu_2 T_1} \tag{44}$$

$$B_s = \frac{1}{2}(p_3 - iq_3)e^{i\nu_2 T_1}$$

where

$$\nu_1 = \sigma_1 \quad \text{and} \quad \nu_2 = \frac{1}{2}(\sigma_1 - \sigma_2)$$

in equations (25)–(27), separate real and imaginary parts, and obtain

$$p_1' + \nu_1 q_1 + \mu_o p_1 + 2\Lambda_1(p_2q_2 + p_3q_3) = 0 \tag{45}$$

$$q_1' - \nu_1 p_1 + \mu_o q_1 - \Lambda_1(p_2^2 + p_3^2 - q_2^2 - q_3^2) = f \tag{46}$$

$$p_2' + \nu_2 q_2 + \mu_s p_2 + \Lambda_2(q_1p_2 - q_2p_1) = 0 \tag{47}$$

$$q_2' - \nu_2 p_2 + \mu_s q_2 - \Lambda_2(p_1p_2 + q_1q_2) = 0 \tag{48}$$

$$p_3' + \nu_2 q_3 + \mu_s p_3 + \Lambda_2(q_1p_3 - q_3p_1) = 0 \tag{49}$$

$$q_3' - \nu_2 p_3 + \mu_s q_3 - \Lambda_2(p_1p_3 + q_1q_3) = 0 \tag{50}$$

Equations (45)–(50) are a generalization of those studied by Miles (1985) and Nayfeh (1987). The local stability of a fixed point with respect to a small perturbation proportional to exp (λT_1) is determined by the zeros of the characteristic equation

$$\begin{vmatrix} \lambda+\mu_o & \nu_1 & 2\Lambda_1 q_2 & 2\Lambda_1 p_2 & 2\Lambda_1 q_3 & 2\Lambda_1 p_3 \\ -\nu_1 & \lambda+\mu_o & -2\Lambda_1 p_2 & 2\Lambda_1 q_2 & -2\Lambda_1 p_3 & 2\Lambda_1 q_3 \\ -\Lambda_2 q_2 & \Lambda_2 p_2 & \lambda+\mu_s+\Lambda_2 q_1 & \nu_2-\Lambda_2 p_1 & 0 & 0 \\ -\Lambda_2 p_2 & -\Lambda_2 q_2 & -\nu_2-\Lambda_2 p_1 & \lambda+\mu_s-\Lambda_2 q_1 & 0 & 0 \\ -\Lambda_2 q_3 & \Lambda_2 p_3 & 0 & 0 & \lambda+\mu_s+\Lambda_2 q_1 & \nu_2-\Lambda_2 p_1 \\ -\Lambda_2 p_3 & -\Lambda_2 q_3 & 0 & 0 & -\nu_2-\Lambda_2 p_1 & \lambda+\mu_s-\Lambda_2 q_1 \end{vmatrix} = 0 \tag{51}$$

To investigate the stability of the linear solution given by equations (40), we put $p_2 = p_3 = q_2 = q_3 = 0$ in equation (51) and, after some algebraic manipulations, obtain

$$[(\lambda+\mu_o)^2+\nu_1^2][(\lambda+\mu_s)^2+\nu_2^2-\Lambda_2^2 a_o^2]^2 = 0 \tag{52}$$

Hence,

$$\lambda = -\mu_o \pm i\nu_1, \quad -\mu_s \pm (\Lambda_2^2 a_o^2-\nu_2^2)^{1/2}, \quad -\mu_s \pm (\Lambda_2^2 a_o^2-\nu_2^2)^{1/2} \tag{53}$$

Consequently, the linear solution is stable if and only if

$$\Lambda_2^2 a_o^2 \le \nu_2^2+\mu_s^2 \tag{54}$$

which, in conjunction with equation (41), implies that the linear solution is stable if $a_o \le a_o^*$ and unstable if $a_o > a_o^*$ or $f > \xi_2 = a_o^*(\mu_o^2+\sigma_1^2)^{1/2}$.

To study the stability of the nonlinear solution given by equations (41)–(43) when $b_s = 0$, we let $p_3 = q_3 = 0$ in equation (51), use equation (41), and obtain

$$\begin{aligned}[(\lambda+\mu_s)^2-\mu_s^2]\{\lambda^4+2(\mu_o+\mu_s)\lambda^3+[\mu_o^2+4\mu_o\mu_s+\nu_1^2 \\ +4\Lambda_1\Lambda_2 a_s^2]\lambda^2+[2\mu_s\mu_o^2+2\mu_s\nu_1^2+4\Lambda_1\Lambda_2(\mu_o+\mu_s)a_s^2]\lambda \\ +4\Lambda_1\Lambda_2 a_s^2[\Lambda_1\Lambda_2 a_s^2+\mu_o\mu_s-\nu_1\nu_2]\} = 0\end{aligned} \tag{55}$$

Hence, either $\lambda = 0$ or $-2\mu_s$ or

$$\begin{aligned}\lambda^4+2(\mu_o+\mu_s)\lambda^3+[\mu_o^2+4\mu_o\mu_s+\nu_1^2+4\Lambda_1\Lambda_2 a_s^2]\lambda^2 \\ +[2\mu_s\mu_o^2+2\mu_s\nu_1^2+4\Lambda_1\Lambda_2(\mu_o+\mu_s)a_s^2]\lambda \\ +4\Lambda_1\Lambda_2 a_s^2[\Lambda_1\Lambda_2 a_s^2+\mu_o\mu_s-\nu_1\nu_2] = 0\end{aligned} \tag{56}$$

The necessary and sufficient conditions that none of the roots of equation (56) have positive real parts are

$$\Lambda_1\Lambda_2 a_s^2+\mu_o\mu_s-\nu_1\nu_2 > 0 \tag{57}$$

$$\begin{aligned}4\mu_o\mu_s(\mu_o^2+\nu_1^2)(4\mu_s^2+4\mu_o\mu_s+\mu_o^2+\nu_1^2)+8(\mu_o \\ +\mu_s)^2\Lambda_1\Lambda_2 a_s^2(\mu_o^2+2\mu_o\mu_s+2\nu_1\nu_2+\nu_1^2) > 0\end{aligned} \tag{58}$$

Condition (57), in conjunction with equation (43), implies that the solution corresponding to the positive sign is stable whereas the solution corresponding to the negative sign is unstable. The violation of condition (58) would imply the existence of a pair of complex-conjugate roots of equation (56) with a positive real part. When $\nu_1\nu_2 > 0$, condition (58) is satisfied for all values of μ_o, μ_s, and f. On the other hand, when $\nu_1\nu_2 < 0$, condition (58) may be violated, depending on the values of μ_o, μ_n, and f.

Next, we present numerical results for the same case considered earlier $\omega_o \approx 2\omega_6$ with $\mu_o = 0.01$ and $\mu_6 = 0.01$.

In Fig. 5, we show a representative variation of the amplitudes of the breathing and 6th flexural modes for the case $\Gamma < 0$, where

$$\Gamma = \frac{1}{2}\sigma_1(\sigma_1-\sigma_2)-\mu_o\mu_s \tag{59}$$

If the shell is excited by a radial load of amplitude f and frequency $\Omega \approx \omega_o$, the linear solution shows that the steady-state amplitude a_6 of the flexural mode is zero, whereas the steady-state amplitude a_o of the breathing mode increases linearly with f. However, including the nonlinear terms shows that above a threshold value ξ_2 of f, where

$$\xi_2 = a_o^*(\mu_o^2+\sigma_1^2)^{1/2} = \Lambda_2^{-1}\left\{(\mu_o^2+\sigma_1^2)\left[\mu_6^2+\frac{1}{4}(\sigma_1-\sigma_2)^2\right]\right\}^{1/2} \tag{60}$$

the linear solution is unstable, a_o remains constant (saturates), and the additional energy spills over into the flexural mode. If the excitation frequency is such that $\sigma_1 = \sigma_2$, then the threshold value ξ_2 of f becomes

$$\xi_2 = \mu_6\Lambda_2^{-1}(\mu_o^2+\sigma_1^2)^{1/2} \tag{61}$$

which can be very small, depending on the damping coefficients, μ_o and μ_6. Consequently, the linear solution is unstable and the shell responds nonlinearly even for small excitations.

Next, we consider the case where $\mu_o = \mu_s = 0.02$, $\sigma_1 = -0.1$, and $\sigma_2 = -0.18$, then $\nu_1 = -0.1$, $\nu_2 = 0.04$, and $\Gamma < 0$. Hence, equation (43) has no real roots when $f < \xi_2$ and it has one real root for all $f \ge \xi_2$, where $\xi_2 \simeq 2.761 \times 10^{-4}$. When $\xi_2 \le f \le \xi_3$, where $\xi_3 \simeq 1.2761 \times 10^{-3}$, both conditions (57) and (58) are satisfied, and hence the finite-amplitude solution is stable. When $f \ge \xi_3$, condition (58) is violated, indicating the existence of a Hopf bifurcation and hence the

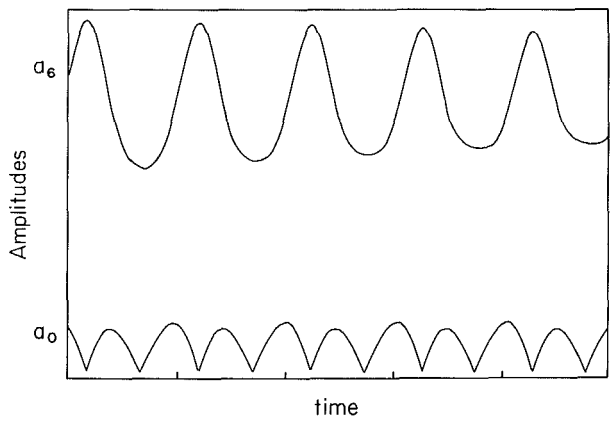


Fig. 6 Hopf bifurcation conditions. Variation of the steady-state amplitudes of the breathing and flexural modes as a function of time when $f = 0.1$.

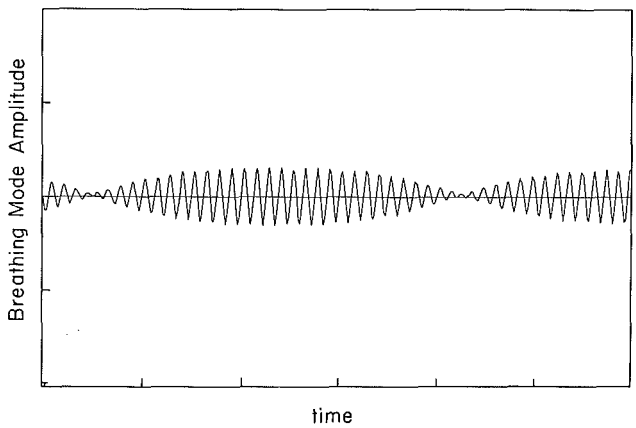


Fig. 7 Hopf bifurcation conditions. Time response history of the breathing mode response when $f = 0.1$.

Table 1 Summary of reponse with varying values of $\hat{\sigma}_1$

$\hat{\sigma}_1$	Trajectory	Phase Trajectory
< -0.075	fixed point	
$[-0.075, -0.031]$	limit cycle/ period T	Fig. 9
$[-0.03, -0.0297]$	limit cycle/ period 2T	Fig. 10
$[-0.0296, -0.02953]$	limit cycle/ period 4T	Fig. 11
$[-0.02950, 0]$	chaos	
>0	fixed point	

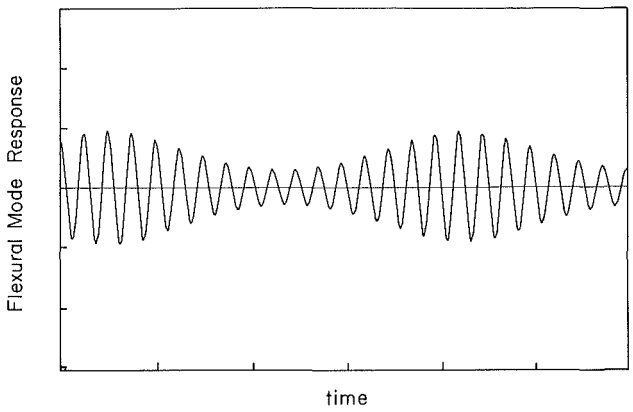


Fig. 8 Hopf bifurcation conditions. Time response history of the flexural mode response when $f = 0.1$.

finite-amplitude solution is unstable. Consequently, the response is periodic and consists of only the breathing mode when $f \leq \xi_2$. When $\xi_2 < f \leq \xi_3$, the response is also periodic having the same period as the excitation but it consists of a combination of the breathing mode and the 6th flexural mode. When $f > \xi_3$, the response is also a combination of the breathing mode and the 6th flexural mode but in this case the amplitudes and phases are not constants. Consequently, the response is either an amplitude and phase-modulated motion or a period multiplying (the period is an integral multiple of the excitation period) motion or a chaotic motion. The critical value $f = \xi_3$ is a Hopf bifurcation point at which the real part of a complex-conjugate pair of the roots of equation (56) changes sign.

Under the above stated conditions of complex-conjugate pair with positive real part, Fig. 6 shows variations of the long-time behavior of the amplitudes a_o and a_6 of the breathing and flexural modes when $f = 0.1$, obtained by numerically integrating equations (45)–(50) using a 6th order Runge Kutta algorithm. Under the same conditions, equations (12)–(14) are integrated and the results are shown in Fig. 7 (breathing mode) and in Fig. 8 (flexural mode).

To illustrate the importance of the Hopf bifurcation, we fix the parameters of the system, vary σ_1, and numerically integrate equations (45)–(48). The results show the system to exhibit a fixed-point response in the phase diagram before entering and after leaving the interval $-0.75 \leq \hat{\sigma}_1 \leq 0$. In this interval, the system exhibits a limit cycle behavior, then a cascade of period-doubling bifurcations starts to develop, leading to chaos. The system goes back to the fixed point behavior for $\hat{\sigma}_1 > 0$. Table 1 summarizes the behavior of the system with reference to the phase trajectory describing every pattern of response.

In Fig. 12, we show a representative variation of the amplitudes of the breathing and 6th flexural modes when $\Gamma > 0$. In addition to the saturation phenomenon, Fig. 12 exhibits the jump phenomenon. When the excitation amplitude f lies in the interval $[\xi_1, \xi_2]$, there are three possible steady-state solutions. Two of these solutions are stable: the trivial solution

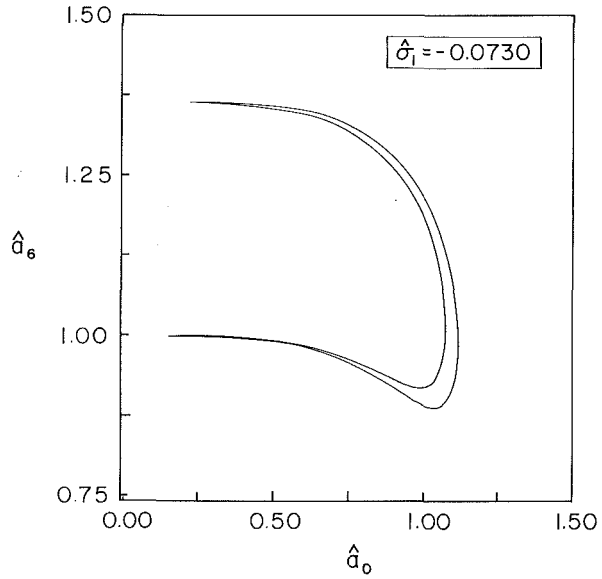


Fig. 9 Projection of the trajectory of the modulation equations on the $\hat{a}_o - \hat{a}_6$, plane, where $\hat{a}_o = a_o\,(f/\Lambda_2)^{-1/2}$, and $\hat{a}_6 = a_6\,(f/\Lambda_1)^{-1/2}$, for $\mu(f\Lambda_2)^{-1/2} = 0.02$, $\sigma_2(f\Lambda_2)^{-1/2} = 0.18$, and $\sigma_1(f\Lambda_2)^{-1/2} = -0.073$

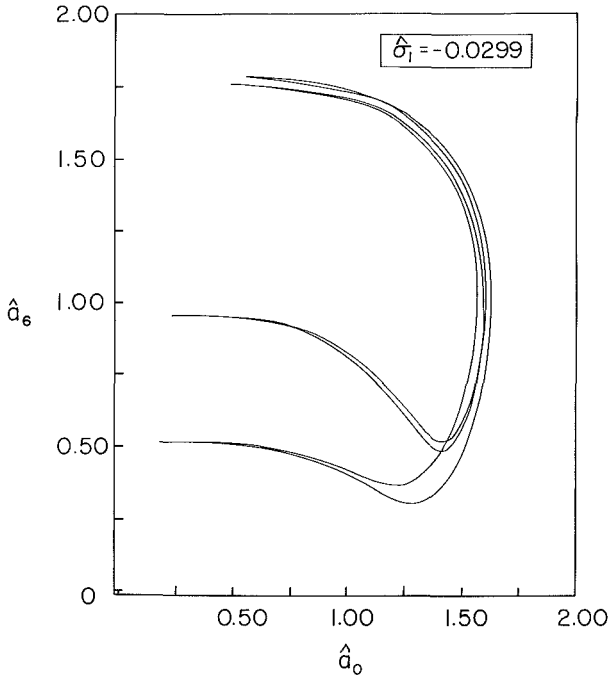


Fig. 10 Projection of the trajectory of the modulation equations on the $\hat{a}_o - \hat{a}_6$, plane, where $\hat{a}_o = a_o\,(f/\Lambda_2)^{-1/2}$, and $\hat{a}_6 = a_6\,(f/\Lambda_1)^{-1/2}$, for $\mu(f\Lambda_2)^{-1/2} = 0.02$, $\sigma_2(f\Lambda_2)^{-1/2} = 0.18$, and $\sigma_1(f\Lambda_2)^{-1/2} = -0.0299$

and the larger amplitude solution. The response that is attained physically depends on the initial conditions. If the excitation amplitude increases from zero, one observes only the breathing mode until f reaches ξ_2. As f increases beyond ξ_2, a_6 jumps up from zero to point C, producing a large wrinkling of the shell. As f increases further, a_o remains constant, whereas a_6 increases slowly along the curve ECD. If f decreases from a value corresponding to point D, a_6 decreases slowly along the curve DCE and a_o remains constant until point E is reached. As f decreases below ξ_1, a_6 jumps down to zero and a_o jumps down to point F. As f decreases further, a_6 remains zero and a_o decreases linearly with f.

If the amplitude of the excitation is set at a value in the in-

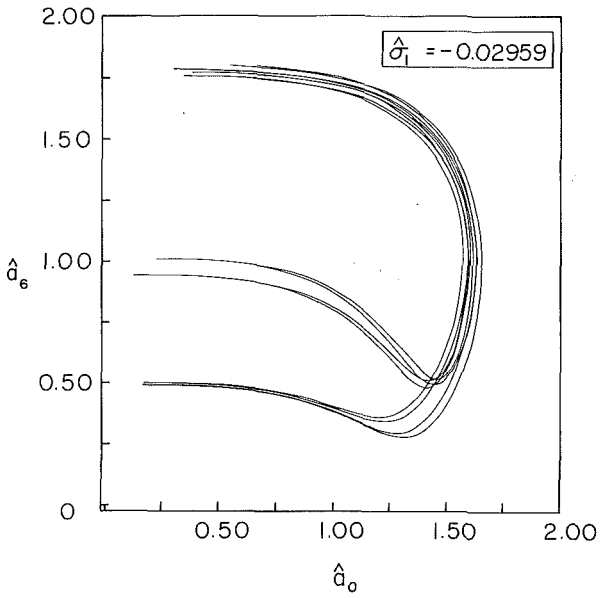


Fig. 11 Projection of the trajectory of the modulation equations on the $\hat{a}_0 - \hat{a}_6$, plane, where $\hat{a}_0 = a_0 (f\Lambda_2)^{-1/2}$, and $\hat{a}_6 = a_6 (f\Lambda_1)^{-1/2}$, for $\mu(f\Lambda_2)^{-1/2} = 0.02$, $\sigma_2(f\Lambda_2)^{-1/2} = 0.18$, and $\sigma_1(f\Lambda_2)^{-1/2} = -0.02959$

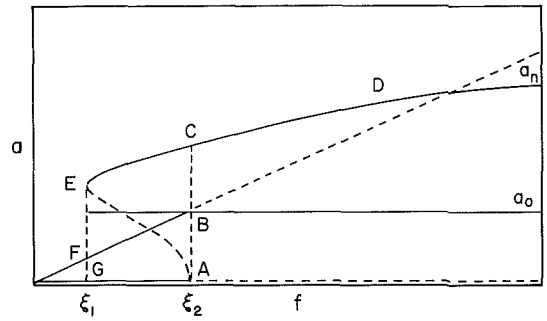


Fig. 12 Modal response amplitudes as functions of the amplitude of the excitation when $\Gamma > 0$

terval $[\xi_1, \xi_2]$ and the shell is initially undisturbed, the response corresponds to the linear solution, in which the shell is breathing without wrinkling. However, if the shell is disturbed, the shell may respond with the nonlinear solution, in which the amplitudes of the breathing and flexural modes increase dramatically, yielding a much larger response.

The instability of the linear solution and the saturation phenomenon were first found analytically and verified numerically by Nayfeh et al. (1973) in the response of ships. Later these phenomena were observed experimentally in the response of a simple model consisting of two beams and two concentrated masses by Haddow et al. (1984) and in the nonlinear vibration laboratory at VPI & SU.

Acknowledgment

This work was supported by the National Science Foundation under Grant No. MSM-852-1748.

References

Atluri, S., 1972, "A Perturbation Analysis of Non-Linear Free Flexural Vibrations of a Circular Cylindrical Shell," *International Journal of Solids and Structures*, Vol. 8, pp. 549–569.

Bieniek, M. P., Fan, T. C., and Lackman, L. M., 1966, "Dynamic Stability of Cylindrical Shells," *AIAA Journal*, Vol. 4, pp. 495–500.

Chen, J. C., 1972, "Nonlinear Vibration of Cylindrical Shells," Ph.D. Dissertation, California Institute of Technology, Pasadena, CA.

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

Goodier, J. N., and McIvor, I. K., 1964, "The Elastic Cylindrical Shell under Nearly Uniform Radial Impulse," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 259–266.

Haddow, A. G., Barr, A. D. S., Mook, D. T., 1984, "Theoretical and Experimental Study of Modal Interaction in a Two-Degree-of-Freedom Structure," *Journal of Sound and Vibration*, Vol. 97, pp. 451–473.

Hatwal, H., Mallik, A. K., and Ghosh, A., 1982, "Non-Linear Vibrations of a Harmonically Excited Autoparametric System," *Journal of Sound and Vibration*, Vol. 81, pp. 153–164.

Hui, D., 1983, "Large-Amplitude Vibrations of Geometrically Imperfect Shallow Spherical Shells with Structural Damping," *AIAA Journal*, Vol. 21, pp. 1736–1741.

McIvor, I. K., 1962, "Dynamic Stability and Nonlinear Oscillations of Cylindrical Shells (Plane Strain) Subjected to Impulsive Pressure," Ph.D. Dissertation, Stanford University, Stanford, CA.

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

McIvor, I. K., and Sonstegard, D. A., 1966, "Axisymmetric Response of a Closed Spherical Shell to a Nearly Uniform Radial Impulse," *The Journal of the Acoustical Society of America*, Vol. 40, pp. 1540–1547.

McIvor, I. K., and Lovell, E. G., 1968, "Dynamic Response of Finite-Length Cylindrical Shells to Nearly Uniform Radical Impulse," *AIAA Journal*, Vol. 6, pp. 2346–2351.

Mente, L. J., 1973, "Dynamic Nonlinear Response of Cylindrical Shells to Asymmetric Pressure Loading," *AIAA Journal*, Vol. 11, pp. 793–800.

Mettler, E., and Weidenhammer, F., 1962, "Zum Problem des Kinetischen Durchschlagens Schwach Gekrümmter Stäbe," Ingenieur Archiv, Vol. 31, pp. 421–432.

Miles, J. W., 1984, "Resonantly Forced Motion of Two Quadrically Coupled Oscillators," *Physica*, Vol. 13D, pp. 247–260.

Miles, J., 1985, "Parametric Excitation of an Internally Resonant Double Pendulum," *Journal of Applied Mathematical Physics* (*ZAMP*), Vol. 36, pp. 337–345.

Mook, D. T., Marshall, L. R., and Nayfeh, A. H., 1974, 'Subharmonic and Superharmonic Resonances in the Pitch and Roll Modes of Ship Motions," *Journal of Hydrodynamics*, Vol. 8, pp. 32–40.

Nayfeh, A. H., Mook, D. T., and Marshall, L. R., 1973, "Nonlinear Coupling of Pitch and Roll Modes in Ship Motion," *Journal of Hydronautics*, Vol. 7, pp. 145–152.

Nayfeh, A. H., 1973. *Perturbation Methods*, Wiley-Interscience, New York.

Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, Wiley-Interscience, New York,

Nayfeh, A. H., 1981, *Introduction to Perturbation Techniques*, Wiley-Interscience, New York.

Nayfeh, A. H., and Zavodney, L. D., 1986, "The Response of Two-Degree-of-Freedom System with Quadratic Nonlinearities to a Combination Parametric Resonance," *Journal of Sound and Vibration*, Vol. 107, pp. 329–350.

Nayfeh, A. H., and Raouf, R. A., 1986, "Nonlinear Forced Response of Circular Cylindrical Shells," ASME *Design and Analysis of Plates and Shells*, Vol. 105, pp. 145–155.

Nayfeh, A. H., 1987, "Parametric Excitation of Two Internally Resonant Oscillators," *Journal of Sound and Vibration*, Vol. 119, in press.

Raouf, R. A., 1985, "Nonlinear Forced Response of Circular Cylindrical Shells," M.S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Reissner, E., 1955, "On Axi-Symmetrical Vibrations of Shallow Spherical Shells," *Quarterly of Applied Mathematics*, Vol. 13, pp. 279–290.

Sethna, P. R., 1965, "Vibrations of Dynamics Systems with Quadratic Nonlinearities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 576–582.

Simmonds, J. G., 1979, "Accurate Nonlinear Equations and a Perturbation Solution for the Free Vibrations of a Circular Elastic Ring," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 156–160.

Yamamoto, T., and Yasuda, K., 1977, "On the Internal Resonance in a Nonlinear Two-Degree-of-Freedom System," *Bulletin of the Japanese Society of Mechanical Engineers*, Vol. 20, pp. 169–175.

Large Elastic Deformation of Shear Deformable Shells of Revolution: Theory and Analysis

L. A. Taber
Department of Mechanical Engineering,
University of Rochester,
Rochester, NY 14627

Large axisymmetric deformation of pressurized shells of revolution is studied. The governing equations include the effects of transverse normal strain and transverse shear deformation for shells composed of an incompressible, hyperelastic material. Asymptotic solutions to the equations are developed which are valid for moderately large strains. Application to Mooney-Rivlin clamped spherical caps reveals that, for large enough bending and stretching, the consequences of shear deformation include: (1) bending moments can decrease *at the edge after the load passes a critical point; (2) even thick shells can behave as membranes; (3) transition points can occur in the shell which divide regions of shell-like behavior from regions of membrane-like behavior.*

Introduction

Large elastic deformation of shells has gained renewed interest in recent years. Although governing equations for large strain have been available since the 1950s, the extreme complexity of these relations (e.g., Naghdi, 1972) has persuaded most researchers to focus on membrane solutions. Much recent effort has, therefore, been devoted toward the development of approximate equations that are more amenable to analysis (Libai and Simmonds, 1981; Taber, 1985; Simmonds, 1986). Although further refinement of these relations, especially the boundary conditions, undoubtedly will occur in the future, the intent here is to gain insight into the fundamental behavior of elastic shells undergoing large bending and stretching. Thus, using these equations at their current stage of development, this paper examines some basic shell bending problems through asymptotic analysis.

A pair of previous publications (Taber, 1987a,b) presented asymptotic expansions for large axisymmetric deformation of rubber-like circular plates and cylindrical shells. These and other recent papers (Libai and Simmonds, 1981; Keppel, 1984; Simmonds, 1986; Brodland, 1986) have relaxed the Kirchhoff hypothesis to allow for thickness changes but not transverse shear strains. Justification for such a theory is based in part on the work of John (1965), who has shown that, at least for small strains, transverse shear stresses are only $O(\sigma t/L)$, where σ is the norm of the stresses in the shell, t is the thickness, and L is the minimum "wavelength" of the deformation pattern. However, if this conclusion can be extended to large strains, then transverse shear deformation can become significant near edges and concentrated loads of even "thin" shells, where L can grow quite small. Indeed, a recent publication by Meroueh (1986) demonstrates this behavior for cylindrical deformation. Thus, the current work extends and generalizes that of Taber (1987a,b) to include the following features: (1) general axisymmetric geometry; (2) incompressible material properties characterized by a general strain energy density function; and (3) transverse shear deformation.

The shell equations employed herein represent essentially a combination of the field equations for a shell of revolution developed by Reissner (1969, 1972) and the two-dimensional strain energy density function of Simmonds (1986). With the latter slightly modified to allow transverse shear strains, this Reissner-Simmonds (R-S) shell theory assumes that lines originally normal to the reference surface remain straight but not necessarily normal after deformation. The results of Meroueh (1986) indicate that this is not a bad assumption, even for very large bending. In addition, these lines can change in length to incorporate transverse normal strains. A key ingredient of this theory is the constitutive coupling between bending and stretching that occurs during large strain in isotropic shells (Libai and Simmonds, 1981; Taber, 1985), which is similar to that encountered in linear theories for unsymmetrically layered laminates. In its current form, R-S shell theory is valid for large membrane and "moderately large" bending and shear strains, i.e., $[\gamma^2, (t\kappa)^2] << 1$, with γ and κ being the largest transverse shear strain and curvature change measure, respectively.

This paper first examines the effects of thickness changes on the governing equations of Reissner (1969, 1972) and then adds transverse shear effects to the strain energy density function of Simmonds (1986). Next, based on these R-S equations, expansions are developed for pressurized shells of revolution. And finally, results are presented for the special case of a clamped, spherical cap. As in the case of a cylinder (Taber,

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February 23, 1987; final revision, March 19, 1987.
Paper No. 87-WA/APM-10.

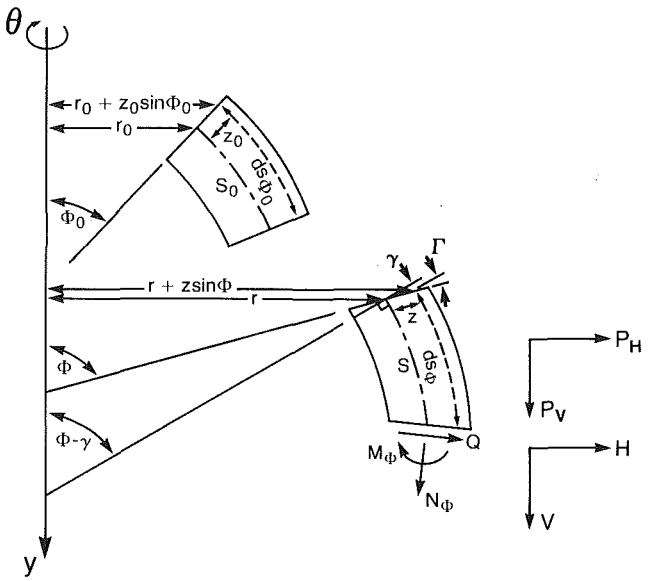


Fig. 1 Geometry for shell element

1987b), the results show the presence of a primary boundary layer, and bending and stretching components of a secondary boundary layer. Details of the primary layer, which was first identified in membranes by Bromberg and Stoker (1945), and of the bending portion of the secondary layer were studied for small strain by Reissner (1959). On the other hand, the secondary stretching component develops as strains grow large (Taber, 1987a,b).

Governing Equations

Geometry. The geometry of a shell element before and after deformation is shown in Fig. 1, in which subscript zero indicates values in the original configuration. While the *undeformed* reference surface S_0 is taken as the middle surface of the shell, two definitions for the *deformed* reference surface S will be explored. First, as in Simmonds (1986), the position of S will be defined through a "dynamic (or static) consistency condition" (Theory I)

$$\int_{-t/2}^{t/2} z\, dz_0 = 0 \tag{1}$$

where t is the undeformed shell thickness, and z_0 and z are the undeformed and deformed transverse coordinates (Fig. 1). With this definition, the material composition of S is not necessarily the same as that of S_0 and can change with the deformation. Second, as in Taber (1985), the reference surface will be assumed to be made up of the same material points, regardless of the deformation (Theory II). In general, S will not be located at the geometric midsurface. These differing criteria for locating the reference surface mainly affect the form of the constitutive relations, which are discussed in a later section.

As a first approximation, transverse shear deformation enters in the manner of Reissner (1969), who employed an extension of Timoshenko beam theory. Accordingly, a line normal to S_0 and at a meridional angle Φ_0 relative to the axis of symmetry remains straight during deformation but rotates an additional amount γ relative to S due to shear, forming an angle Φ with the shell axis (Fig. 1). Furthermore, transverse normal strains are accounted for through the assumption that a point in the shell, originally a distance z_0 from S_0, moves to a distance z from S along this nonorthogonal line, where the (engineering) shear strain is Γ.

For axisymmetric deformation, it is convenient to work in terms of the stretch ratios

$$\Lambda_\Phi = ds_\Phi/ds_{\Phi_0}, \quad \Lambda_\theta = ds_\theta/ds_{\theta_0}, \quad \Lambda_z = \partial z/\partial z_0, \tag{2}$$

and, with the kinematic assumptions outlined above, the geometry of Fig. 1 gives the relations

$$\cos\Phi_0 ds_{\Phi_0} = d(r_0 + z_0\sin\Phi_0), \quad \sin\Phi_0 ds_{\Phi_0} = d(y_0 - z_0\cos\Phi_0)$$

$$ds_{\theta_0} = (r_0 + z_0\sin\Phi_0)d\theta \tag{3a}$$

$$\cos(\Phi-\Gamma)ds_\Phi = d(r + z\sin\Phi), \quad \sin(\Phi-\Gamma)ds_\Phi = d(y - z\cos\Phi)$$

$$ds_\theta = (r + z\sin\Phi)d\theta. \tag{3b}$$

After substitution of equations (3) into (2), the two forms for ds_{Φ_0} and ds_Φ provide a pair of equations to be solved for Λ_Φ and Γ, giving

$$\bar{\Lambda}_\Phi = \Lambda_\Phi\cos\Gamma = \frac{\bar{\lambda}_\Phi + zk_\Phi}{1 + z_0k_{\Phi_0}}, \quad \Lambda_\theta = \frac{\lambda_\theta + zk_\theta}{1 + z_0k_{\theta_0}}$$

$$\bar{\Gamma} = \Lambda_\Phi\sin\Gamma = \frac{\bar{\gamma} + z'}{1 + z_0k_{\Phi_0}} \tag{4}$$

where the bars indicate modified strain measures and

$$(\,)' \equiv d(\,)/(ds_{\Phi_0})_{s_0} \equiv d(\,)/dx.$$

At the reference surface, the strain measures are

$$\bar{\lambda}_\Phi = \lambda_\Phi\cos\gamma = r'\cos\Phi + y'\sin\Phi, \quad \lambda_\theta = r/r_0$$

$$\bar{\gamma} = \lambda_\Phi\sin\gamma = r'\sin\Phi - y'\cos\Phi \tag{5}$$

and the undeformed and deformed curvature measures are

$$k_{\Phi_0} = \Phi_0', \quad k_\Phi = \Phi'; \quad k_{\theta_0} = \sin\Phi_0/r_0, \quad k_\theta = \sin\Phi/r_0. \tag{6}$$

Outside of the appearance of z', equations (4) are equivalent to those of Reissner (1969, 1972), who neglected thickness changes. In addition, these relations do not include several higher-order terms as given by the more rigorous derivation of Simmonds (1986).

To render the analysis more manageable, we now introduce two further assumptions. The first is the usual thin-shell approximation $(tk_{\Phi_0}, tk_{\theta_0}) << 1$, and the second is $\Gamma^2 << 1$. Thus, to this point of the development, the analysis is valid for large bending and membrane strains but only moderately large transverse shear strains, and equations (4) can be written

$$\bar{\Lambda}_\Phi \approx \Lambda_\Phi \approx \lambda_\Phi + z\kappa_\Phi, \quad \Lambda_\theta \approx \lambda_\theta + z\kappa_\theta$$

$$\bar{\Gamma} \approx \Lambda_\Phi\Gamma \approx \bar{\gamma} \approx \lambda_\Phi\gamma \tag{7}$$

where

$$\kappa_\Phi = k_\Phi - \lambda_\Phi^2\lambda_\theta k_{\Phi_0}, \quad \kappa_\theta = k_\theta - \lambda_\Phi\lambda_\theta^2 k_{\theta_0} \tag{8}$$

are curvature change measures (Taber, 1985). Note also that the z' term of $\bar{\Gamma}$ has been dropped. Later, it will be shown that this term should not be included in a first approximation shear deformation theory.

Equilibrium. With stress and moment resultants (Fig. 1) defined per unit *undeformed* length of the reference surface, vertical and horizontal force equilibrium and moment equilibrium yield[1] (Reissner, 1969, 1972)

$$(r_0V)' + r_0p_v = 0, \quad (r_0H)' - N_\theta + r_0p_H = 0$$

$$(r_0M_\Phi)' - M_\theta\cos\Phi - r_0(Q\bar{\lambda}_\phi - N_\Phi\bar{\gamma}) = 0. \tag{9}$$

Also, the geometry gives

$$N_\Phi = H\cos\Phi + V\sin\Phi, \quad Q = H\sin\Phi - V\cos\Phi \tag{10}$$

in terms of the vertical and horizontal force components V

[1]The moment equilibrium equation ignores the contribution due to the moment about the normal to the reference surface, which was included by Reissner (1969) but shown by Simmonds (1986) to be of higher order.

and H, respectively. For a uniform internal pressure p, the vertical and horizontal surface loads per unit area of S_0 are

$$p_V = -p\lambda_\Phi\lambda_\theta\cos(\Phi-\gamma), \quad p_H = p\lambda_\Phi\lambda_\theta\sin(\Phi-\gamma) \tag{11}$$

and equation $(9)_1$ can be integrated to give

$$V = pr^2/2r_0. \tag{12}$$

Constitutive Relations. For a shell composed of a hyperelastic material, the two-dimensional strain energy density w is defined per unit area of S_0. Via the principle of virtual work, Reissner (1972) determined that the constitutive relations can be written

$$N_\Phi = \frac{\partial w}{\partial \bar{\lambda}_\Phi}, \quad N_\theta = \frac{\partial w}{\partial \lambda_\theta}, \quad Q = \frac{\partial w}{\partial \bar{\gamma}},$$

$$M_\Phi = \frac{\partial w}{\partial \kappa_\Phi}, \quad M_\theta = \frac{\partial w}{\partial \kappa_\theta}. \tag{13}$$

The appearance of $\bar{\lambda}_\Phi$ and $\bar{\gamma}$ instead of λ_Φ and γ in these equations is due to the fact that, in general, N_Φ and Q are not parallel and perpendicular, respectively, to S (Fig. 1).

Boundary Conditions. Recently, Gregory and Wan (1985) have shown that, at least for linear plate theory, the exact distribution of edge tractions across the plate thickness is more important than previously thought. Even the solution in the plate interior can be affected significantly. For shells, it may be that the curvature effects contain this behavior within the edge zones, but this requires further study, especially for large deformations. Since a detailed investigation of the boundary conditions is beyond the scope of the present paper, only approximate conditions consistent with the field equations will be used here. The principle of virtual work shows that the appropriate boundary conditions are to specify M_Φ or Φ, and H or r, and V or y at each shell edge (Reissner, 1969).

Strain Energy Density Function

The forms of the preceding equations are not significantly different from those of a small-strain shell theory. For large strain, the major additional complexity lies in the constitutive behavior contained in the two-dimensional strain energy density function w. As discussed by Reissner (1974) and Libai and Simmonds (1983), two methods can be used to determine w for a shell material. One way involves experiments on two-dimensional samples of the material (direct method), and the other obtains w through transverse integration of the three-dimensional energy density W (reduction method). Here, following the analysis and much of the notation of Simmonds (1986), we employ the reduction method to compute an approximate form for w consistent with the strain measures derived earlier. Specifically, w is developed as an asymptotic series in powers of the small parameter

$$\epsilon = t/2L. \tag{14}$$

Three-Dimensional Relations. Upon introduction of the nondimensional quantities

$$(z_0^*, z^*) = (2/t)(z_0, z), \quad (\kappa_\Phi^*, \kappa_\theta^*) = L(\kappa_\Phi, \kappa_\theta),$$
$$(\bar{\Gamma}^*, \bar{\gamma}^*) = (\bar{\Gamma}, \bar{\gamma})/\epsilon \quad W^* = W/C, \quad w^* = w/Ct, \tag{15}$$

where L is the deformation wavelength and C is a material constant with units of a Young's modulus, equations $(2)_3$ and (7) become

$$\Lambda_\Phi = \lambda_\Phi + \epsilon z^*\kappa_\Phi^*, \quad \Lambda_\theta = \lambda_\theta + \epsilon z^*\kappa_\theta^*, \quad \Lambda_z = \partial z^*/\partial z_0^*,$$
$$\bar{\Gamma} = \epsilon\bar{\gamma}^*. \tag{16}$$

The current derivation is valid for any incompressible, elastic material that allows axisymmetric deformation without twisting—for example, a polar orthotropic material. However, this paper focuses later on isotropic shells for which

$$W = W(I_1, I_2), \quad I_3 = 1 \tag{17}$$

where W is the strain energy density per unit undeformed volume. The strain invariants are

$$I_1 = G_\Phi + G_\theta + G_z, \quad I_2 = G_\Phi G_\theta + G_\Phi G_z + G_\theta G_z - G_\Gamma,$$
$$I_3 = G_\Phi G_\theta G_z - G_\Gamma G_\theta \tag{18}$$

where equations (16) give

$$G_\Phi = \Lambda_\Phi^2 = \lambda_\Phi^2 + 2\lambda_\Phi\kappa_\Phi^* z^*\epsilon + \kappa_\Phi^{*2} z^{*2}\epsilon^2 = \overset{0}{G}_\Phi + \overset{1}{G}_\Phi\epsilon + \overset{2}{G}_\Phi\epsilon^2$$
$$G_\theta = \Lambda_\theta^2 = \lambda_\theta^2 + 2\lambda_\theta\kappa_\theta^* z^*\epsilon + \kappa_\theta^{*2} z^{*2}\epsilon^2 = \overset{0}{G}_\theta + \overset{1}{G}_\theta\epsilon + \overset{2}{G}_\theta\epsilon^2$$
$$G_z = \Lambda_z^2 = (\partial z^*/\partial z_0^*)^2, \quad G_\Gamma = \bar{\Gamma}^2 = \epsilon^2\bar{\Gamma}^{*2}. \tag{19}$$

Now, on enforcement of the incompressibility condition $I_3 = 1$, equation $(18)_3$ can be solved for G_z, and then

$$W = W(G_\Phi, G_\theta, G_\Gamma) \tag{20}$$

and

$$I_1 = G_\Phi + G_\theta + 1/G_\Phi G_\theta + G_\Gamma/G_\Phi,$$
$$I_2 = G_\Phi G_\theta + 1/G_\Phi + 1/G_\theta + G_\Gamma G_\theta/G_\Phi. \tag{21}$$

Reduction to Two Dimensions. Simmonds (1985) has shown that, for moderately large bending strains $[(t\kappa)^2 << 1]$, the form of the two-dimensional strain energy function for a shell is the same as that for a flat plate. Therefore, the reduction to two dimensions is given by

$$w^* = \frac{1}{2}\int_{-1}^{1} W^* dz_0^*. \tag{22}$$

Since $G_\Gamma << 1$ for moderate transverse shear strains, W can be expanded in the Taylor series

$$W^* = [W^*]_{G_\Gamma=0} + [W^*,_{G_\Gamma}]_{G_\Gamma=0}G_\Gamma + \ldots \tag{23}$$

which is substituted into equation (22). The first term of the resulting integral, which was computed by Simmonds (1986), corresponds to a first-approximation shell theory that neglects shear deformation, and the second term provides a correction due to shear.

In his approach, Simmonds (1986) sought an expansion about the membrane state given by $\epsilon \to 0$ through a series of the form

$$A(G_\Phi, G_\theta) = \overset{0}{A} + \overset{0}{A}_{(1,0)}(G_\Phi - \overset{0}{G}_\Phi) + \overset{0}{A}_{(0,1)}(G_\theta - \overset{0}{G}_\theta)$$
$$+\frac{1}{2}[\overset{0}{A}_{(2,0)}(G_\Phi - \overset{0}{G}_\Phi)^2 + \overset{0}{A}_{(0,2)}(G_\theta - \overset{0}{G}_\theta)^2$$
$$+2\overset{0}{A}_{(1,1)}(G_\Phi - \overset{0}{G}_\Phi)(G_\theta - \overset{0}{G}_\theta)] + \ldots \tag{24}$$

where

$$A_{(m,n)} \equiv \frac{\partial^{m+n}A}{\partial G_\Phi^m \partial G_\theta^n}, \quad \overset{0}{A} \equiv (A)_{\epsilon=0}.$$

After substitution of equations (19) into (24), setting $A = [W^*]_{G_\Gamma=0}$ and $A = [W^*,_{G_\Gamma}]_{G_\Gamma=0}$ gives these terms as functions of z^* and ϵ. Subsequent substitution into equations (22) and (23) then yields

$$w^* = (\overset{0}{W^*})_{G_\Gamma=0} + \Big[(\lambda_\Phi\kappa_\Phi^* \overset{0}{W}^*_{(1,0)}$$
$$+\lambda_\theta\kappa_\theta^* \overset{0}{W}^*_{(0,1)})_{G_\Gamma=0}\int_{-1}^{1} z^* dz_0^*\Big]\epsilon$$

$$+\frac{1}{2}\Big[\{(\overset{0}{W}{}^{*}_{(1,0)}+2\lambda_\Phi^2\overset{0}{W}{}^{*}_{(2,0)})\kappa_\Phi^{*2}$$

$$+(\overset{0}{W}{}^{*}_{(0,1)}+2\lambda_\theta^2\overset{0}{W}{}^{*}_{(0,2)})\kappa_\theta^{*2}$$

$$+4\overset{0}{W}{}^{*}_{(1,1)}\lambda_\Phi\lambda_\theta\kappa_\Phi^*\kappa_\theta^*\}_{G_\Gamma=0}\int_{-1}^{1}z^{*2}dz_0^*$$

$$+2g^{-1}(\overset{0}{W}{}^{*},_{G_\Gamma})_{G_\Gamma=0}\bar{\gamma}^{*2}\Big]\epsilon^2+O(\epsilon^3) \tag{25}$$

where

$$\frac{1}{2}\int_{-1}^{1}\bar{\Gamma}^{*2}dz_0^*\equiv\bar{\gamma}^{*2}/g \tag{26}$$

has been introduced, with g being the reciprocal "form factor in shear." For linear shell theory, the transverse shear stress distribution is parabolic over the shell thickness, leading, in some derivations, to the value $g=1.2$ (Reissner, 1952), which is used here. The determination of an accurate value for large strain is left to future study.

Next, the incompressibility condition provides z in terms of z_0. As in Simmonds (1986), the asymptotic expansion

$$z^*=z^{(0)*}+\epsilon z^{(1)*}+\ \ldots \tag{27}$$

is substituted into equations $(18)_3$ and (19) with $I_3=1$. Equating coefficients of like powers of ϵ gives

$$z^{(0)*}=z_0^*/\lambda_\Phi\lambda_\theta,$$

$$z^{(1)*}=-(3z_0^{*2}-q)(\lambda_\Phi\kappa_\theta^*+\lambda_\theta\kappa_\Phi^*)/6\lambda_\Phi^3\lambda_\theta^3 \tag{28}$$

where q is introduced to distinguish between Theory I, which satisfies equation (1), and Theory II, which does not, i.e.,

$$q=\begin{cases}1, \text{ Theory I}\\ 0, \text{ Theory II.}\end{cases}$$

At this point, note that, if z' is kept in $\bar{\Gamma}$ as given by equation (4), then derivatives of the strains would appear in the expression for w. For a first-approximation shell theory (including shear deformation), we assume that stresses depend on only the local strains and, therefore, drop these terms.

Finally, substitution of equations (27) and (28) into equation (25) and subsequent integration give, in terms of new nondimensional variables,

$$w^*=w_0+w_1\bar{\gamma}^2/g+\beta^{-2}(w_2\hat{\kappa}_\Phi^2+w_3\hat{\kappa}_\Phi\hat{\kappa}_\theta+w_4\hat{\kappa}_\theta^2) \tag{29}$$

where the $w_n(\lambda_\Phi,\lambda_\theta)$ are

$$w_0=(\overset{0}{W}{}^{*})_{G_\Gamma=0},\quad w_1=(\overset{0}{W}{}^{*},_{G_\Gamma})_{G_\Gamma=0}$$

$$w_2=\alpha^2/(12\lambda_\Phi^2\lambda_\theta^2)[q\overset{0}{W}{}^{*}_{(1,0)}+2\lambda_\Phi^2\overset{0}{W}{}^{*}_{(2,0)}]_{G_\Gamma=0}$$

$$w_3=\alpha^2/(12\lambda_\Phi^3\lambda_\theta^3)[4\lambda_\Phi^2\lambda_\theta^2\overset{0}{W}{}^{*}_{(1,1)}$$

$$-(1-q)(\lambda_\Phi^2\overset{0}{W}{}^{*}_{(1,0)}+\lambda_\theta^2\overset{0}{W}{}^{*}_{(0,1)})]_{G_\Gamma=0}$$

$$w_4=\alpha^2/(12\lambda_\Phi^2\lambda_\theta^2)[q\overset{0}{W}{}^{*}_{(0,1)}+2\lambda_\theta^2\overset{0}{W}{}^{*}_{(0,2)}]_{G_\Gamma=0}. \tag{30}$$

To facilitate the following asymptotic analysis, the new nondimensional quantities are defined as

$$(\hat{\kappa}_\Phi,\hat{\kappa}_\theta)=a(\kappa_\Phi,\kappa_\theta)$$

$$\alpha=pa/Ct,\quad \beta=pa^2/Ct^2 \tag{31}$$

where a is a characteristic length such as a reference radius of curvature of the shell. For $\bar{\gamma}=0$ and $q=1$, equation (29) reduces to the expression given by Simmonds (1986), while $\bar{\gamma}=q=0$ produces the relation given in Taber (1985). Specialization to a Mooney-Rivlin material is provided in Appendix A.

Asymptotic Analysis

Nondimensional Equations. The foregoing analysis is based on the nondimensional quantities

$$(r^*,r_0{}^*,x^*,y^*)=(r,r_0,x,y)/a,$$

$$(\kappa_\Phi^*,\kappa_\theta^*)=(\hat{\kappa}_\Phi,\hat{\kappa}_\theta)=a(\kappa_\Phi,\kappa_\theta)$$

$$(N_\Phi^*,N_\theta^*,Q^*,H^*,V^*)=(N_\Phi,N_\theta,Q,H,V)/pa$$

$$(M_\Phi^*,M_\theta^*)=(M_\Phi,M_\theta)/Ct^2,\quad w^*=w/Ct \tag{32}$$

along with α and β as defined in equation (31). In addition, we introduce the notation

$$c_0=\cos\Phi_0,\quad s_0=\sin\Phi_0,\quad c=\cos\Phi,\quad s=\sin\Phi$$

$$\phi=\sin\omega,\quad \eta=\cos\omega \tag{33}$$

where $\omega=\Phi-\Phi_0$ is the rotation of a meridional *face* of the shell (Fig. 1). After substitution of equations (32) and (33) into equations (5), (6), and (8)–(13), removal of the stars yields the governing relations for a pressurized shell in the form

$$r'=\lambda_\Phi c+\bar{\gamma}s,\quad y'=\lambda_\Phi s-\bar{\gamma}c,\quad \lambda_\theta=r/r_0$$

$$\kappa_\Phi=\phi'/\eta+(1-\lambda_\Phi^2\lambda_\theta\Phi_0'),\quad \kappa_\theta=(s-\lambda_\Phi\lambda_\theta^2 s_0)/r_0 \tag{34a}$$

$$N_\Phi=\frac{1}{\alpha}\frac{\partial w}{\partial\lambda_\Phi},\quad N_\theta=\frac{1}{\alpha}\frac{\partial w}{\partial\lambda_\theta},\quad Q=\frac{1}{\alpha}\frac{\partial w}{\partial\bar{\gamma}}$$

$$M_\Phi=\frac{\beta}{\alpha}\frac{\partial w}{\partial\kappa_\Phi},\quad M_\theta=\frac{\beta}{\alpha}\frac{\partial w}{\partial\kappa_\theta} \tag{34b}$$

$$V=r^2/2r_0,\quad (r_0H)'-N_\theta+r_0\lambda_\theta(\lambda_\Phi s-\bar{\gamma}c)=0$$

$$\beta^{-1}[(r_0M_\Phi)'-cM_\theta]-r_0(Q\lambda_\Phi-N_\Phi\bar{\gamma})=0 \tag{34c}$$

$$s=\eta s_0+\phi c_0,\quad c=\eta c_0-\phi s_0,\quad \eta^2+\phi^2=1$$

$$N_\Phi=Hc+Vs,\quad Q=Hs-Vc. \tag{34d}$$

Formal Expansions. The asymptotic analysis follows those in Taber (1987a,b), which ignore transverse shear strains in treating the special cases of circular plates and cylinders composed of neo-Hookean material. This work extends those analyses to general shells of revolution composed of a general incompressible, hyperelastic material and also adds the effect of transverse shear deformation.

Briefly, each dependent variable is expanded in the form

$$\{\psi(x),\kappa_\Phi(x),\kappa_\theta(x)\}=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\beta^{-m}e^{n\beta\xi(x)}$$

$$\times\{\psi^{(mn)}(x),\beta\kappa_\Phi^{(mn)}(x),\beta\kappa_\theta^{(mn)}(x)\} \tag{35}$$

where ψ represents any dependent variable except κ_Φ and κ_θ. These expansions are valid for moderately large membrane strains, so that $\alpha=O(1)$, and thus, for a thin shell, $\beta>>1$ (equation (31)). Similar expansions for small strain were given by Ranjan and Steele (1980) and Steele (1980) for cases without and with transverse shear deformation, respectively. Note that, while the approximation for w (equation (29)) limits the magnitude of the bending and shear strains, the asymptotic analysis limits the magnitude of the membrane strains so that all strains can be only moderately large. In equation (35), the $n=0$ terms give the interior solution, and the exponential ($n>0$) terms provide the edge-zone solution. For $\beta>8$, $\xi(x)$ is a real-valued decay function of O(1) (Taber, 1987b), and so the $\psi^{(mn)}$ will be real.

As shown by Taber (1987a,b), the main features of the solution are given by the $n=0$, 1, and 2 terms in equation (35). After substitution into equations (29) and (34), like powers of

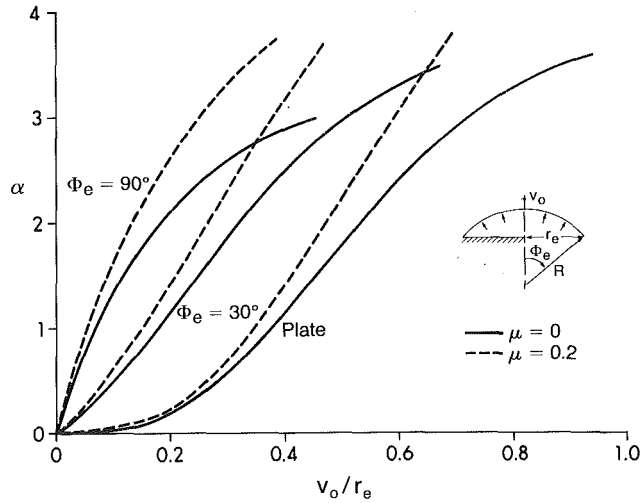


Fig. 2 Load-deflection curves for pressurized, spherical membrane caps

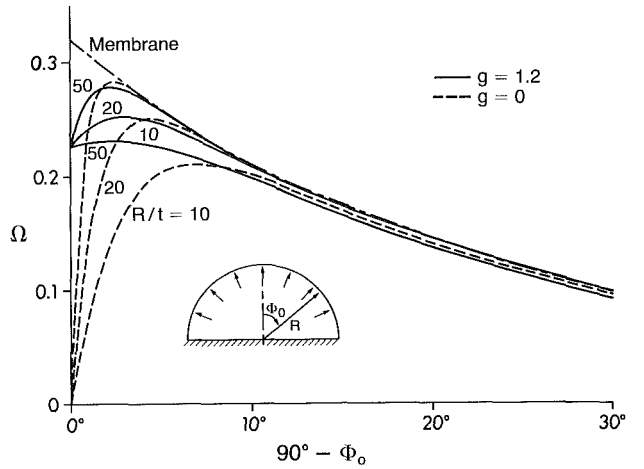


Fig. 3 Meridional rotation profiles near edge of clamped hemispherical cap ($\mu = 0$, $\alpha = 2$)

β, $\beta e^{\beta\xi}$, etc. are equated, giving a first-approximation solution of the form

$$\psi = \psi^{(00)} + \beta^{-1}\psi^{(10)} + e^{\beta\xi}\psi^{(01)} + e^{2\beta\xi}\psi^{(02)}. \tag{36}$$

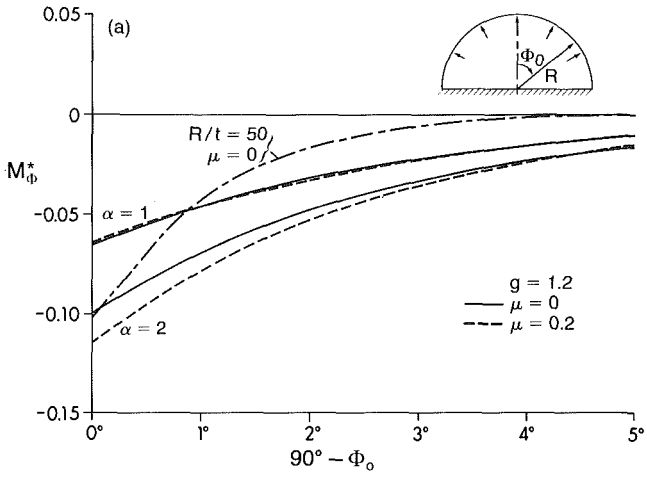
Here, the (00) terms represent the solution of nonlinear membrane theory, the (10) terms give the interior bending moments, and the (01) and (02) terms represent bending and membrane components, respectively, of the edge-zone solution. Details of the solution procedure are very similar to those expounded upon in Taber (1987a,b) and will not be repeated here. However, the terms necessary for the solution (36) are presented in Appendix B. For $g = 0$, this solution contains no shear deformation.

Results for a Clamped Spherical Cap

This section applies the preceding development to the special case of a pressurized spherical cap clamped around its edge. For this geometry, $a = R$, and the material is rubber-like with a Mooney-Rivlin strain energy density function

$$W = C[(I_1 - 3) + \mu(I_2 - 3)] \tag{37}$$

where C and μ are material constants, and I_1 and I_2 are given by equations (21). Most of the features discussed in Taber (1987a,b) for neo-Hookean ($\mu = 0$) plates and cylinders without shear deformation ($g = 0$) also occur for spherical



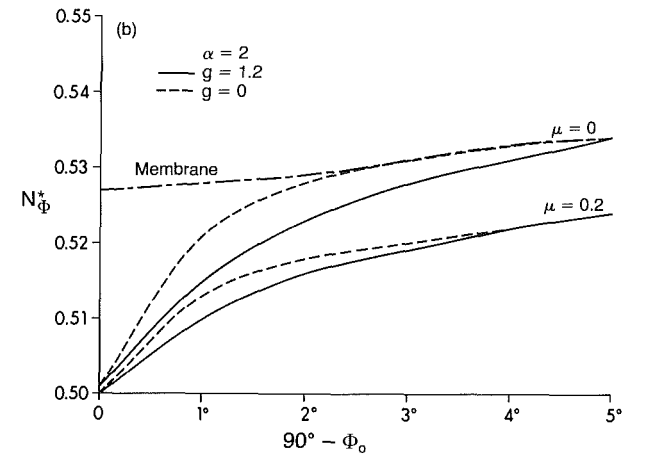


Fig. 4 Bending moment (*a*) and membrane stress resultant (*b*) distributions near edge of clamped hemispherical cap ($R/t = 20$)

geometry. This paper, therefore, focuses on the effects of shear deformation ($g = 1.2$) and the value of μ.

While curves of load versus maximum deflection for membranes (Fig. 2) differ little from those of shells, the local deformation pattern near the edge depends strongly on bending and shear effects. For example, Fig. 3 shows the meridional rotation $\Omega = \omega - \gamma$ of the reference surface near the clamped edges of a hemispherical membrane and shells of three different thicknesses. The primary boundary layer of Bromberg and Stoker (1945) appears but is not pronounced in the membrane solution. On the other hand, the boundary condition of zero rotation for $g = 0$ leads to the development of the readily apparent bending component of the secondary boundary layer. Inclusion of transverse shear deformation ($g = 1.2$) relaxes the edge restraint considerably, letting the deformed shape approach that of a membrane, even for a thick shell with $R/t = 10$. In addition, the curves for $R/t = 50$ indicate that $L/t \approx 2$, and so, although this shell fits the classical definition of a "thin shell," such a short wavelength leads to the significant shear deformation effects in the edge zone.

Bending moments and membrane stress resultants (Fig. 4) illustrate the bending and the stretching components, respectively, of the secondary boundary layer. Note that the latter component is required to satisfy the equilibrium condition $N_\Phi^* = 0.5$ at the clamped edge of hemispherical shells (Taber, 1987b). The force resultant distributions depend on the decay function, which is (see Appendices A and B)

$$\xi(x) = \int_{x_e}^{x} \left[\frac{6\lambda_\Phi^{(00)7}\lambda_\theta^{(00)4}N_\Phi^{(00)}}{\alpha(3 + \lambda_\Phi^{(00)4}\lambda_\theta^{(00)2})(1 + \mu\lambda_\theta^{(00)2})}\, G \right]^{1/2} dx \tag{38a}$$

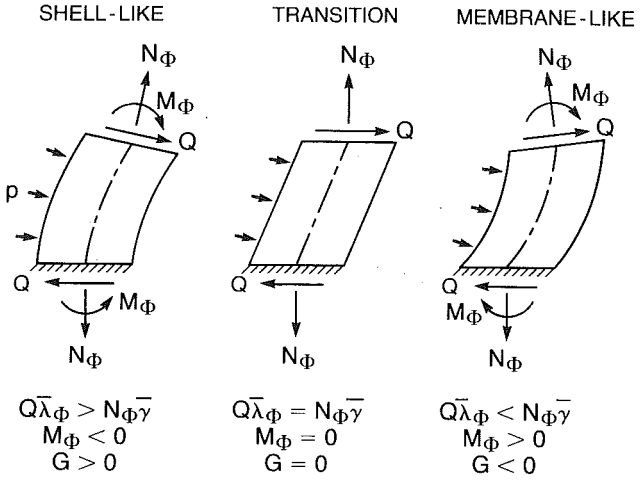


Fig. 5 Transition from shell-like to membrane-like behavior at edge of clamped, pressurized cylinder with shear deformation

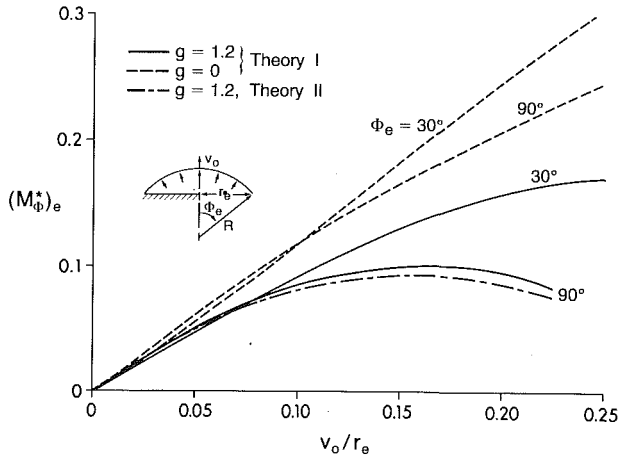


Fig. 6 Bending moment at clamped edges of spherical caps ($R/t = 20$, $\mu = 0$)

near the edge $x = x_e$ of a Mooney-Rivlin shell, where

$$G = 1 - \tilde{\gamma}^{(01)} N_\Phi^{(00)} / \lambda_\Phi^{(00)} Q^{(01)}$$
$$= 1 - g\alpha \lambda_\Phi^{(00)} N_\Phi^{(00)} / 2(1 + \mu \lambda_\theta^{(00)2}). \quad (38b)$$

These relations show that, at a given pressure for the case $g = 0$, membrane stretching narrows the secondary boundary layer. (The primary layer, which is contained in the interior solution, actually *widens* with increasing pressure (Reissner, 1959).) Increasing μ decreases the magnitude of this stretching and, therefore, widens this layer, but only a small amount for $\mu = 0.2$. The behavior of a shell with shear deformation allowed ($g > 0$) is somewhat more complicated. The shear term G in equation (38) indicates that shear deformation mediates the effect of membrane stretching; ξ decreases and the edge zone is wider than for $g = 0$.

The physical basis behind this behavior is illustrated for a cylinder in Fig. 5. When $g = 0$, the zero rotation condition dictates a negative M_Φ applied to the edge, with N_Φ having no effect on the moment. For $g > 0$ and $G > 0$, however, N_Φ contributes a moment that opposes the moment due to shear, thereby reducing M_Φ (see also equation $(9)_3$). Figure 6 clearly shows this effect; for large deflections, shear deformation dramatically reduces the edge moment. Furthermore, for $g = 1.2$, M_Φ actually peaks at a critical deflection and begins to decrease. Eventually, a point is reached at which $G = 0$ at the shell edge, where now $\xi' = M_\Phi = 0$ (equations (38) and (*B*4)), and the secondary boundary layer dissipates (see below).

The condition $G = 0$ defines a "membrane equivalent state" of the shell in which the bending moment, but not the transverse shear stress, vanishes, i.e., the moment due to N_Φ exactly balances that due to Q. In a pressurized cap, this condition actually is met first at a point in the shell interior. Then, as the load increases, this transition point ($G = \xi' = 0$), which divides the exponentially decreasing edge-zone solution ($G > 0$, ξ' real) from an oscillatory solution ($G < 0$, ξ' imaginary), moves toward the shell edge.

At very large deflections, $G < 0$ throughout the shell, and the edge-zone solution becomes completely oscillatory in x. Now, the edge effects propagate into the interior of the shell, which behaves as a "thick-walled membrane." The interior bending stresses, which have been $O(\beta^{-1})$ as given by the interior solution (Appendix B), now are O(1) and are given by what were edge-effect terms (exponential (01) terms in equation (36)). An in-depth study of this behavior is left to future research.

Finally, Fig. 6 also shows the effects of reference surface location on the bending moment. When shear deformation is included, the difference in bending moments given by Theory I and Theory II is quite small, even for very large deflections. Thus, the choice of reference surface definition does not appear to be important for moderately large strains.

Acknowledgment

This work was supported by the National Science Foundation under grant No. MSM-8611490.

References

Adkins, J. E., and Rivlin, R. S., 1952, "Large Elastic Deformations of Isotropic Materials IX. The Deformation of Thin Shells," *Philosophical Transactions of the Royal Society of London*, Vol. 244, Series A, pp. 505–531.

Brodland, G. W., 1986, "Nonlinear Deformation of Uniformly Loaded Circular Plates," *Solid Mechanics Archives*, Vol. 11, pp. 219–256.

Bromberg, E., and Stoker, J. J., 1945, "Non-linear Theory of Curved Elastic Sheets," *Quarterly of Applied Mathematics*, Vol. 3, pp. 246–265.

Gregory, R. D., and Wan, F. Y. M., 1985, "On Plate Theories and Saint-Venant's Principle," *International Journal of Solids and Structures*, Vol. 21, pp. 1005–1024.

John, F., 1965, "Estimates for the Derivatives of the Stresses in a Thin Shell and Interior Shell Equations," *Communications on Pure and Applied Mathematics*, Vol. 18, pp. 235–267.

Keppel, W. J., 1984, "Finite Axisymmetric Deformation of a Thin Shell of Revolution," Ph.D. Dissertation, University of Arizona.

Libai, A., and Simmonds, J. G., 1981, "Large-Strain Constitutive Laws for the Cylindrical Deformation of Shells," *International Journal of Non-Linear Mechanics*, Vol. 16, pp. 91–103.

Libai, A., and Simmonds, J. G., 1983, "Nonlinear Elastic Shell Theory," *Advances in Applied Mechanics*, Vol. 23, Hutchinson, J. W., and Wu, T. Y., eds., Academic Press, New York, pp. 271–371.

Meroueh, K. A., 1986, "On a Formulation of a Nonlinear Theory of Plates and Shells with Applications," *Computers and Structures*, Vol. 24, pp. 691–705.

Naghdi, P. M., 1972, "The Theory of Plates and Shells," *Encyclopedia of Physics*, 2nd ed., Vol. VIa/2, Flugge, S., ed., Springer-Verlag, New York, pp. 425–640.

Ranjan, G. V., and Steele, C. R., 1980, "Nonlinear Corrections for Edge Bending of Shells," ASME Journal of Applied Mechanics, Vol. 47, pp. 861–865.

Reissner, E., 1952, "Stress Strain Relations in the Theory of Thin Elastic Shells," *Journal of Mathematics and Physics*, Vol. 31, pp. 109–119.

Reissner, E., 1959, "The Edge Effect in Symmetric Bending of Shallow Shells of Revolution," *Communications on Pure and Applied Mathematics*, Vol. 12, pp. 385–398.

Reissner, E., 1969, "On Finite Symmetrical Deflections of Thin Shells of Revolution," ASME Journal of Applied Mechanics, Vol. 36, pp. 267–270.

Reissner, E., 1972, "On Finite Symmetrical Strain in Thin Shells of Revolution," ASME Journal of Applied Mechanics, Vol. 39, pp. 1137–1138.

Reissner, E., 1974, "Linear and Nonlinear Theory of Shells," *Thin Shell Structures*, Fung, Y. C., and Sechler, E. E., eds., Prentice-Hall, Englewood Cliffs, New Jersey.

Simmonds, J. G., 1985, "The Strain Energy Density of Rubber-Like Shells," *International Journal of Solids and Structures*, Vol. 21, pp. 67–77.

Simmonds, J. G., 1986, "The Strain Energy Density of Rubber-Like Shells of

Revolution Undergoing Torsionless, Axisymmetric Deformation (Axishells)," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 593–596.

Steele, C. R., 1980, "Asymptotic Solutions Without Special Functions for Steep and Shallow Shells," *Mechanics Today*, Vol. 5, Nemat-Nasser, S., ed., Pergamon, New York, pp. 483–494,

Taber, L. A., 1985, "On Approximate Large Strain Relations for a Shell of Revolution," *International Journal of Non-Linear Mechanics*, Vol. 20, pp. 27–39.

Taber, L. A., 1987a, "Asymptotic Expansions for Large Elastic Strain of a Circular Plate," *International Journal of Solids and Structures*, in press.

Taber, L. A., 1987b, "On Boundary Layers in a Pressurized Mooney Cylinder," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 280–286.

APPENDIX A

For a Mooney-Rivlin material with W given by equation (37), the terms required for the approximation for w given by equations (29) and (30) are

$$\overset{0}{(W^*)}_{G_\Gamma=0}=\lambda_\Phi^2+\lambda_\theta^2+(\lambda_\Phi\lambda_\theta)^{-2}-3 +\mu(\lambda_\Phi^2\lambda_\theta^2+\lambda_\Phi^{-2}+\lambda_\theta^{-2}-3)$$

$$\overset{0}{(W^*_{(1,0)})}_{G_\Gamma=0}=(1-\lambda_\Phi^{-4}\lambda_\theta^{-2})(1+\mu\lambda_\theta^2)$$

$$\overset{0}{(W^*_{(0,1)})}_{G_\Gamma=0}=(1-\lambda_\Phi^{-2}\lambda_\theta^{-4})(1+\mu\lambda_\Phi^2)$$

$$\overset{0}{(W^*_{(1,1)})}_{G_\Gamma=0}=(1+\mu\lambda_\Phi^4\lambda_\theta^4)/(\lambda_\Phi\lambda_\theta)^4$$

$$\overset{0}{(W^*_{(2,0)})}_{G_\Gamma=0}=2(1+\mu\lambda_\theta^2)/(\lambda_\Phi^6\lambda_\theta^2)$$

$$\overset{0}{(W^*_{(0,2)})}_{G_\Gamma=0}=2(1+\mu\lambda_\Phi^2)/(\lambda_\Phi^2\lambda_\theta^6)$$

$$\overset{0}{(W^*_{;G_\Gamma})}_{G_\Gamma=0}=(1+\mu\lambda_\theta^2)/\lambda_\Phi^2. \tag{A1}$$

APPENDIX B

The terms for a first approximation solution (36) follow.

Interior Solution. The first-order terms in β provide the system of equations

$$\kappa_\Phi^{(00)}=\kappa_\theta^{(00)}=\tilde\gamma^{(00)}=M_\Phi^{(00)}=M_\theta^{(00)}=0$$

$$r^{(00)\prime}=\lambda_\Phi^{(00)}c^{(00)},\quad y^{(00)\prime}=\lambda_\Phi^{(00)}s^{(00)},\quad \lambda_\theta^{(00)}=r^{(00)}/r_0,$$

$$N_\Phi^{(00)}=\alpha^{-1}[\partial w_0/\partial\lambda_\Phi]_0=H^{(00)}c^{(00)}+V^{(00)}s^{(00)},$$

$$N_\theta^{(00)}=\alpha^{-1}[\partial w_0/\partial\lambda_\theta]_0,$$

$$Q^{(00)}=H^{(00)}s^{(00)}-V^{(00)}c^{(00)}=0,\quad V^{(00)}=r^{(00)2}/2r_0,$$

$$(r_0H^{(00)})'-N_\theta^{(00)}+r_0\lambda_\Phi^{(00)}\lambda_\theta^{(00)}s^{(00)}=0,$$

$$s^{(00)}=\eta^{(00)}s_0+\phi^{(00)}c_0,\quad c^{(00)}=\eta^{(00)}c_0-\phi^{(00)}s_0,$$

$$\eta^{(00)}=(1-\phi^{(00)2})^{1/2} \tag{B1}$$

where w_0 is given by equation (30) and

$$[\ ]_0\equiv[\ ]_{(\lambda_\Phi^{(00)},\lambda_\theta^{(00)})}. \tag{B2}$$

These are the governing relations of nonlinear membrane theory (Adkins and Rivlin, 1952) and contain the primary boundary layer (Taber, 1987b). The terms involving the general strain energy density function w were obtained by observation of the solutions for specific forms of w.

The second-order β terms contribute

$$M_\Phi^{(10)}=\alpha^{-1}[2w_2\kappa_\Phi^{(10)}+w_3\kappa_\theta^{(10)}]_0,$$

$$M_\theta^{(10)}=\alpha^{-1}[w_3\kappa_\Phi^{(10)}+2w_4\kappa_\theta^{(10)}]_0$$

$$\kappa_\Phi^{(10)}=\phi^{(00)\prime}/\eta^{(00)}+(1-\lambda_\Phi^{(00)2}\lambda_\theta^{(00)}\Phi_0'),$$

$$\kappa_\theta^{(10)}=(s^{(00)}-\lambda_\Phi^{(00)}\lambda_\theta^{(00)2}s_0)/r_0$$

$$Q^{(10)}=\tilde\gamma^{(10)}=0. \tag{B3}$$

Edge-Zone Solution. The first-order $e^{\beta\xi}$ terms give

$$r^{(01)}=y^{(01)}=\lambda_\Phi^{(01)}=\lambda_\theta^{(01)}=\kappa_\theta^{(01)}=N_\Phi^{(01)}=N_\theta^{(01)}=V^{(01)}=H^{(01)}=0$$

$$\kappa_\Phi^{(01)}=\xi'\phi^{(01)}/\eta^{(00)},\quad \tilde\gamma^{(01)}=g\alpha Q^{(01)}/[2w_1]_0,$$

$$Q^{(01)}=N_\Phi^{(00)}\phi^{(01)}/\eta^{(00)},\quad M_\Phi^{(01)}=2\alpha^{-1}[w_2]_0\kappa_\Phi^{(01)},$$

$$M_\theta^{(01)}=\alpha^{-1}[w_3]_0\kappa_\Phi^{(01)},\quad \eta^{(01)}=-\phi^{(00)}\phi^{(01)}/\eta^{(00)},$$

$$s^{(01)}=c^{(00)}\phi^{(01)}/\eta^{(00)},\quad c^{(01)}=-s^{(00)}\phi^{(01)}/\eta^{(00)} \tag{B4}$$

along with the decay function

$$(\xi')^2=\alpha G\lambda_\Phi^{(00)}N_\Phi^{(00)}/[2w_2]_0 \tag{B5}$$

where

$$G=1-\tilde\gamma^{(01)}N_\Phi^{(00)}/\lambda_\Phi^{(00)}Q^{(01)}. \tag{B6}$$

With $\phi^{(01)}$ determined by the boundary conditions, these terms provide the bending component of the secondary boundary layer. The stretching component is given by the $e^{2\beta\xi}$ terms

$$r^{(02)}=y^{(02)}=\lambda_\theta^{(02)}=\kappa_\Phi^{(02)}=\kappa_\theta^{(02)}=\tilde\gamma^{(02)}=0$$

$$M_\Phi{}^{(02)}=M_\theta^{(02)}=Q^{(02)}=V^{(02)}=H^{(02)}=0$$

$$\lambda_\Phi^{(02)}=-\left[\frac{\partial w_2}{\partial\lambda_\Phi}+\frac{w_2}{\lambda_\Phi G}\left(1+\frac{(1-G)\lambda_\Phi}{w_1}\frac{\partial w_1}{\partial\lambda_\Phi}\right)\right]_0 \times\frac{\kappa_\Phi^{(01)2}}{[\partial^2w_0/\partial\lambda_\Phi^2]_0}$$

$$N_\Phi{}^{(02)}=-(N_\Phi^{(00)}/2\eta^{(00)2})\phi^{(01)2}$$

$$N_\theta^{(02)}=\alpha^{-1}\left[\lambda_\Phi^{(02)}\frac{\partial^2w_0}{\partial\lambda_\Phi\partial\lambda_\theta}+\frac{\tilde\gamma^{(01)2}}{g}\frac{\partial w_1}{\partial\lambda_\theta}+\kappa_\Phi^{(01)2}\frac{\partial w_2}{\partial\lambda_\theta}\right]_0$$

$$\phi^{(02)}=-(\phi^{(00)}/2\eta^{(00)2})\phi^{(01)2},\quad \eta^{(02)}=-\phi^{(01)2}/2\eta^{(00)}$$

$$s^{(02)}=-(s^{(00)}/2\eta^{(00)2})\phi^{(01)2},$$

$$c^{(02)}=-(c^{(00)}/2\eta^{(00)2})\phi^{(01)2}. \tag{B7}$$

D. J. Segalman[1]
Earth Simulation Research,
Plano, TX 75075
Mem. ASME

Calculation of Damping Matrices for Linearly Viscoelastic Structures

A technique is presented for the systematic calculation of "damping" and "stiffness" matrices to represent the linearly viscoelastic properties of structures. The technique generates explicit expressions for these matrices in terms of the measurable viscoelastic properties of the components of the structure.

Introduction

Presented here is a systematic approach for the calculation of damping and stiffness matrices for the calculation of the small amplitude motion of linearly viscoelastic structures. The problem is described as follows.

The governing equations for the small amplitude dynamics of viscoelastic structures, even linearly viscoelastic structures, are *integro-differential* equations in time. Not only are such equations difficult to solve analytically, but their numerical solution is computationally very expensive. On the other hand, the increasing use of plastics in advanced structures makes consideration of the viscoelastic properties of components increasingly important in structural modeling.

It is because of the prohibitive difficulty of solving the integro-differential equations of viscoelastic dynamics that it is common to approximate those equations by simpler ones which are simply second order differential equations in time. The viscoelastic nature of the structure is accounted for by the inclusion of a "damping" term.

Only for very restricted subsets of linear viscoelasticity have systematic methods been developed for the calculation of damping matrices that could be employed in corresponding second-order systems (Biot, 1955; Golla and Hughes, 1985; Bagley and Torvik, 1983). However, there has not been a general and systematic method for the calculation of the damping matrices from arbitrary linear viscoelastic models. One such method is suggested in this paper.

The approach taken is to match the perturbation solution for a "slightly viscoelastic" structure to the perturbation solution for a corresponding "slightly damped" structure. Requiring the two perturbation results to agree results in expressions for the damping (and stiffness) matrices in terms of the viscoelastic properties.

Discretization of Linearly Viscoelastic Structures

A linearly viscoelastic structure is one that is composed of linearly elastic and linearly viscoelastic materials. The stress response at a particle, X, of a linearly viscoelastic material is

$$\sigma_{ij}(t,X) = \int_0^\infty \zeta_{ijmn}(\tau,X)\dot{e}_{mn}[(t-\tau),X]d\tau \tag{1}$$

where $\sigma_{ij}(t, X)$ is the ij component of the Cauchy stress at time t and particle X; $e_{ij}(t, X)$ is the ij component of the Cauchy strain at time t and particle X; and $\zeta_{ijmn}(\ ,X)$ is the $ijmn$ component of the relaxation tensor at particle X; (It is a material property.) Above and in the rest of this paper, summation occurs over repeated indices.

Gurtin and Sternberg (1962) enumerate conditions on both the relaxation tensor and the strain history that guarantee the convergence of the integral in equation (1). The constraint on strain history is that the strain be continuous over the interval $[-\infty, t]$.

Discretization

The development in this section should be familiar to those who have worked through standard finite element derivations. It is presented here as the most concise method of introducing and defining quantities which are used in later portions of this paper.

A structure B of particles X is considered. It is assumed that the configuration at any time, t can be specified satisfactorily by a linear combination of basis functions, $\{h_{i|r}(X)\}$, and a corresponding set of generalized displacements, $\{w_{|r}\}$. The physical displacements, $y_i(t, X)$, of particles $X \subset B$ are linearly determined by the generalized displacements:

$$y_i(t,X) = h_{i|r}(X)w_{|r}(t) \tag{2}$$

where subscripts $|r$ refer to the generalized coordinate and again, there is summation over repeated indices.

The strain field is obtained from the above displacement field and is also expressed in terms of the generalized displacements:

$$e_{ij}(t,X) = k_{ij|r}(X)w_r(t) \tag{3}$$

To assure convergence of all relevant integrals, only strain and displacement fields are considered here which are bounded (finite) throughout the body.

In terms of the generalized displacements, the Lagrange equations of motion reduce to:

[1]Currently Member of Technical Staff, Sandia National Laboratories, Albuquerque, NM 87185.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13-18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 14, 1986; final revision, November 22, 1986.
Paper No. 87-WA/APM-6.

$$P_{|rs}\ddot{w}_{|s}(t) = Q_{|r}(t) \tag{4}$$

for each r, where $Q_{|r}$ is a generalized force associated with the rth degree of freedom defined by the virtual work equation

$$\delta W = Q_{|r}\delta w_{|r} \tag{5}$$

$$P_{|rs} = \int_B \rho(X) h_{i|r}(X) h_{i|s}(X) dX; \tag{6}$$

and $\rho(X)$ is mass density at X so that $\rho(X)dX$ has units of mass.

The virtual work done (internally) through the stresses within the material is

$$\delta W^I = -\int_B \sigma_{ij}(t,X)\delta e_{ij}(t,X)dX \tag{7}$$

$$= \left[\int_0^\infty \Gamma_{|rs}(\tau)\dot{w}_{|r}(t-\tau)d\tau\right]\delta w_{|s} \tag{8}$$

where

$$\Gamma_{|rs}(\tau) = \int_B \zeta_{ijmn}(\tau,X) k_{ij|r}(X) k_{mn|s}(X) dX \tag{9}$$

(Note that the above virtual work consists of both recoverable and unrecoverable parts.)

The generalized force originating from the stress response of the material in the structure and associated with this virtual work term is

$$Q^I_{|r}(t) = -\int_0^\infty \Gamma_{|rs}(\tau)\dot{w}_{|s}(t-\tau)d\tau \tag{10}$$

Substituting into the Lagrange equations of motion,

$$P_{|rs}\ddot{w}_{|s}(t) + \int_0^\infty \Gamma_{|rs}(\tau)\dot{w}_{|s}(t-\tau)d\tau = Q^E_{|r}(t) \tag{11}$$

where the right-hand term represents all other applied forces, including body forces and tractions applied at boundaries.

The convergence of the infinite integral in the above equation is induced from the convergence of the integral in equation (1) and the boundedness of the strain functions, $k_{ij|r}$. The restriction on admissible strain histories used in equation (1) induces a similar restriction in the generalized degrees of freedom; the above equation is restricted to histories $w_{|r}(t)$ over which $w_{|r}(t)$ is continuous over its history.

It should be observed that the mass matrix P ($= [P_{|rs}]$) is determined entirely by the mapping from generalized displacements to particle displacements and by the distribution of mass in the structure. The matrix of relaxation functions $\Gamma(\tau)$ ($= [\Gamma_{|rs}(\tau)]$) is entirely determined by the mapping from the generalized displacements to the strain field in the structure and by the viscoelastic properties of the materials in the structure. The stiffness matrix Γ^e ($= [\Gamma^e_{|rs}]$) which would be calculated from static elastic properties of the materials in the structure is $\Gamma(\infty)$.

The Problem of Free Vibration

$$P_{|rs}\ddot{w}_{|s}(t) + \int_0^\infty \Gamma_{|rs}(\tau)\dot{w}_{|s}(t-\tau)d\tau = 0 \tag{12}$$

We look for solutions $w_{|r}(t) = \text{Re}\,[A_{|r}e^{-\alpha t}]$ where in general, $A_{|r}$ and α are complex.

$$e^{-\alpha t}\alpha^2 P_{|rs}A_{|s} + e^{-\alpha t}A_{|s}\int_0^\infty \Gamma_{|rs}(\tau)(-\alpha)e^{\alpha\tau}d\tau = 0 \tag{13}$$

It is now convenient to define the complex stiffness of a structure by

$$\Gamma^*_{|rs}(\omega) = \int_0^\infty \iota\,\omega\,\Gamma_{|rs}(\tau)e^{-\iota\omega\tau}d\tau. \tag{14}$$

(This corresponds to the definition for the complex modulus of a viscoelastic material.)

Then
$$[\alpha^2 P_{|rs} + \Gamma^*_{|rs}(\iota\alpha)]A_{|s} = 0 \tag{15}$$

In matrix form:
$$[\alpha^2 P + \Gamma^*(\iota\alpha)]A = 0 \tag{16}$$

Because of the dependence of Γ^* on α, this is a nonlinear eigenvalue problem.

Perturbation for "Small Viscoelasticitiy"

A second order approximation for the equations of motion for a viscoelastic structure is derived for the special case where the elasticity of the viscoelastic structure dominates the transient response:

$$\Gamma_{|rs}(\tau) = \Gamma^e_{|rs} + \Delta\Gamma_{|rs}(t). \tag{17}$$

The assumption of small viscoelasticity is that

$$\Delta\Gamma_{|rs}(t) \ll \Gamma^e_{|rs} \tag{18}$$

for all t.

The standard ϵ notation will be used to emphasize this assumption:

$$\Gamma(\tau) = \Gamma^e + \epsilon\Delta\Gamma(t). \tag{19}$$

where ϵ is assumed to be much less than 1. Here and for the rest of this paper, matrix notation is employed.

Also defined is

$$\Delta\Gamma^*(\omega) = \int_0^\infty \iota\omega\Delta\Gamma(\tau)e^{-\iota\omega\tau}d\tau \tag{20}$$

so that

$$\Gamma^*(\omega) = \Gamma^e + \epsilon\Delta\Gamma^*(\omega). \tag{21}$$

Expressed in terms of the above notation, the equation of motion for the case of free vibration is an eigenproblem:

$$\left[(\alpha^n)^2 P + \Gamma^e + \epsilon\Delta\Gamma^*(\alpha^n)\right]A^n = 0 \tag{22}$$

(no sum on the n's) where (α^n, A^n) is the nth eigensolution. Here and in what follows, a superscript identifies distinct eigensolutions. The above is a nonlinear eigenproblem because of the dependence of $\Delta\Gamma^*$ on α.

Note that since the real part as well as the imaginary part of $\Delta\Gamma^*$ may be nonzero, the viscoelasticity of the structure will not only add damping, but will also alter the apparent stiffness of the structure. In acoustics, this stiffening is referred to as "dispersion." This stiffening is accommodated in the second order approximation later in this paper through modification of the effective stiffness matrix.

A perturbation solution to the above viscoelastic eigenproblem is sought in the following form:

$$\alpha^n = -\iota\omega^n + \epsilon\beta^n \tag{23}$$

$$A^n = x^n + \epsilon v^n \tag{24}$$

where ω^n and x^n are on the order of Γ^e.

Substitution of the assumed solution into the viscoelastic eigen equation and retaining only terms up to first order in ϵ yields the following matrix equations:

$$[-(\omega^m)^2 P + \Gamma^e]x^m = 0 \tag{25}$$

for the zeroth order terms and

$$[-(\omega^n)^2 P + \Gamma^e]v^n = [2\iota\omega^n\beta^n P - \Delta\Gamma^*(\omega^n)]x^n \tag{26}$$

(no sum on n) for first order terms.

It will be useful to observe that, by construction, the mass matrix P is symmetric. The elastic stiffness matrix Γ^e is symmetric by virtue of the Maxwell reciprocal theorem.

The first of these matrix equations is contracted with v^n and the second is contracted with x^m. Then terms involving the

mass matrix P are resolved out to generate the following equation:

$$[1-(\omega^n/\omega^m)^2](x^m)^T\Gamma^e v^n \tag{27}$$

$$=[(2\iota\beta^n/\omega^n)(x^m)^T\Gamma^e x^n-(x^m)^T\Delta\Gamma^* x^n \tag{28}$$

Assuming only simple eigensolutions for the elastic problem, there are only two cases which must be considered:

(1) $n=m$: in which case

$$(x^n)^T\Delta\Gamma^* x^n=\frac{2\iota\beta^n}{\omega^n}(x^n)^T\Gamma^e x^n \tag{29}$$

where there is no summation over n.

(2) $n\neq m$: in which case

$$(x^m)^T\Delta\Gamma^*(\omega^n)x^n=-[1-(\omega^n/\omega^m)^2](x^m)^T\Gamma^e v^n \tag{30}$$

Approximation by a Second Order System

The equations of motion are next approximated by the following second order system:

$$P\ddot{w}+\epsilon\Delta C\dot{w}+[\Gamma^e+\epsilon\Delta\Gamma^e]w=Q \tag{31}$$

where the matrices P and Γ^e are as defined above, and ΔC and $\Delta\Gamma^e$ are selected to give the same perturbation solutions as derived in the previous section for free vibration.

$$P\ddot{w}+\epsilon\Delta C\dot{w}+(\Gamma^e+\epsilon\Delta\Gamma^e)w=0 \tag{32}$$

Again, a solution $w(t)=\mathrm{Re}\,[Ae^{-\alpha t}]$ is assumed:

$$[\alpha^2 P-\alpha\epsilon\Delta C+(\Gamma^e+\epsilon\Delta\Gamma^e)]A=0. \tag{33}$$

Again, the solution is sought as a perturbation to the elastic problem:

$$\alpha^n=-\iota\omega^n+\epsilon\beta^n \tag{34}$$

and

$$A^n=x^n+\epsilon v^n \tag{35}$$

The following is derived employing manipulations similar to the previous section:

$$[1-(\omega^n/\omega^m)^2](x^m)^T\Gamma^e v^n=(2\iota\beta^n/\omega^n)(x^m)^T\Gamma^e x^n$$
$$-(x^m)^T[\Delta\Gamma^e+\iota\omega^n\Delta C]x^n \tag{36}$$

from which can be derived:

$$(x^n)^T[\Delta\Gamma^e+\iota\omega^n\Delta C]x^n=\frac{2\iota\beta^n}{\omega^n}(x^n)^T\Gamma^e x^n \tag{37}$$

for all n, and

$$(x^m)^T[\Delta\Gamma^e+\iota\omega^n\Delta C]x^n$$
$$=-[1-(\omega^n/\omega^m)^2](x^m)^T\Gamma^e v^n \tag{38}$$

for $n\neq m$. Requiring that this perturbation be identical to the previous one results in the following equation:

$$[\Delta\Gamma^e+\iota\omega^n\Delta C]x^n=\Delta\Gamma^*(\omega^n)x^n \tag{39}$$

for all n. Explicit expressions for the ΔC and $\Delta\Gamma^e$ can be derived with the help of a set of basis vectors $\{z^n\}$ orthogonal to the eigenvectors relative to the stiffness matrix:

$$z^n=\frac{\Gamma^e x^n}{(x^n)^T\Gamma^e x^n} \tag{40}$$

(no sum on the n's) for each n. Postmultiplying both sides by $(z^n)^T$ and summing on n,

$$\Delta\Gamma^e\sum_n x^n(z^n)^T=Re\sum_n\Delta\Gamma^*(\omega^n)x^n(z^n)^T \tag{41}$$

Since $\sum_n x^n(z^n)^T$ is the identity matrix,

$$\Delta\Gamma^e=\mathrm{Re}\Big[\sum_n\Delta\Gamma^*(\omega^n)x^n(z^n)^T\Big] \tag{42}$$

Similarly,

$$\Delta C=\sum_n\mathrm{Im}\Big[\frac{1}{\omega^n}\Delta\Gamma^*(\omega^n)x^n(z^n)^T\Big] \tag{43}$$

Some Computational Simplifications

Examination of the above two equations does not provide any indication that the damping and stiffness matrices are symmetric. Indeed, there is no reason to expect damping and stiffness matrices that derive from linear viscoelasticity to be symmetric, since they do not originate with dissipation or strain energy potentials. In addition, the damping and stiffness matrices derived above can not be expected to preserve the normal modes of the perfectly elastic problem. Though physically reasonable, the nonsymmetry and mode coupling of the above matrices are usually undesirable features in numerical calculation. It is for that reason that the following further approximations are introduced so that more amenable matrices are achieved.

Comments on Application to Finite Element Analysis

On Traditional Methods of Selecting Damping Methods. Traditional approaches to treating damping—in particular viscous damping, Rayleigh damping, and modal damping—have been employed both for computational ease and because of the absence of a rational method for incorporating known viscoelastic properties of materials. The method presented here is one such rational method.

The method presented here is more general and more direct than those presented in Biot (1955), Golla and Hughes (1985), and Bagley and Torvik (1983) in that it is not restricted to any specific subset of linear viscoelasticity. In fact, the method presented here simply requires complex moduli of the materials in the appropriate range of frequencies.

On the Cost of Assembling These Damping Matricies. In the context of finite element analysis, this method would involve the calculation of complex stiffness matrices Γ^* at each resonant frequency using code almost identical to the generation of the elastic stiffness matrix. Also calculated at each resonant frequency must be the dual vectors z^n. Both of the above operations has its own numerical cost. However, the above operations need not be performed for all eigenfrequencies; the process may be restricted to just those frequencies or ranges of frequencies of interest.

On the Symmetry of Damping and Stiffness Matrices Derived in the Above Manner. There is no reason to expect coefficient matrices derived from linear viscoelasticity to be symmetric except in one special case: where the relaxation tensors, $\zeta_{ijmn}(\tau, X)$, everywhere in the structure consist only of heavy side functions at zero (elastic components) and delta functions at zero (Newtonian viscosity). In this special case, where strain energy and dissipation potentials exist, the method presented here does indeed generate symmetric matrices.

On Preservation of the Normal Modes of the Elastic Problem. In general, the damping and stiffness matrices generated in the manner presented here do not preserve the eigenmodes of the elastic problem. However, these matrices can be modified slightly in a manner that both results in symmetric matrices and preserves the eigenmodes of the elastic problem.

The revised matrices are obtained by premultiplying each component, $\Delta\Gamma^*(\omega^n)\,x^n\,(z^n)^T$, of the complex matrix from which the damping matrix is derived, by appropriate terms to make it orthogonal to all but the corresponding eigenmode:

$$\Delta\Gamma^e = \mathrm{Re}\left[\sum_n z^n (x^n)^T \Delta\Gamma^*(\omega^n) x^n (z^n)^T\right] \tag{44}$$

Similarly,

$$\Delta C = \sum_n \mathrm{Im}\left[\frac{1}{\omega^n} z^n (x^n)^T \Delta\Gamma^*(\omega^n) x^n (z^n)^T\right. \tag{45}$$

Each of the matrices is symmetric and preserves the desired eigenmode.

On the Assumption of "Small Viscoelasticity." This assumption was employed in connecting viscoelasticity to damping, and the derivation suggests that the two can only be connected for general deformation histories in the case where the assumption is good. The author is developing numerical experiments to provide a better notion of just how "small" the viscoelasticity of a structure need be in order for it to be modelled adequately by a damped system.

References

Bagley, R. L., and Torvik, P. J., 1983, "Fractional Calculus—A Different Approach to Finite Element Analysis of Viscoelastically Damped Structures," *AIAA Journal*, Vol. 21, No. 5, pp. 741–748.

Biot, M. A., 1955, "Variational Principles in Irreversible Thermodynamics with Application to Viscoelasticity," *Physical Review*, Vol. 97, No. 6, pp. 1463–1469.

Golla, D. F., and Hughes, P. C., 1985, "Dynamics of Viscoelastic Structures—A Time-Domain, Finite Element Formulation," ASME JOURNAL OF APPLIED MECANICS, Vol. 52, pp. 897–906.

Gurtin, M. E., and Sternberg, E., 1962, "On the Linear Theory of Viscoelasticity," *Arch. Ration. Mech. Anal.*, Vol. 11, pp. 291–356.

An Improved Shear-Deformation Theory for Moderately Thick Multilayered Anisotropic Shells and Plates

M. Di Sciuva
Dipartimento di Ingegneria Aeronautica e Spaziale, Politecnico di Torino, Turin, Italy 10129

The general linear equations governing the motion of moderately thick multilayered anisotropic shells are derived by making use of the principle of virtual work in conjunction with an a priori assumed displacement field. The assumed displacement field is piecewise linear in the u *and* v *components and fulfills the static and geometric continuity conditions between the contiguous layers; furthermore, it takes into account the distortion of the deformed normal. Shear and rotatory inertia terms have also been considered in the formulation. Particularization of the resulting equations to the flat multilayered anisotropic plates is straightforward; thus, only the final expressions are given. The proposed approach gives, as particular cases, the linear equations of motion of the classical shells theory based on the Kirchhoff-Love kinematic hypothesis and those of the shear deformation theory for which it is assumed that the deformed normal do not distort.*

Introduction

An increasing number of structural designs, especially in the aerospace, automobile, and petrochemical industries, are extensively utilizing fiber composite laminated plates and shells as structural elements. Because the solution of the three-dimensional linear problem with general boundary conditions involves considerable mathematical difficulties, in recent years some approximate bidimensional linear theories for multilayered plates and shells have been developed by making use of the axiomatic approach. This approach generally utilizes the principle of virtual work in conjunction with an assumed displacement field. An integration with respect to the thickness coordinate supply the governing differential equations and consistent boundary conditions in terms of unknown generalized coordinates which are independent of the thickness coordinate.

At the present time many of the existing methods of analysis for multilayered anisotropic plates and shells are direct extensions of those developed earlier for homogeneous isotropic and orthotropic plates and shells. In fact, many approaches utilize a displacement field which do not account for the equilibrium requirements at the interfaces.

In the Classical Lamination Theory (C.L.T.), the well-known Kirchhoff-Love kinematic hypothesis is assumed to be verified (Reissner and Stavsky, 1961; Dong et al., 1962; Ambartsumyan, 1964). A theoretical unification of the thin shell theories commonly used (Donnel's, Love's, Sanders's, and Flugge's theories) as well as a numerical comparison has been presented by Soldatos (1984).

The range of applicabilitiy of the C.L.T. solution has been well established for laminated flat plates by Pagano (1969, 1970). To the best of the author's knowledge no analogous solutions exist for curved plates and closed shells.

These analyses have indicated that a theory which accounts for the transverse shear deformation effects would be adequate to predict the gross behavior of the laminate.

A Mindlin-type first-order transverse shear deformation theory (S.D.T.) has been first developed by Whitney and Pagano (1970) for multilayered anisotropic plates and by Dong and Tso (1972) for multilayered anisotropic shells.

Both of the previous approaches considered all layers as one equivalent single anisotropic layer; thus, these approaches are inadequate to model the warpage of cross sections, that is, the distortion of the deformed normal due to transverse shear stresses. Furthermore, the assumption of nondeformable normal results in incompatible shearing stresses between every two adjacent layers. Also, the latter approach requires the introduction of an arbitrary shear correction factor which is dependent on the lamination parameters for obtaining accurate results (Whitney and Pagano, 1970).

The exact analyses performed by Pagano (1969, 1970) on composite flat plates have indicated that the distortion of the deformed normal is dependent not only on the laminate thickness, but also on the orientation and degree of orthotropy of the individual layers. Therefore, the hypothesis of nondeformable normals, while acceptable for isotropic plates and shells is often quite unacceptable for multilayered anisotropic plates and shells with very large ratio of Young's

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 23, 1985; final revision, May 30, 1986.

modulus to shear modulus, even if they are relatively thin. Thus, a transverse shear deformation theory which also accounts for the distortion of the deformed normal would be quite accurate in predicting the elastic linear behavior (deflection, thickness distribution of the in-plane displacements, natural frequencies, etc.) of multilayered anisotropic plates and shells.

Higher order theories, in which a displacement field of polynomial form a degree greater than one is assumed, have been developed (Whitney and Sun, 1973, 1974; Librescu, 1975; Lo et al., 1977; Bhimaraddi, 1985) for the purpose of removing the inaccuracies in the classical lamination and first-order shear deformation theories. From an engineering point of view, such approaches have little equal because of the difficulties in obtaining solutions to the system of governing partial differential equations and in prescribing boundary conditions. In this context mention should be made of the approach proposed by Reddy which has developed a higher-order but simple shear deformation theory of laminated plates (Reddy, 1984) and shells (Reddy, 1985). The developed theory is simple in the sense that it contains the same dependent unknowns as in the first-order shear deformation theory. The u and v displacements are expanded as cubic functions of the thickness coordinate (as in Lo et al., 1977), and the transverse displacement is assumed to be constant. However, in contrast to Lo et al., the number of generalized coordinates is reduced to five, setting to zero the transverse shearing stresses at the top and bottom surfaces.

Obviously, these higher-order shear deformation theories, as well as the first-order theory, will not fulfill the continuity conditions for the transverse shearing stresses at the interfaces.

To develop a theory for composite laminates which allows the contact conditions for the displacements and the transferse shearing stresses at the interfaces to be satisfied simultaneously, the following two axiomatic approaches have been proposed and developed.

The first approach has been utilized by Sun and Whitney (1973) and Srinivas (1973) to formulate a refined theory for multilayered plates, and by Zukas and Vinson (1971) and Waltz and Vinson (1976) for multilayered shells. Firstly, distinct transverse shear deformations are allowed to exist within each layer; thus, initially, for each layer it is assumed that the kinematic hypothesis of the first-order shear deformation theory is verified. Secondly, these deformations are constrained in order that the shearing stresses be continuous at the interfaces of the layers.

The second approach has been originated by Ambartsumyan (1964, 1969). Following this approach, the distributions of the transverse shearing stresses in each layer are assumed to be known. The fulfillment of both the continuity conditions between the adjacent layers and the boundary conditions on the bounding surfaces allows to obtain some unknown parameters. The approach has been utilized by Ambartsumyan (1969) to formulate a refined plate theory for symmetric cross-ply laminates and extended by Whitney (1969) to the symmetric laminates in which the material axes of each layer have arbitrary orientation with respect to the plate axes. Extensions to the laminated cylindrical shells consisting of orthotropic layers have been suggested by Hsu and Wang (1970), and to symmetrically layered general orthotropic shells by Rath and Das (1973).

Although the two latter approaches are very accurate, they are quite cumbersome and computationally more demanding, especially in the case of multiple layers, because the number of equations in the final system increases with increasing the number of layers. In fact, all the numerical results refer to two layered shells. Further, it is very difficult to utilize these approaches for constructing plate and shell finite elements via the finite element displacement method.

The approach proposed in this paper utilizes a displacement field which fulfills a priori the static and geometric continuity conditions between contiguous layers. It is worth mentioning that the number of partial differential equations in the resulting system is independent of the number of layers; in addition, the order of the system is the same as in the first-order shear deformation theory.

The reduction of the three dimensional problem to the bidimensional one is accomplished by assuming a displacement field which allows piecewise linear variation of the in-plane displacements u and v, and the constant value of the transverse displacement w through the thickness of the laminate. Thus, the boundary conditions on the external bounding surfaces are not fulfilled, as well as in the first-order transverse shear deformation theories.

In a recent series of papers, the writer has employed this approach to obtain a refined shear deformation theory governing the linear elastostatic behavior of multilayered orthotropic (Di Sciuva, 1984a) and anisotropic (Di Sciuva, 1984b) plates. The approach has been extended also to the formulation of the geometrically nonlinear equations of motion of multilayered orthotropic plates (Di Sciuva, 1986).

Numerical tests carried out on the cylindrical bending of a three-layered symmetric cross-ply (Di Sciuva, 1984a) and angle-ply (Di Sciuva, 1984b) strip and on the bending, vibration, and buckling of a three-layered, symmetric cross-ply, square plate simply supported on all edges (Di Sciuva, 1986) prove that the proposed approach does work. Thus, by making use of this approach, in the following a refined linear theory governing the elastodynamic behavior of moderately thick anisotropic shells is developed.

The chief advantage of the assumed displacement field rests on its capability to model the distortion of the deformed normal and to satisfy the contact conditions ad initio, without increasing the number and order of the partial differential equations with respect to the first-order transverse shear deformation theory. Furthermore, it is feasible to employ this formulation for constructing plate and shell finite elements via the finite element displacement method (Di Sciuva, 1985a–b).

The present paper is structured as follows. After discussing the general linear strain-displacement relations in the Section 2, we give the complete expression for the displacement field in Section 3. By making use of the principle of virtual work, the linearized differential equations of motion and related boundary conditions for a general shell are derived in Section 4. In Sections 5 and 6 the previous results are particularized to the shells of revolution and to the flat plates, respectively.

A sequel to the present part containing numerical results and comparisons with other results and solutions for curved plates and closed shells of revolution is in preparation.

1 Geometrical Preliminaries

Consider the space surrounding an arbitrary surface S, hereafter designated the shell reference surface, which is defined by two curvilinear orthogonal coordinates (α, β) coinciding with its lines of principal curvature (Gaussian curvilinear coordinates). Let $\vec{t}_\alpha$ and $\vec{t}_\beta$ be the unit vectors in the directions of α and β, respectively,

$$\vec{t}_\alpha = A^{-1}\partial_\alpha \vec{r} \quad \vec{t}_\beta = B^{-1}\partial_\beta \vec{r}$$

where $\vec{r} = \vec{r}(\alpha, \beta)$ is the position vector of a point on the reference surface;

$$A^2 = \partial_\alpha \vec{r} \cdot \partial_\alpha \vec{r}; \quad B^2 = \partial_\beta \vec{r} \cdot \partial_\beta \vec{r}$$

are the coefficients of the first fundamental form of the shell reference surfaces (surface metric coefficients). The symbol ∂_α stands for partial derivative with respect to α. The unit vector perpendicular to S is denoted by $\vec{n}$, which is chosen so that $\vec{t}_\alpha$, $\vec{t}_\beta$, and $\vec{n}$ form a right-handed orthogonal system, $\vec{n} = \vec{t}_\alpha \Lambda \vec{t}_\beta$.

The radii of curvature in the directions of α and β are denoted by R_α and R_β, respectively, and are taken to be positive when the centers of curvature lie in the positive direction of $\vec{n}$.

Let ζ be a rectilinear coordinate measured along the normal $\vec{n}$ to S. The following relations hold in the given triorthogonal system of curvilinear coordinates:

Square of a line element:

$$(d\eta)^2 = H_\alpha^2 (d\alpha)^2 + H_\beta^2 (d\beta)^2 + (d\zeta)^2$$

Area of an infinitesimal rectangle on the reference surface:

$$dS = AB d\alpha d\beta$$

Area of an infinitesimal rectangle on the edge surface S_e:

$$dS_e = H(\zeta) d\zeta dC$$

Volume of an infinitesimal parallelepiped:

$$dV = H_\alpha H_\beta d\alpha d\beta d\zeta$$

where[1]

$$H_\alpha = A(1 - R_\alpha^{-1}\zeta) \quad H_\beta = B(1 - R_\beta^{-1}\zeta)$$

$$H(\zeta) = ((mH_\alpha/A)^2 + (lH_\beta/B)^2)^{1/2}$$

H_α and H_β are the coefficients of the second fundamental form of the shell reference surface (Lamé coefficients).

2 Strain-Displacement Relations

Following Washizu (1968), it is found that the following strain-displacement relations hold in the Gaussian curvilinear coordinates:

$$H_\alpha e_{\alpha\alpha} = \partial_\alpha u + v B^{-1} \partial_\beta A - w A R_\alpha^{-1} \tag{1}$$

$$H_\beta e_{\beta\beta} = \partial_\beta v + u A^{-1} \partial_\alpha B - w B R_\beta^{-1} \tag{2}$$

$$e_{\zeta\zeta} = \partial_\zeta w \tag{3}$$

$$H_\alpha H_\beta e_{\alpha\beta} = H_\alpha^2 \partial_\beta (H_\alpha^{-1} u) + H_\beta^2 \partial_\alpha (H_\beta^{-1} v) \tag{4}$$

$$H_\alpha e_{\alpha\zeta} = \partial_\alpha w + H_\alpha^2 \partial_\zeta (H_\alpha^{-1} u) \tag{5}$$

$$H_\beta e_{\beta\zeta} = \partial_\beta w + H_\beta^2 \partial_\zeta (H_\beta^{-1} v) \tag{6}$$

Here, u, v, and w are the displacement components of an arbitrary point in the direction of the $\vec{t}_\alpha$, $\vec{t}_\beta$ and $\vec{n}$.

The approximate displacement field is assumed to be of the following form

$$u(\alpha,\beta,\zeta) = (1 - R_\alpha^{-1}\zeta) u^0 + \zeta(\gamma_{\alpha\zeta}^0 - A^{-1}\partial_\alpha w^0) + \Sigma_k \psi_k (\zeta - \zeta_k) Y(\zeta - \zeta_k) \tag{7}$$

$$v(\alpha,\beta,\zeta) = (1 - R_\beta^{-1}\zeta) v^0 + \zeta(\gamma_{\beta\zeta}^0 - B^{-1}\partial_\beta w^0) + \Sigma_k \theta_k (\zeta - \zeta_k) Y(\zeta - \zeta_k) \tag{8}$$

$$w(\alpha,\beta,\zeta) = w^0 \tag{9}$$

where u^0, v^0, and w^0 are displacements of a point on the shell reference surface; $\gamma_{\kappa\zeta}^0$ and $\gamma_{\beta\zeta}^0$ are the values of the shear rotations in the (α, ζ) and (β, ζ) planes, respectively; $Y(\zeta - \zeta_k)$ is the Heaviside unit function and $\psi_k(\alpha, \beta)$ and $\theta_k(\alpha, \beta)$ are functions to be determined by satisfying the contact conditions on the transverse shearing stresses at the interfaces k. In addition, Σ_k denotes the summation for k ranging from 1 to $N-1$, N being the number of the layers.

Substituting the above expressions for the displacements into equations (1)-(6), we obtain the following strain-displacement relations for a linearized theory of layered anisotropic shells including the transverse shear deformation effects and the distortion of the deformed normal:

$$H_\alpha e_{\alpha\alpha} = e_{\alpha\alpha}^0 + \zeta K_\alpha + \Sigma_k (\partial_\alpha \psi_k + \theta_k B^{-1} \partial_\beta A)(\zeta - \zeta_k) Y(\zeta - \zeta_k) \tag{1a}$$

$$H_\beta e_{\beta\beta} = e_{\beta\beta}^0 + \zeta K_\beta + \Sigma_k (\partial_\beta \theta_k + \psi_k A^{-1} \partial_\alpha B)(\zeta - \zeta_k) Y(\zeta - \zeta_k) \tag{2a}$$

$$\begin{aligned} H_\alpha H_\beta e_{\alpha\beta} = {} & H_\alpha \{\partial_\beta u^0 - v^0 A^{-1} \partial_\alpha B \\ & + \zeta[\partial_\beta(\gamma_{\alpha\zeta}^0 - A^{-1}\partial_\alpha w^0 - u^0 R_\alpha^{-1}) - \\ & - A^{-1}(\gamma_{\beta\zeta}^0 - B^{-1}\partial_\beta w^0 - v^0 R_\beta^{-1})\partial_\alpha B]\} \\ & + H_\beta \{\partial_\alpha v^0 - u^0 B^{-1}\partial_\beta A + \\ & + \zeta[\partial_\alpha(\gamma_{\beta\zeta}^0 - B^{-1}\partial_\beta w^0 - v^0 R_\beta^{-1}) \\ & - B^{-1}(\gamma_{\alpha\zeta}^0 - A^{-1}\partial_\alpha w^0 - u^0 R_\alpha^{-1})\partial_\beta A]\} + \\ & + \Sigma_k [H_\alpha(\partial_\beta \psi_k - \theta_k A^{-1}\partial_\alpha B) + H_\beta(\partial_\alpha \theta_k \\ & - \psi_k B^{-1}\partial_\beta A)](\zeta - \zeta_k) Y(\zeta - \zeta_k) \end{aligned} \tag{4a}$$

$$H_\alpha e_{\alpha\zeta} = A\gamma_{\alpha\zeta}^0 + A\Sigma_k \psi_k (1 - R_\alpha^{-1}\zeta_k) Y(\zeta - \zeta_k) \tag{5a}$$

$$H_\beta e_{\beta\zeta} = B\gamma_{\beta\zeta}^0 + B\Sigma_k \theta_k (1 - R_\beta^{-1}\zeta_k) Y(\zeta - \zeta_k) \tag{6a}$$

where

$$e_{\alpha\alpha}^0 = \partial_\alpha u^0 + v^0 B^{-1}\partial_\beta A - w^0 A R_\alpha^{-1};$$

$$e_{\beta\beta}^0 = \partial_\beta v^0 + u^0 A^{-1}\partial_\alpha B - w^0 B R_\beta^{-1} \tag{10}$$

$$\begin{aligned} K_\alpha = {} & -\partial_\alpha(A^{-1}\partial_\alpha w^0 + u^0 R_\alpha^{-1}) - B^{-1}(B^{-1}\partial_\beta w^0 \\ & + v^0 R_\beta^{-1})\partial_\beta A \\ & + \partial_\alpha \gamma_{\alpha\zeta}^0 + \gamma_{\beta\zeta}^0 B^{-1}\partial_\beta A \end{aligned} \tag{11}$$

$$\begin{aligned} K_\beta = {} & -\partial_\beta(B^{-1}\partial_\beta w^0 + v^0 R_\beta^{-1}) - A^{-1}(A^{-1}\partial_\alpha w^0 \\ & + u^0 R_\alpha^{-1})\partial_\alpha B \\ & + \partial_\beta \gamma_{\beta\zeta}^0 + \gamma_{\alpha\zeta}^0 A^{-1}\partial_\alpha B \end{aligned} \tag{12}$$

3 Expressions for Ψ_k and Θ_k

Let us consider a shell of constant thickness h consisting of N parallel thin layers of anisotropic materials perfectly bonded together. The thickness of each layer is assumed to be constant and the material to possess a plane of elastic symmetry parallel to the reference surface ζ = const. The material properties and the thickness of each layer may be entirely different. The shell reference surface S is superposed on the top bounding surface of the shell[2] and the unit vector $\vec{n}$ is directed inwards in the material of the shell. Owing to the existence of a plane of elastic symmetry, the constitutive relations for any individual layer are given by

$$\begin{bmatrix} \sigma_{\alpha\alpha} \\ \sigma_{\beta\beta} \\ \sigma_{\zeta\zeta} \\ \sigma_{\alpha\beta} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{36} \\ C_{16} & C_{26} & C_{36} & C_{66} \end{bmatrix} \begin{bmatrix} e_{\alpha\alpha} \\ e_{\beta\beta} \\ e_{\zeta\zeta} \\ e_{\alpha\beta} \end{bmatrix}$$

$$\begin{bmatrix} \sigma_{\alpha\zeta} \\ \sigma_{\beta\zeta} \end{bmatrix} = \begin{bmatrix} C_{44} & C_{45} \\ C_{45} & C_{55} \end{bmatrix} \begin{bmatrix} e_{\alpha\zeta} \\ e_{\beta\zeta} \end{bmatrix} \tag{13}$$

where C_{ij} are the elastic coefficients and the usual notation for stresses and engineering strain components has been adopted.

[1] (l, m) stands for the director cosines of the outward normal to C, where C is the intersection curve between the reference surface S and the edge surface S_e.

[2] Obviously, other choices are possible. For example, for symmetric laminates several partial mathematical simplifications are achieved if the middle surface is selected as reference surface.

If a given layer consists of a fiber reinforced composite material, it is treated as an orthotropic sheet, the axes of material symmetry being parallel and normal to the fiber direction. Thus only 9, rather 13, of coefficients C_{ij} are independent in this case.

As usual, the normal stress $\sigma_{\zeta\zeta}$ is assumed to be small in comparison with other normal stresses and is neglected. As a consequence, local effects, such as boundary layer phenomena, presence of load or geometric discontinuities, etc., are beyond the capability of the proposed approach. Taking into account this assumption and eliminating $\sigma_{\zeta\zeta}$ from the above relations, the constitutive relations for each layer assume the following contracted form

$$\begin{bmatrix} \sigma_{\alpha\alpha} \\ \sigma_{\beta\beta} \\ \sigma_{\alpha\beta} \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & Q_{16} \\ Q_{12} & Q_{22} & Q_{26} \\ Q_{16} & Q_{26} & Q_{66} \end{bmatrix} \begin{bmatrix} e_{\alpha\alpha} \\ e_{\beta\beta} \\ e_{\alpha\beta} \end{bmatrix}$$

$$\begin{bmatrix} \sigma_{\alpha\zeta} \\ \sigma_{\beta\zeta} \end{bmatrix} = \begin{bmatrix} Q_{44} & Q_{45} \\ Q_{45} & Q_{55} \end{bmatrix} \begin{bmatrix} e_{\alpha\zeta} \\ e_{\beta\zeta} \end{bmatrix} \tag{13a}$$

where $Q_{ij} = C_{ij} - C_{i3}/C_{33}$ for $i, j = 1, 2, 6$ and $Q_{ij} = C_{ij}$ for $i, j = 4, 5$.

For shells consisting of layers perfectly bonded together (the layers of the shell function concurrently without slippage), we know from elasticity theory that the displacements and stresses at the interface k between the kth and $(k + 1)$th bonded layers must satisfy the following contact conditions

$$u_k = u_{k+1} \quad v_k = v_{k+1} \quad w_k = w_{k+1} \tag{14}$$

$$\sigma_{\alpha\zeta,k} = \sigma_{\alpha\zeta,k+1} \quad \sigma_{\beta\zeta,k} = \sigma_{\beta\zeta,k+1} \quad \sigma_{\zeta\zeta,k} = \sigma_{\zeta\zeta,k+1} \tag{15}$$

In examining the relations (7)–(9) it is not difficult to realize that the displacements u and v are continuous functions of the ζ coordinate for all values of the $\psi_k(\alpha, \beta)$ and $\theta_k(\alpha, \beta)$. It follows that the expressions for $\psi_k(\alpha, \beta)$ and $\theta_k(\alpha, \beta)$ may be found from the contact conditions for the transverse shearing stresses $\sigma_{\alpha\zeta}$ and $\sigma_{\beta\zeta}$.

According to the previous relations, the transverse shearing stresses at the interface k are given by the following relations

where

$$\sigma^-_{\alpha\zeta} = Q_{44}(k)e^-_{\alpha\zeta} + Q_{45}(k)e^-_{\beta\zeta} \qquad \sigma^-_{\beta\zeta} = Q_{45}(k)e^-_{\alpha\zeta} + Q_{55}(k)e^-_{\beta\zeta} \tag{16a}$$

$$\sigma^+_{\alpha\zeta} = Q_{44}(k+1)e^+_{\alpha\zeta} + Q_{45}(k+1)e^+_{\beta\zeta} \qquad \sigma^+_{\beta\zeta} = Q_{45}(k+1)e^+_{\alpha\zeta} + Q_{55}(k+1)e^+_{\beta\zeta} \tag{16b}$$

$$e^-_{\alpha\zeta} = [\gamma^0_{\alpha\zeta} + \sum^{k-1}_q \psi_q(1 - R_\alpha^{-1}\zeta_q)]/(1 - R_\alpha^{-1}\zeta_k); \quad e^+_{\alpha\zeta} = e^-_{\alpha\zeta} + \psi_k \tag{17a}$$

$$e^-_{\beta\zeta} = [\gamma^0_{\beta\zeta} + \sum^{k-1}_q \theta_q(1 - R_\beta^{-1}\zeta_q)]/(1 - R_\beta^{-1}\zeta_k); \quad e^+_{\beta\zeta} = e^-_{\beta\zeta} + \theta_k \tag{17b}$$

Here, the symbols $-$ or $+$ refer to the values of the functions for $\zeta = \zeta_k - 0$ and $\zeta = \zeta_k + 0$, respectively, where ζ_k is the value of the ζ coordinate at the interface k.

By substituting the expressions for $e^+_{\alpha\zeta}$ and $e^+_{\beta\zeta}$ into equations (16b) and satisfying the contact conditions for the transverse shearing stresses, yields

$$\begin{bmatrix} \psi_k \\ \theta_k \end{bmatrix} = \frac{1}{R} \begin{bmatrix} -Q_{55}(k+1) & Q_{45}(k+1) \\ Q_{45}(k+1) & -Q_{44}(k+1) \end{bmatrix} \begin{bmatrix} \Delta Q^k_{44} & \Delta Q^k_{45} \\ \Delta Q^k_{45} & \Delta Q^k_{55} \end{bmatrix} \begin{bmatrix} e^-_{\alpha\zeta} \\ e^-_{\beta\zeta} \end{bmatrix} \tag{18}$$

where we have posed

$$\Delta Q^k_{ij} = Q_{ij}(k+1) - Q_{ij}(k);$$

$$R = Q_{44}(k+1)Q_{55}(k+1) - Q^2_{45}(k+1)$$

It is readily realized, by substitution of the expressions for $e^-_{\alpha\zeta}$ and $e^-_{\beta\zeta}$ into equation (18), that ψ_k and θ_k are known functions of the generalized coordinates $\gamma^0_{\alpha\zeta}$ and $\gamma^0_{\beta\zeta}$; therefore, we pose

$$\psi_k = a_k\gamma^0_{\alpha\zeta} + c_k\gamma^0_{\beta\zeta}; \quad \theta_k = d_k\gamma^0_{\alpha\zeta} + b_k\gamma^0_{\beta\zeta} \tag{19}$$

where a_k, b_k, c_k, and d_k are known constants only depending on the transverse shear mechanical properties of the various layers.

4 Governing Equations

General Equation of Dynamics. Consider the motion of an elastic body under prescribed surface tractions and boundary conditions. If the body is assumed to execute an arbitrary set of infinitesimal virtual displacements δu, δv, and δw from the actual configuration, the following variational equation of motion holds (Washizu, 1968)

$$\delta\Phi - \delta W + \int_V \mu(\ddot{u}\delta u + \ddot{v}\delta v + \ddot{w}\delta w)dV = 0 \tag{20}$$

In the above, μ is the material mass density, W the work done by the applied external forces, and Φ the strain energy. In addition, the overdot indicates differentiation with respect to time t and V is the volume of the body under consideration.

Strain Energy. If $V(s)$ stands for the volume of the sth layer, the variation of the strain energy is given by (the index s is dropped for sake of simplicity)

$$\delta\Phi = \sum_{s=1}^{N} \int_{V(s)} [\sigma_{\alpha\alpha}\delta e_{\alpha\alpha} + \sigma_{\beta\beta}\delta e_{\beta\beta} + \sigma_{\alpha\beta}\delta e_{\alpha\beta} + \sigma_{\alpha\zeta}\delta e_{\alpha\zeta} + \sigma_{\beta\zeta}\delta e_{\beta\zeta}]H_\alpha H_\beta d\alpha d\beta d\zeta \tag{21}$$

By taking into account equations (1a)–(6a) and the expressions (19) for ψ_k and θ_k, we obtain after some straightforward manipulations

$$\begin{aligned}
\delta\Phi = -\int_S \{ & [\partial_\alpha(BN_\alpha) + \partial_\beta(AN_{\beta\alpha}) + N_{\alpha\beta}\partial_\beta A - N_\beta\partial_\alpha B - ABR_\alpha^{-1}\tilde{Q}_\alpha]\delta u^0 + \\
& + [\partial_\alpha(BN_{\alpha\beta}) + \partial_\beta(AN_\beta) + N_{\beta\alpha}\partial_\alpha B - N_\alpha\partial_\beta A - ABR_\beta^{-1}\tilde{Q}_\beta]\delta v^0 + \\
& + [AB(R_\alpha^{-1}N_\alpha + R_\beta^{-1}N_\beta) + \partial_\alpha(B\tilde{Q}_\alpha) + \partial_\beta(A\tilde{Q}_\beta)]\delta w^0 + \\
& + AB[(\tilde{Q}_\alpha - Q_\alpha) + (\tilde{Q}^a_\alpha - Q^a_\alpha) + (\tilde{Q}^d_\beta - Q^d_\beta)]\delta\gamma^0_{\alpha\zeta} + \\
& + AB[(\tilde{Q}_\beta - Q_\beta) + (\tilde{Q}^b_\beta - Q^b_\beta) + (\tilde{Q}^c_\alpha - Q^c_\alpha)]\delta\gamma^0_{\beta\zeta}\}d\alpha d\beta + \\
& + \int_C \{[(N_\alpha - R_\alpha^{-1}M_\alpha)l + (N_{\beta\alpha} - R_\alpha^{-1}M_{\beta\alpha})m]\delta u^0 \\
& + [(N_{\alpha\beta} - R_\beta^{-1}M_{\alpha\beta})l + (N_\beta - R_\beta^{-1}M_\beta)m]\delta v^0 + (\tilde{Q}_\alpha l + \tilde{Q}_\beta m)\delta w^0 \\
& - A^{-1}(M_\alpha l + M_{\beta\alpha}m)\delta\partial_\alpha w^0 - B^{-1}(M_{\alpha\beta}l + M_\beta m)\delta\partial_\beta w^0 + [(M_\alpha + M^a_\alpha \\
& + M^d_{\alpha\beta})l + (M_{\beta\alpha} + M^a_{\beta\alpha} + M^d_\beta)m]\delta\gamma^0_{\alpha\zeta} + \\
& + [(M_\beta + M^b_\beta + M^c_{\beta\alpha})m + (M_{\alpha\beta} + M^b_{\alpha\beta} + M^c_\alpha)l]\delta\gamma^0_{\beta\zeta}\}dC
\end{aligned}$$

where we have introduced the following notation for the stress resultants

$$(N_\alpha, N_{\alpha\beta}) = B^{-1} < H_\beta (\sigma_{\alpha\alpha}, \sigma_{\alpha\beta}) > \qquad (N_\beta, N_{\beta\alpha}) = A^{-1} < H_\alpha (\sigma_{\beta\beta}, \sigma_{\beta\alpha}) >$$

$$(M_\alpha, M_{\alpha\beta}) = B^{-1} < H_\beta (\sigma_{\alpha\alpha}, \sigma_{\alpha\beta}) \zeta > \qquad (M_\beta, M_{\beta\alpha}) = A^{-1} < H_\alpha (\sigma_{\beta\beta}, \sigma_{\beta\alpha}) \zeta >$$

$$Q_\alpha = B^{-1} < H_\beta \sigma_{\alpha\zeta} > \qquad Q_\beta = A^{-1} < H_\alpha \sigma_{\beta\zeta} >$$

$$AB\tilde{Q}_\alpha = \partial_\alpha (BM_\alpha) + \partial_\beta (AM_{\beta\alpha}) + M_{\alpha\beta}\partial_\beta A - M_\beta \partial_\alpha B$$

$$AB\tilde{Q}_\beta = \partial_\beta (AM_\beta) + \partial_\alpha (BM_{\alpha\beta}) + M_{\beta\alpha}\partial_\alpha B - M_\alpha \partial_\beta A$$

$$\begin{bmatrix} (M_\alpha^a, M_\alpha^b, M_\alpha^c, M_\alpha^d) \\ (M_{\alpha\beta}^a, M_{\alpha\beta}^b, M_{\alpha\beta}^c, M_{\alpha\beta}^d) \end{bmatrix} = B^{-1} < H_\beta \begin{bmatrix} \sigma_{\alpha\alpha} \\ \sigma_{\alpha\beta} \end{bmatrix} \Sigma_k (a_k, b_k, c_k, d_k)(\zeta - \zeta_k) >$$

$$\begin{bmatrix} (M_\beta^a, M_\beta^b, M_\beta^c, M_\beta^d) \\ (M_{\beta\alpha}^a, M_{\beta\alpha}^b, M_{\beta\alpha}^c, M_{\beta\alpha}^d) \end{bmatrix} = A^{-1} < H_\alpha \begin{bmatrix} \sigma_{\beta\beta} \\ \sigma_{\beta\alpha} \end{bmatrix} \Sigma_k (a_k, b_k, c_k, d_k)(\zeta - \zeta_k) >$$

$$(Q_\alpha^a, Q_\alpha^c) = B^{-1} < H_\beta \sigma_{\alpha\zeta} \Sigma_k (a_k, c_k)(1 - R_\alpha^{-1}\zeta_k) >$$

$$(Q_\beta^b, Q_\alpha^d) = A^{-1} < H_\alpha \sigma_{\beta\zeta} \Sigma_k (b_k, d_k)(1 - R_\beta^{-1}\zeta) >$$

For notational convenience, $< >$ is defined by

$$< . . > = \sum_{s=1}^{N} \int_{\zeta_{s-1}}^{\zeta_s} (. .) d\zeta$$

and Σ_k stands for $\sum_{k=1}^{s-1}$. If we introduce the following quantities

$$N_{\alpha\nu} = N_\alpha l + N_{\alpha\beta} m \qquad N_{\beta\nu} = N_{\alpha\beta} l + N_\beta m$$

$$M_{\alpha\nu} = M_\alpha l + M_{\beta\alpha} m \qquad M_{\beta\nu} = M_{\alpha\beta} l + M_\beta m$$

$$M_\nu = M_{\alpha\nu} l + M_{\beta\nu} m \qquad M_{\nu t} = -M_{\alpha\nu} m + M_{\beta\nu} l$$

$$V_n = Q_\alpha l + Q_\beta m$$

and take into account that

$$A^{-1}\partial_\alpha = l\partial_\nu - m\partial_t; \quad B^{-1}\partial_\beta = m\partial_\nu + l\partial_t$$

an integration by parts on terms like $(. .)\, \delta\partial_t w^0$, yields

$$\begin{aligned}
\delta\Phi = -\int_S \{ & [\partial_\alpha (BN_\alpha) + \partial_\beta (AN_{\beta\alpha}) + N_{\alpha\beta}\partial_\beta A \\
& - N_\beta \partial_\beta B - ABR_\alpha^{-1}\tilde{Q}_\alpha]\delta u^0 + \\
& + [\partial_\alpha (BN_{\alpha\beta}) + \partial_\beta (AN_\beta) + N_{\beta\alpha}\partial_\alpha B \\
& - N_\alpha \partial_\beta A - ABR_\beta^{-1}\tilde{Q}_\beta]\delta v^0 + \\
& + [AB(R_\alpha^{-1}N_\alpha + R_\beta^{-1}N_\beta) + \partial_\alpha (B\tilde{Q}_\alpha) + \partial_\beta (A\tilde{Q}_\beta)]\delta w^0 + \\
& + AB[(\tilde{Q}_\alpha - Q_\alpha) + (\tilde{Q}_\alpha^a - Q_\alpha^a) + (\tilde{Q}_\beta^d - Q_\beta^d)]\delta\gamma_{\alpha\zeta}^0 + \\
& + AB[(\tilde{Q}_\beta - Q_\beta) + (\tilde{Q}_\beta^b - Q_\beta^b) + (\tilde{Q}_\alpha^c - A_\alpha^c)]\delta\gamma_{\beta\zeta}^0 \} d\alpha d\beta + \\
& + \int_C \{ (N_{\alpha\beta} - R_\alpha^{-1}M_{\alpha\nu})\delta u^0 + (N_{\beta\nu} - R_\beta^{-1}M_{\beta\nu})\delta v^0 \\
& + (V_n + M_{\nu t,t})\delta w^0 - \\
& - M_\nu \delta\partial_\nu w^0 + (M_{\alpha\nu} + M_{\alpha\nu}^a + M_{\beta\nu}^d)\delta\gamma_{\beta\zeta}^0 \\
& + (M_{\beta\nu} + M_{\beta\nu}^b + M_{\alpha\nu}^c)\delta\gamma_{\beta\zeta}^0 \} dC - \\
& - M_{\nu t}\delta w^0 /_{C_p}
\end{aligned} \tag{22}$$

Work of External Forces. Let $\bar{p}_\alpha$, $\bar{p}_\beta$, and $\bar{p}_\zeta$ denote the α, β, and ζ components, respectively, of the surface tractions on S_p (S_p being that part of the surface of the shell on which tractions are prescribed).

If we suppose that the boundary surfaces of the shell are loaded only by normal loads $\bar{p}_\zeta$, the virtual work is given by

$$\delta W_b = \int_S \bar{q}_\zeta \delta w^0 dS \tag{23}$$

where $\bar{q}_\zeta = \bar{p}_\zeta(0) + \bar{p}_\zeta(h)\ (1 - R_\alpha^{-1}h)\ (1 - R_\beta^{-1}h)$ and $\bar{p}_\zeta(0)$ and $\bar{p}_\zeta(h)$ identify the quantity $\bar{p}_\zeta$ at the outer ($\zeta = 0$) and inner ($\zeta = h$) bounding surfaces of the shell, respectively.

In writing δW_b we have assumed that on the boundary surfaces only surface tractions, and not displacements, are prescribed.

Let $\bar{p}_\alpha(s)$, $\bar{p}_\beta(s)$, and $\bar{p}_\zeta(s)$ denote the α, β, and ζ components, respectively, of the edge forces. If S_p^e denotes that part of the edge surface S_e of the shell on which surface tractions are prescribed, the virtual work done by these external forces is

$$\delta W_e = \sum_{s=1}^{N} \int_{S_p^e} [\bar{p}_\alpha(s)\delta u + \bar{p}_\beta(s)\delta v + \bar{p}_\zeta(s)\delta w] dS_e$$

If equations (7)–(9) are substituted into the above, the expression for δW_e becomes, after an integration by parts,

$$\begin{aligned}
\delta W_e = \int_{C_p} & [(\bar{N}_{\alpha\nu} - R_\alpha^{-1}\bar{M}_{\alpha\nu})\delta u^0 \\
& + (\bar{N}_{\beta\nu} - R_\beta^{-1}\bar{M}_{\beta\nu})\delta v^0 + (\bar{V}_n + \bar{M}_{\nu t,t})\delta w^0 - \\
& - \bar{M}_\nu \partial_\nu \delta w^0 + (\bar{M}_{\alpha\nu} + \bar{M}_{\alpha\nu}^a + \bar{M}_{\beta\nu}^d)\delta\gamma_{\alpha\zeta}^0 \\
& + (\bar{M}_{\beta\nu} + \bar{M}_{\beta\nu}^b + \bar{M}_{\alpha\nu}^c)\delta\gamma_{\beta\zeta}^0] dC \\
& - \bar{M}_{\nu t}\delta w^0 /_{C_p}
\end{aligned}$$

where

$$(\bar{N}_{\alpha\nu}, \bar{N}_{\beta\nu}, \bar{V}_n) = < H(\zeta)[\bar{p}_\alpha(s), \bar{p}_\beta(s), \bar{p}_\zeta(s)] >$$

$$(\bar{M}_{\alpha\nu}, \bar{M}_{\beta\nu}) = < \zeta H(\zeta)[\bar{p}_\alpha(s), \bar{p}_\beta(s)] >$$

$$\begin{bmatrix} \bar{M}_{\alpha\nu}^a \\ \bar{M}_{\alpha\nu}^c \end{bmatrix} = < \bar{p}_\alpha(s) H(\zeta) \Sigma_k \begin{bmatrix} a_k \\ c_k \end{bmatrix} (\zeta - \zeta_k) >$$

$$\begin{bmatrix} \bar{M}_{\beta\nu}^b \\ \bar{M}_{\beta\nu}^d \end{bmatrix} = < \bar{p}_\beta(s) H(\zeta) \Sigma_k \begin{bmatrix} b_k \\ d_k \end{bmatrix} (\zeta - \zeta_k) >$$

Work of Inertial Forces. Using the approximate expressions for δu, δv, and δw (as well as for $\ddot{u}$, $\ddot{v}$, and $\ddot{w}$) given by the equations (7)–(9), an approximate expression for the virtual variation of the work given by the inertial forces is obtained as

$$\delta W_{in} = -\sum_{s=1}^{N} \int_S \int_{\zeta_{s-1}}^{\zeta_s} \mu(s)(\ddot{u}\delta u + \ddot{v}\delta v + \ddot{w}\delta w) H_\alpha H_\beta d\alpha d\beta d\zeta =$$

$$= -\int_S \{ (M - 2R_\alpha^{-1}P + R_\alpha^{-2}J)\ddot{u}^0 + [(P + P^a)$$

$$-R_\alpha^{-1}(J+J^a)]\ddot{\gamma}^0_{\alpha\zeta}+(P^c-R_\alpha^{-1}J^c)\ddot{\gamma}^0_{\beta\zeta}-$$

$$-A^{-1}(P-R_\alpha^{-1}J)\partial_\alpha\ddot{w}^0\}\delta u^0 ABd\alpha d\beta-$$

$$-\int_S\{(M-2R_\beta^{-1}P+R_\beta^{-2}J)\ddot{v}^0+(P^d-R_\beta^{-1}J^d)\ddot{\gamma}^0_{\alpha\zeta}$$

$$+[(P+P^b)-R_\beta^{-1}(J+J^b)]\ddot{\gamma}^0_{\beta\zeta}-$$

$$-B^{-1}(P-R_\beta^{-1}J)\partial_\beta\ddot{w}^0\}\delta v^0 ABd\alpha d\beta-$$

$$-\int_S\{ABM\ddot{w}^0+\partial_\alpha[B(P-R_\alpha^{-1}J)\ddot{u}^0+$$

$$+B(J+J^a)\ddot{\gamma}^0_{\alpha\zeta}+BJ^c\ddot{\gamma}^0_{\beta\zeta}-A^{-1}BJ\partial_\alpha\ddot{w}^0]+$$

$$+\partial_\beta[A(P-R_\beta^{-1}J)\ddot{v}^0+AJ^d\ddot{\gamma}^0_{\alpha\zeta}+A(J+J^d)\ddot{\gamma}^0_{\beta\zeta}$$

$$-AB^{-1}J\partial_\beta\ddot{w}^0]\}\delta w^0 d\alpha d\beta-$$

$$-\int_S\{[(P+P^a)-R_\alpha^{-1}(J+J^a)]\ddot{u}^0+(P^d-R_\beta^{-1}J^d)\ddot{v}^0$$

$$+(J+2J^a+J^{aa}+J^{dd})\ddot{\gamma}^0_{\alpha\zeta}+$$

$$+(J^c+J^d+J^{ac}+J^{bd})\ddot{\gamma}^0_{\beta\zeta}-A^{-1}(J+J^a)\partial_\alpha\ddot{w}^0$$

$$-B^{-1}J^d\partial_\beta\ddot{w}^0\}\delta\gamma^0_{\alpha\zeta}ABd\alpha d\beta-$$

$$-\int_S\{(P^c-R_\alpha^{-1}J^c)\ddot{u}^0+[(P+P^b)-R_\beta^{-1}(J+J^b)]\ddot{v}^0$$

$$+(J^c+J^d+J^{ac}+J^{bd})\ddot{\gamma}^0_{\alpha\zeta}+$$

$$+(J+2J^b+J^{bb}+J^{cc})\ddot{\gamma}^0_{\beta\zeta}-A^{-1}J^c\partial_\alpha\ddot{w}^0$$

$$-B^{-1}(J+J^b)\partial_\beta\ddot{w}^0\}\delta\gamma^0_{\beta\zeta}ABd\alpha d\beta+$$

$$+\int_C\{[(P-R_\alpha^{-1}J)\ddot{u}^0+(J+J^a)\ddot{\gamma}^0_{\alpha\zeta}+J^c\ddot{\gamma}^0_{\beta\zeta}$$

$$-A^{-1}J\partial_\alpha\ddot{w}^0]l+$$

$$+[(P-R_\beta^{-1})\ddot{v}^0+J^d\ddot{\gamma}^0_{\alpha\zeta}+(J+J^b)\ddot{\gamma}^0_{\beta\zeta}$$

$$-B^{-1}\partial_\beta\ddot{w}^0]m\}\delta w^0 dC \tag{25}$$

where

$$[M\quad P\quad J]=(AB)^{-1}<[1\ \ \zeta\ \ \zeta^2]H_\alpha H_\beta\mu(s)>$$

$$\begin{bmatrix}[P^a\ \ P^c]\\ [J^a\ \ J^c]\end{bmatrix}=(AB)^{-1}<\begin{bmatrix}1\\ \zeta\end{bmatrix}H_\alpha H_\beta\mu(s)[A_k\ \ C_k]>$$

$$\begin{bmatrix}[P^b\ \ P^d]\\ [J^b\ \ J^d]\end{bmatrix}=(AB)^{-1}<\begin{bmatrix}1\\ \zeta\end{bmatrix}H_\alpha H_\beta\mu(s)[B_k\ \ D_k]>$$

$$[J^{aa}J^{bb}J^{cc}J^{dd}]=(AB)^{-1}<H_\alpha H_\beta\mu(s)[A_kA_r\ B_kB_r\ C_kC_r\ D_kD_r]>$$

$$[J^{ac}J^{bd}]=(AB)^{-1}<H_\alpha H_\beta\mu(s)[A_kC_r\ B_kD_r]>$$

with $(A_k, B_k, C_k, D_k) = \Sigma_k\ (a_k, b_k, c_k, d_k)(\zeta-\zeta_k)$.

Equations of Motion and Boundary Conditions. Inserting the expressions for $\delta\Phi$, δW_b, δW_e, and δW_{in} given by equations (22) through (25) into equation (20), we obtain the differential equations of motion and the boundary conditions in terms of force and moment resultants.

The vanishing of the coefficients of the virtual variations in the surface integral gives us the following differential equations of motion:

$$\partial_\alpha(BN_\alpha)+\partial_\beta(AN_{\beta\alpha})+N_{\alpha\beta}\partial_\beta A-N_\beta\partial_\alpha B-ABR_\alpha^{-1}\tilde{Q}_\alpha$$
$$=AB\{M-2R_\alpha^{-1}P+R_\alpha^{-2}J)\ddot{u}^0+$$
$$+[(P+P^a)-R_\alpha^{-1}(J+J^a)]\ddot{\gamma}^0_{\alpha\zeta}+(P^c-R_\alpha^{-1}J^c)\ddot{\gamma}^0_{\beta\zeta}$$
$$-A^{-1}(P-R_\alpha^{-1}J)\partial_\alpha\ddot{w}^0\} \tag{26}$$

$$\partial_\alpha(BN_{\alpha\beta})+\partial_\beta(AN_\beta)+N_{\beta\alpha}\partial_\alpha B-N_\alpha\partial_\beta A-ABR_\beta^{-1}\tilde{Q}_\beta$$
$$=AB\{(M-2R_\beta^{-1}P+R_\beta^{-2}J)\ddot{v}^0+$$
$$+(P^d-R_\beta^{-1}J^d)\ddot{\gamma}^0_{\alpha\zeta}+[(P+P^b)-R_\beta^{-1}(J+J^b)]\ddot{\gamma}^0_{\beta\zeta}$$
$$-B^{-1}(P-R_\beta^{-1}J)\partial_\beta\ddot{w}^0\} \tag{27}$$

$$\partial_\alpha(B\tilde{Q}_\alpha)+\partial_\beta(A\tilde{Q}_\beta)+AB(R_\alpha^{-1}N_\alpha+R_\beta^{-1}N_\beta)+AB\tilde{q}_\zeta$$
$$=ABM\ddot{w}^0+\partial_\alpha[B(P-R_\alpha^{-1}J)\ddot{u}^0+$$
$$+B(J+J^a)\ddot{\gamma}^0_{\alpha\zeta}+BJ^c\ddot{\gamma}^0_{\beta\zeta}-A^{-1}BJ\partial_\alpha\ddot{w}^0]+$$
$$+\partial_\beta[A(P-R_\beta^{-1}J)\ddot{v}^0+AJ^d\ddot{\gamma}^0_{\alpha\zeta}+$$
$$+A(J+J^d)\ddot{\gamma}^0_{\beta\zeta}-AB^{-1}J\partial_\beta\ddot{w}^0] \tag{28}$$

$$(\tilde{Q}_\alpha-Q_\alpha)+(\tilde{Q}^a_\alpha-Q^a_\alpha)+(\tilde{Q}^d_\beta-Q^d_\beta)=[(P+P^a)-R_\alpha^{-1}(J$$
$$+J^a)]\ddot{u}^0+(P^d-R_\beta^{-1}J^d)\ddot{v}^0+$$
$$+(J+2J^a+J^{aa}+J^{dd})\ddot{\gamma}^0_{\alpha\zeta}+(J^c+J^d+J^{ac}+J^{bd})\ddot{\gamma}^0_{\beta\zeta}$$
$$-A^{-1}(J+J^a)\partial_\alpha\ddot{w}^0-B^{-1}J^d\partial_\beta\ddot{w}^0 \tag{29}$$

$$(\tilde{Q}_\beta-Q_\beta)+(\tilde{Q}^b_\beta-Q^b_\beta)+(\tilde{Q}^c_\alpha-Q^c_\alpha)=(P^c-R_\alpha^{-1}J^c)\ddot{u}^0$$
$$+[(P+P^b)-R_\beta^{-1}(J+J^b)]\ddot{v}^0+$$
$$+(J^c+J^d+J^{ac}+J^{bd})\ddot{\gamma}^0_{\alpha\zeta}+(J+2J^b+J^{bb}$$
$$+J^{cc})\ddot{\gamma}^0_{\beta\zeta}-A^{-1}J^c\partial_\alpha\ddot{w}^0-B^{-1}(J+J^b)\partial_\beta\ddot{w}^0 \tag{30}$$

From the vanishing of the terms in the line integral, the following natural (mechanical) and prescribed (geometrical) boundary conditions are obtained:

Natural Boundary Conditions on C_p	or	Prescribed Boundary Condition on C_u
$N_{\alpha\nu}-R_\alpha^{-1}M_{\alpha\nu}=\bar{N}_{\alpha\nu}-R_\alpha^{-1}\bar{M}_{\alpha\nu}$		$u^0=\bar{u}^0$
$N_{\beta\nu}-R_\beta^{-1}M_{\beta\nu}=\bar{N}_{\beta\nu}-R_\beta^{-1}\bar{M}_{\beta\nu}$		$v^0=\bar{v}^0$
$V_n+M_{\nu t,t}=\bar{V}_n+\bar{M}_{\nu t,t}+[(P-R_\alpha^{-1}J)\ddot{u}^0+(J+J^a)\ddot{\gamma}^0_{\alpha\zeta}-$ $-J^c\ddot{\gamma}^0_{\beta\zeta}-A^{-1}J\partial_\alpha\ddot{w}^0]l+[(P-R_\beta^{-1}J)\ddot{v}^0+$ $+J^d\ddot{\gamma}^0_{\alpha\zeta}+(J+J^b)\ddot{\gamma}^0_{\beta\zeta}-B^{-1}\partial_\beta\ddot{w}^0]m$		$w^0=\bar{w}^0$
$M_\nu=\bar{M}_\nu$		$\partial_\nu w^0=\overline{\partial_\nu w^0}$
$M^*_{\alpha\nu}=\bar{M}^*_{\alpha\nu}$		$\gamma^0_{\alpha\zeta}=\bar{\gamma}^0_{\alpha\zeta}$
$M^*_{\beta\nu}=\bar{M}^*_{\alpha\nu}$		$\gamma^0_{\beta\zeta}=\bar{\gamma}^0_{\beta\zeta}$

where we have introduced the following notation

$$M^*_{\alpha\nu} = M_{\alpha\nu} + M^a_{\alpha\nu} + M^d_{\beta\nu} \quad M^*_{\beta\nu} = M_{\beta\nu} + M^b_{\beta\nu} + M^c_{\alpha\nu}$$

Analoguos expressions hold for $\bar{M}^*_{\alpha\nu}$ and $\bar{M}^*_{\beta\nu}$.

5 Shells of Revolution

In the engineering applications of thin shells, a shell whose reference surface is in the form of a surface of revolution has extensive usage. Thus, our concern in the rest of this paper will be a theory of the shells of revolution.

Let R_o be the radius of a latitude circle (also called a parallel). The position of a meridian is defined by the angle ω (the azimuth of the meridian plane), measured from some datum meridian plane; the equation of the meridian is $R_o = R_o(x)$.

The position of a latitude circle is defined by the angle ϕ, made by the normal to the reference surface and the axis of rotation. If the description of the reference surface of the shell is based on the independent variables ϕ and ω for the square of a linear element on the reference surface, we have

$$(d\eta)^2 = R^2_\phi (d\phi)^2 + R^2_0 (d\omega)^2$$

where R_ϕ stands for the radius of curvature of the meridian.

The first term on the right-hand side of the above equation represents the square of the differential length of arc along a meridian, and the second terms represents the square of the differential length of arc along a parallel.

If we associate α with ϕ and β with ω, for the coefficients of the first quadratic form we obtain the following expressions

$$A = R_\phi \quad B = R_0 = R_\omega \sin\phi$$

where R_ω stands for the second radius of curvature of the surface. Furthermore, for a surface of revolution the relation $\partial_\phi R_0 = R_\phi \cos\phi$ holds.

With the parameters that we have identified above and with the observation that in a shell of revolution R_0, R_ϕ, and R_ω are independent of ω, the above equations reduce to:

Displacement Field.

$$u(\phi,\omega,\zeta) = (1 - R_\phi^{-1}\zeta)u^0 + \zeta(\gamma^0_{\phi\zeta} - R_\phi^{-1}\partial_\phi w^0) + \Sigma_k \psi_k(\zeta - \zeta_k) Y(\zeta - \zeta_k) \tag{7R}$$

$$v(\phi,\omega,\zeta) = (1 - R_\omega^{-1}\zeta)v^0 + \zeta(\gamma^0_{\omega\zeta} - R_0^{-1}\partial_\omega w^0) + \Sigma_k \theta_k(\zeta - \zeta_k) Y(\zeta - \zeta_k) \tag{8R}$$

$$w(\phi,\omega,\zeta) = w^0 \tag{9R}$$

Strain-Displacement Relations.

$$R_\phi e_{\phi\phi} = e^0_{\phi\phi} + \zeta K_\phi + \Sigma_k \partial_\phi \psi_k(\zeta - \zeta_k) Y(\zeta - \zeta_k) \tag{1R}$$

$$R_0 e_{\omega\omega} = e^0_{\omega\omega} + \zeta K_\omega + \Sigma_k (\partial_\omega \theta_k + \psi_k R_\phi^{-1}\partial_\phi R_0)(\zeta - \zeta_k) Y(\zeta - \zeta_k) \tag{2R}$$

$$\begin{aligned} R_\phi R_0 e_{\phi\omega} = {} & R_\phi \partial_\omega u^0 - v^0 \partial_\phi R_0 + R_\phi \zeta \partial_\omega (\gamma^0_{\phi\zeta} - R_\phi^{-1}\partial_\phi w^0 - u^0 R_\phi^{-1}) \\ & - (\gamma^0_{\omega\zeta} - R_0^{-1}\partial_\omega w^0 - v^0 R_\omega^{-1})\partial_\phi R_0 + R_0 \partial_\phi v^0 \\ & + R_0 \zeta \partial_\phi (\gamma^0_{\omega\zeta} - R_0^{-1}\partial_\omega w^0 - v^0 R_\omega^{-1}) \\ & + \Sigma_k (R_\phi \partial_\omega \psi_k - \theta_k \partial_\phi R_0 + R_0 \partial_\phi \theta_k)(\zeta - \zeta_k) Y(\zeta - \zeta_k) \end{aligned} \tag{4R}$$

$$e_{\phi\zeta} = \gamma^0_{\phi\zeta} + \Sigma_k \psi_k Y(\zeta - \zeta_k) \tag{5R}$$

$$e_{\omega\zeta} = \gamma^0_{\omega\zeta} + \Sigma_k \theta_k Y(\zeta - \zeta_k) \tag{6R}$$

where

$$e^0_{\phi\phi} = \partial_\phi u^0 - w^0 \quad e^0_{\omega\omega} = \partial_\omega v^0 + u^0 R_\phi^{-1}\partial_\phi R_0 - w^0 R_0 R_\omega^{-1} \tag{10R}$$

$$K_\phi = -\partial_\phi (R_\phi^{-1}\partial_\phi w^0 + u^0 R_\phi^{-1}) + \partial_\phi \gamma^0_{\phi\zeta} \tag{11R}$$

$$\begin{aligned} K_\omega = {} & -\partial_\omega (R_0^{-1}\partial_\omega w^0 + v^0 R_\omega^{-1}) \\ & - R_\phi^{-1}(R_\phi^{-1}\partial_\phi w^0 + u^0 R_\phi^{-1})\partial_\phi R_0 + \\ & + \partial_\omega \gamma^0_{\omega\zeta} + \gamma^0_{\phi\zeta} R_\phi^{-1}\partial_\phi R_0 \end{aligned} \tag{12R}$$

Equations of Motion.

$$\begin{aligned} & \partial_\phi (R_0 N_\phi) + R_\phi \partial_\omega N_{\omega\phi} - R_\phi N_\omega \cos\phi - R_0 \tilde{Q}_\phi \\ & = R_\phi R_0 \{ (M - 2R_\phi^{-1}P + R_\phi^{-2}J)\ddot{u}^0 + \\ & + [(P + P^a) - R_\phi^{-1}(J + J^a)]\ddot{\gamma}^0_{\phi\zeta} + (P^c - R_\phi^{-1}J^c)\ddot{\gamma}^0_{\omega\zeta} - \\ & - R_\phi^{-1}(P - R_\phi^{-1}J)\partial_\phi \ddot{w}^0 \} \end{aligned} \tag{26R}$$

$$\begin{aligned} & \partial_\phi (R_0 N_{\phi\omega}) + R_\phi \partial_\omega N_\omega + R_\phi N_{\omega\phi} \cos\phi - R_0 R_\phi R_\omega^{-1} \tilde{Q}_\omega \\ & = R_\phi R_0 \{ (M - 2R_\omega^{-1}P + R_\omega^{-2}J)\ddot{v}^0 + \\ & + (P^d - R_\omega^{-1}J^d)\ddot{\gamma}^0_{\phi\zeta} + [(P + P^b) - R_\omega^{-1}(J + J^b)]\ddot{\gamma}^0_{\omega\zeta} - \\ & - R_0^{-1}(P - R_\omega^{-1}J)\partial_\omega \ddot{w}^0 \} \end{aligned} \tag{27R}$$

$$\begin{aligned} & \partial_\phi (R_0 \tilde{Q}_\phi) + R_\phi \partial_\omega \tilde{Q}_\omega + R_0 R_\phi (R_\phi^{-1} N_\phi + R_\omega^{-1} N_\omega) \\ & + R_0 R_\phi \bar{q}_\zeta = R_\phi R_0 M \ddot{w}^0 + \\ & + \partial_\phi [R_0 (P - R_\phi^{-1}J)\ddot{u}^0 + R_0 (J + J^a)\ddot{\gamma}^0_{\phi\zeta} + R_0 J^c \ddot{\gamma}^0_{\omega\zeta} - \\ & - R_\phi^{-1} R_0 J \partial_\phi \ddot{w}^0] + \partial_\omega [R_\phi (P - R_\omega^{-1}J)\ddot{v}^0 + R_\phi J^d \ddot{\gamma}^0_{\phi\zeta} + \\ & + R_\phi (J + J^d)\ddot{\gamma}^0_{\omega\zeta} - R_\phi R_0^{-1} J \partial_\omega \ddot{w}^0] \end{aligned} \tag{28R}$$

$$\begin{aligned} & (\tilde{Q}_\phi - Q_\phi) + (\tilde{Q}^a_\phi - Q^a_\phi) + (\tilde{Q}^d_\omega - Q^d_\omega) \\ & = [(P + P^a) - R_\phi^{-1}(J + J^a)]\ddot{u}^0 + \\ & + (P^d - R_\omega^{-1}J^d)\ddot{v}^0 + (J + 2J^a + J^{aa} + J^{dd})\ddot{\gamma}^0_{\phi\zeta} \\ & + (J^c + J^d + J^{ac} + J^{bd})\ddot{\gamma}^0_{\omega\zeta} - \\ & - R_\phi^{-1}(J + J^a)\partial_\phi \ddot{w}^0 - R_0^{-1} J^d \partial_\omega \ddot{w}^0 \end{aligned} \tag{29R}$$

$$\begin{aligned} & (\tilde{Q}_\omega - Q_\omega) + (\tilde{Q}^b_\omega - Q^b_\omega) + (\tilde{Q}^c_\phi - Q^c_\phi) = (P^c - R_\phi^{-1}J^c)\ddot{u}^0 + \\ & + [(P + P^b) - R_\omega^{-1}(J + J^b)]\ddot{v}^0 + (J^c + J^d + J^{ac} + J^{bd})\ddot{\gamma}^0_{\phi\zeta} + \\ & + (J + 2J^b + J^{bb} + J^{cc})\ddot{\gamma}^0_{\omega\zeta} - R_\phi^{-1} J^c \partial_\phi \ddot{w}^0 \\ & - R_0^{-1}(J + J^b)\partial_\omega \ddot{w}^0 \end{aligned} \tag{30R}$$

where

$$q_\zeta = \bar{p}_\zeta(0) + \bar{p}_\zeta(h)$$

$$R_0 R_\phi \tilde{Q}_\phi = \partial_\phi (R_0 M_\phi) + R_\phi \partial_\omega M_{\omega\phi} - M_\omega \partial_\phi R_0$$

$$R_0 R_\phi \tilde{Q}_\omega = R_\phi \partial_\omega M_\omega + \partial_\phi (R_0 M_{\phi\omega}) + M_{\omega\phi}\partial_\phi R_0$$

6 Flat Plates

It is not difficult to realize that the previous relations contain, as a particular case, the linear equations of motion for multilayered anisotropic plates. For this purpose, let us to choose

$$x = \alpha;\ y = \beta;\ z = \zeta$$

with the plate reference surface belonging to the plane (x, y). Then,

$$R_\alpha = R_\beta = \infty;\ H_\alpha = H_\beta = 1;\ A = B = 1$$

and equations (26) to (38) yield

$$\partial_x N_x + \partial_y N_{xy} = M\ddot{u}^0 + (P + P^a)\ddot{\gamma}^0_{xz} + P^c \ddot{\gamma}^0_{yz} - P\partial_x \ddot{w}^0 \tag{26P}$$

$$\partial_x N_{xy} + \partial_y N_y = M\ddot{v}^0 + P^d \ddot{\gamma}^0_{xz} + (P + P^b)\ddot{\gamma}^0_{yz} - P\partial_y \ddot{w}^0 \tag{27P}$$

$$\begin{aligned} & \partial_x \tilde{Q}_x + \partial_y \tilde{Q}_y + \bar{q}_z = M\ddot{w}^0 + P\partial_x \ddot{u}^0 + P\partial_y \ddot{v}^0 \\ & + (J + J^a)\partial_x \ddot{\gamma}^0_{xz} + J^d \partial_y \ddot{\gamma}^0_{xz} + \\ & + J^c \partial_x \ddot{\gamma}^0_{yz} + (J + J^d)\partial_y \ddot{\gamma}^0_{yz} - J\partial_{xx}\ddot{w}^0 - J\partial_{yy}\ddot{w}^0 \end{aligned} \tag{28P}$$

$$(\tilde{Q}_x - Q_x) + (\tilde{Q}_x^a - Q_x^a) + (\tilde{Q}_y^d - Q_y^d) = (P + P^a)\ddot{u}^0 + P^d\ddot{v}^0 + (J + 2J^a + J^{aa} + J^{dd})\ddot{\gamma}_{xz}^0 + (J^c + J^d + J^{ac} + J^{bd})\ddot{\gamma}_{yz}^0 - (J + J^a)\partial_x\ddot{w}^0 - J^d\partial_y\ddot{w}^0 \quad (29P)$$

$$(\tilde{Q}_y - Q_y) + (\tilde{Q}_y^b - Q_y^b) + (\tilde{Q}_x^c - Q_x^c) = P^c\ddot{u}^0 + (P + P^b)\ddot{v}^0 + (J^c + J^d + J^{ac} + J^{bd})\ddot{\gamma}_{xz}^0 + (J + 2J^b + J^{bb} + J^{cc})\ddot{\gamma}_{yz}^0 - J^c\partial_x\ddot{w}^0 - (J + J^b)\partial_y\ddot{w}^0 \quad (30P)$$

where

$$\tilde{Q}_x = \partial_x\tilde{M}_x + \partial_y\tilde{M}_{xy} \text{ and } \tilde{Q}_y = \partial_y\tilde{M}_y + \partial_{xy}\tilde{M}_x$$

These equations are exactly those already derived by Di Sciuva (1984b).

7 Concluding Remarks

In this paper, a system of linear partial differential equations governing the motion of moderately thick multilayered anisotropic shells and plates has been derived.

The reduction of the three-dimensional problem to the bi-dimensional one is accomplished assuming a displacement field which allows piecewise linear variation of the u and v displacement (thus allowing for the distortion of the deformed normal to the reference surface), and constant value of the w displacement. The assumed displacement field also allows the contact conditions at the interfaces to be satisfied simultaneously and ab initio. No other simplifying assumptions, such as that regarding the smallness of the ratio h/R (Love's first approximation) have been adopted. The shear and rotatory inertia terms are also included in the derivation.

On the basis of the numerical results obtained in studying multilayered anisotropic plates (Di Sciuva, 1984a–b; 1986), it is hoped that the present theory should give better results than the conventional shear deformation theory and do not require the introduction of the shear correction factors. Moreover, the present approach may be employed to develop refined shell finite elements via the finite element displacement method (Di Sciuva, 1985a–b), which has been found to be not feasible by the conventional shear deformation approach. Here, for refined shell finite elements we mean elements constructed on the basis of the formulations which take into account the distortion of the deformed normal to the reference surface.

References

Ambartsumyan, S. A., 1964, "Theory of Anisotropic Shells," NASA TTF-118.

Ambartsumyan, S. A., 1969, "Theory of Anisotropic Plates," Translated from Russian by Cheron and Ed. by J. E. Ashton, Technomic Publishing Co.

Bhimaraddi, A., 1985, "Dynamic Response of Orthotropic, Homogeneous, and Laminated Cylindrical Shells," *AIAA J.*, Vol. 27, No. 11, pp. 1834–1837.

Di Sciuva, M., 1984a, "A Refinement of the Transverse Shear Deformation Theory for Multilayered Orthotropic Plates," *L'Aerotecnica Missili e Spazio*, Vol. 63, No. 2, pp. 84–91.

Di Sciuva, M., 1984b, "A Refined Transverse Shear Deformation Theory for Multilayered Anisotropic Plates," *Atti Accademia delle Scienze di Torino*, Vol. 118, pp. 279–295.

Di Sciuva, M., 1985a, "Development of an Anisotropic, Multilayered, Shear-Deformable Rectangular Plate Element," *Computer and Struct.*, Vol. 21, No. 4, pp. 789–796.

Di Sciuva, M., 1985b, "Evaluation of Some Multilayered, Shear-Deformable Plate Elements," *Proceedings of the 26th Structures, Structural Dynamics and Materials Conference*, AIAA/ASME/ASCE/AHS-Paper 85-0717, pp. 394–400.

Di Sciuva, M., 1986, "Bending, Vibration and Buckling of Simply Supported Thick Multilayered Orthotropic Plates: An Evaluation of a New Displacement Model," *J. Sound and Vib.*, Vol. 105, No. 3, pp. 425–442.

Dong, S. B., Pister, K. S., and Taylor, R. L., 1962, "On the Theory of Laminated Anisotropic Shells and Plates," *J. Aerospace Sciences*, Vol. 29, pp. 969–975.

Dong, G. B., and Tso, F. K. W., 1972, "On a Laminated Orthotropic Shell Theory Including Transverse Shear Deformation," JOURNAL OF APPLIED MECHANICS, Vol. 39, 1091–1097.

Hsu, T. M., and Wang, J. T. S., 1970, "A Theory of Laminated Cylindrical Shells Consisting of Layers of Orthoropic Laminae," *AIAA J.*, Vol. 8, No. 12, pp. 2141–2146.

Librescu, L., 1975, *Elastostatics and Kinetics of Anisotropic and Heterogeneous Shell-Type Structures*, Nordhoff International, Leyden.

Lo, K. H., Christensen, R. M., and Wu, E. M., 1977, "A High-Order Theory of Plate Deformation, Part 2: Laminated Plates," JOURNAL OF APPLIED MECHANICS, Vol. 44, No. 4, pp. 669–676.

Pagano, N. J., 1969, "Exact Solutions for Composite Laminate in Cylindrical Bending," *J. Comp. Mater.*, Vol. 3, pp. 398–411.

Pagano, N. J., 1970, "Exact Solutions for Rectangular Bidirectional Composites and Sandwich Plates," *J. Comp. Mater.*, Vol. 4, pp. 20–34.

Rath, B. K., and Das, Y. C., 1973, "Vibration of Layered Shells," *J. Sound and Vib.*, Vol. 28, No. 4, pp. 737–757.

Reddy, J. N., 1984, "A Simple Higher-Order Theory for Laminated Composite Plates," JOURNAL OF APPLIED MECHANICS, Vol. 51, No. 4, pp. 745–752.

Reddy, J. N., and Liu, C. F., 1985, "A Higher-Order Shear Deformation Theory of Laminated Elastic Shells," *Int. J. Eng. Sci.*, Vol. 23, No. 3, pp. 319–330.

Reissner, E., and Stavsky, Y., 1961, "Bending and Stretching of Certain Types of Heterogeneous Aelotropic Elastic Plates," JOURNAL OF APPLIED MECHANICS, Vol. 28, pp. 402–408.

Soldatos, K. P., 1984, "A Comparison of Some Shell Theories Used for the Dynamic Analysis of Cross-Ply Laminated Circular Cylindrical Panels," *J. Sound and Vib.*, Vol. 97, No. 2, pp. 305–319.

Srinivas, S., 1973, "A Refined Analysis of Composite Laminates," *J. Sound and Vib.*, Vol. 30, No. 4, pp. 495–507.

Sun, C. T., and Whitney, J. M., 1973, "Theories for the Dynamic Response of Laminated Plates," *AIAA J.*, Vol. 11, No. 2, pp. 178–183.

Waltz, T. L., and Vinson, J. R., 1976, "Interlaminar Stresses in Laminated Cylindrical Shells of Composite Materials," *AIAA J.*, Vol. 14, No. 9, pp. 1213–1218.

Washizu, K., 1968, *Variational Methods in Elasticity and Plasticity*, Pergamon Press, N.Y.

Whitney, J. M., 1969, "The Effect of Transverse Shear Deformation on the Bending of Laminated Plates," *J. Comp. Mat.*, Vol. 3, pp. 534–547.

Whitney, J. M., and Pagano, N. J., 1970, "Shear Deformation in Heterogeneous Anisotropic Plates," JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 1031–1036.

Whitney, J. M., and Sun, C. T., 1973, "A Higher Order Theory for Extensional Motion of Laminated Composites," *J. Sound and Vib.*, Vol. 30, No. 1, pp. 85–97.

Whitney, J. M., and Sun, C. T., 1974, "A Refined Theory for Laminated Anisotropic Cylindrical Shells," JOURNAL OF APPLIED MECHANICS, Vol. 41, No. 2, pp. 471–476.

Zukas, J. A., and Vinson, J. R., 1971, "Laminated Transversely Isotropic Cylidnrical Shells," JOURNAL OF APPLIED MECHANICS, Vol. 38, pp. 400–407.

Elastic-Plastic Analysis of Pressurized Cylindrical Shells

G. N. Brooks
Assistant Professor,
Department of Engineering Science
and Mechanics,
University of Tennessee,
Knoxville, TN 37996-2030

Plasticity in shells is often contained near the ends of a segment where the bending stresses are significant. Outside of this local neighborhood the behavior is elastic. Thus, an axisymmetric shell can be divided along its axis into a purely elastic region away from an end and the local region where plasticity is present. The moment-curvature relation in the elastic-plastic region is calculated using the Tresca yield condition. Use of the Tresca yield condition greatly simplifies this derivation because the principal directions are known. This moment-curvature relationship is "exact" in the sense that only the standard assumptions of thin shell theory are made. The solutions of the elastic and plastic regions are matched at their intersection for an efficient numerical solution. The technique is used here to study the semi-infinite clamped cylindrical shell with an internal pressure loading.

Introduction

The typical pressure vessel is a thin-walled shell of revolution. It is well-known that the stress in such a vessel is essentially given by the "membrane" solution except in the local vicinity of stiffening rings, or any other geometric discontinuity, where additional "bending" stresses occur. A purely elastic design for a vessel made of a ductile material to be loaded once is overly conservative. Plastic limit analysis, such as detailed in Hodge (1959, 1963), is useful but does not give any information about residual stresses or about the strains present. For a long pressurized cylinder with an end ring stiffener, limit analysis predicts that collapse occurs when the membrane stress reaches the yield stress. However, in the linear elastic solution the stresses at a discontinuity may far exceed the membrane stress. Thus, moderately ductile materials may fail before the load predicted by limit analysis.

As a section of a shell becomes plastic, generally one of the surfaces yields first. Then, as the load is increased the plastic zone spreads through the thickness and the other surface may also begin to yield. To avoid the difficulty of accurately considering the elastic-plastic behavior through the thickness, early investigators, such as Hodge (1963), used approximations to yield surface in terms of shell stress resultants for rotationally symmetric shells. It was assumed that the entire shell section yields simultaneously, producing an elastic-perfectly plastic relationship between the shell strain measures, the curvatures and midsurface strains, and the moment and force resultants. This is sometime referred to as the "sandwich" shell approximation since it is similar to the behavior of an idealized shell made of two thin layers of material separated by a constant distance.

The yield condition is then a function of the stress resultants of the shell. This is much simpler than considering the state of stress at each material point through the thickness. Hodge also developed the two-moment limited interaction yield condition by neglecting the interaction between orthogonal forces and between orthogonal moments. This approximate yield condition has been used to study plate (Lance and Onat, 1962) and shell (Leckie, 1965) problems.

The objective of this investigation is to develop an efficient method of elastic-plastic analysis of cylindrical shells without resorting to approximations of the yield condition in terms of the stress resultants. The method is then applied to the problem of the semi-infinite clamped cylinder loaded by internal pressure as illustrated in Fig. 1. The results can be applied to many pressure vessels and tanks with fixed ends.

To obtain "exact" relationships between the strain measures and the stress resultants, the elastic-plastic constitutive relations valid for the material at a point must be integrated through the thickness incorporating the usual assumptions of shell theory. The Tresca yield condition is used here because for axisymmetric shells the principal directions are known and remain constant. It is shown in a following section that for plane stress, with some modest restrictions, the current state of stress, can be found as a function of the current applied strains. Assuming that the shell is in a state of plane stress and that the Kirchhoff-Love hypothesis is correct, this solution is integrated through the thickness to give "constitutive" relationships between the strain measures and stress resultants.

These elastic-plastic constitutive relations are combined with the well-known equilibrium and strain-displacement equations, valid for small deflections, to give the complete description of the elastic-plastic pressurized cylinder.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 10, 1986; final revision, December 27, 1986.

Paper No. 87-WA/APM-11.

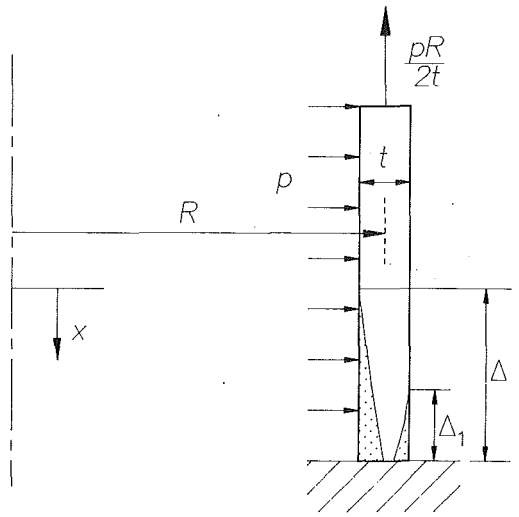


Fig. 1 Geometry of a clamped-end, pressurized semi-infinite cylindrical shell. Δ_1 is the extent of plasticity, if any, on the outside, while Δ is the length of the plastic zone.

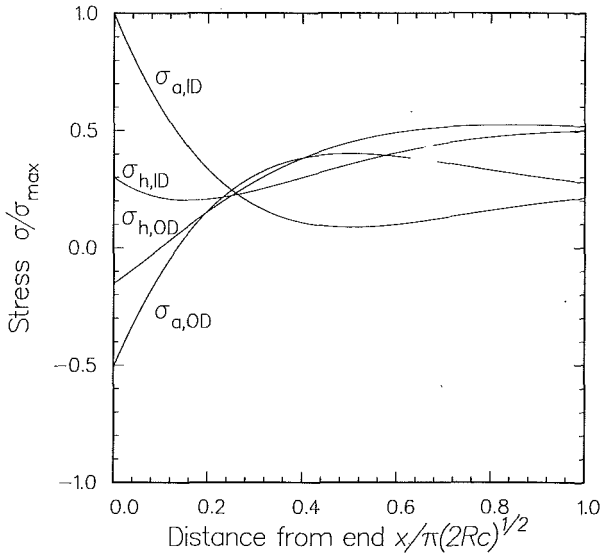


Fig. 2 Axial and hoop stresses on the inner and outer diameter of an elastic, clamped-end pressurized cylinder

Elastic Solution

As a preliminary to the elastic-plastic analysis, the linear solution for the problem of Fig. 1 is presented here. It will be used to determine where yielding first occurs. It is also needed for the elastic-plastic solution where the cylinder is conceptually divided into elastic and plastic regions.

The linearized form of Reissner's (1950) equations for the axisymmetric deformation of shells of revolution are given in matrix form by Steele and Skogh (1970). These are modified here for the cylinder in nondimensional form. Axial position is $x[2Rc]^{1/2}$ where x is dimensionless, the terms in the bracket give it the proper dimensional value, and $c = t/(12(1 - \nu^2))^{1/2}$. The dependent variables are the axial bending moment resultant $M_a[Yt^2/4]$, the transverse shear $H[Yt(6(1 - \nu^2)c/R)^{1/2}/4]$, the rotation of the shell midsurface $\chi[Y(3R(1 - \nu^2)/2c)^{1/2}/2E]$, and the radial displacement $h[YR(3(1 - \nu^2))^{1/2}/2E]$, where Y is the yield stress. Also in the equation is the pressure $p[Yt/R]$ and the axial force resultant $N_a[Yt]$. The advantage of this form of the equations is that each of the variables that may be prescribed as a boundary condition is present in the vector unknown.

$$\frac{d}{dx}\begin{bmatrix} M_a \\ H \\ \chi \\ h \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}\begin{bmatrix} M_a \\ H \\ \chi \\ h \end{bmatrix} + \frac{4(\nu N_a - p)}{\left[3(1-\nu^2)\right]^{1/2}}\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{1}$$

For internal pressure loading of a closed-end cylinder, the axial stress resultant is found from static equilibrium to be

$$N_a = \frac{p}{2} \tag{2}$$

the solution of equation (1) which satisfies clamped boundary conditions at $x = 0$ is

$$\begin{bmatrix} M_a \\ H \\ \chi \\ h \end{bmatrix} = \frac{(2-\nu)p}{\left[3(1-\nu^2)\right]^{1/2}}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \text{Real}\ \frac{(2-\nu)p}{\left[3(1-\nu^2)\right]^{1/2}}\begin{bmatrix} 1-i \\ 2 \\ 2i \\ 1+i \end{bmatrix} e^{(1+i)x} \tag{3}$$

When $x = -\pi$, the exponential in the complementary solution is only four percent of its value at the edge and thus is essentially insignificant. This corresponds to a physical distance of $\pi(2Rc)^{1/2}$ which will be referred to as the "elastic decay distance" hereafter.

The stresses on the inner and outer diameter are plotted in Fig. 2. It is seen that the maximum stress is the axial stress on the inner diameter at the clamped edge. Since the hoop and axial stresses have the same sign, yielding will initiate when $\sigma_a = Y$ on the inner diameter. The pressure at which this occurs is

$$p_{el} = \frac{2}{1 + (2-\nu)\left[3/(1-\nu^2)\right]^{1/2}} \tag{4}$$

The elastic solution suggests that as the pressure is increased, the outer diameter will eventually yield under the condition $\sigma_a = -Y$. This will be verified by the elastic-plastic analysis.

Elastic-Plastic Moment-Curvature Relationships

Use of the Tresca yield condition is convenient for the study of axisymmetric elastic-plastic shells because for those problems the principal stress directions are known and unchanging. The Tresca yield surface for plane stress is shown in Fig. 3. Assuming that the yielding material at a given point in the shell stays on the same side of the yield surface, the plane stress problem can be integrated to give current stresses as a function of the current strains. This solution, combined with the standard geometric assumptions of small deformation shell theory, was then integrated through the thickness of the shell to give the elastic-plastic moment-curvature relationships.

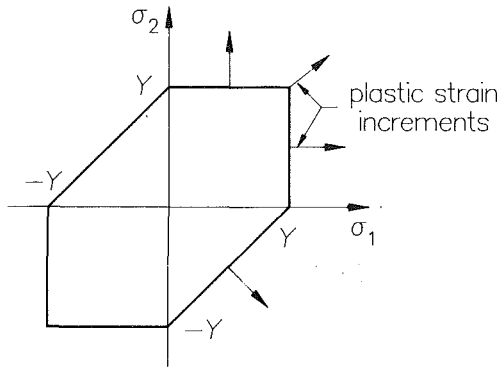


Fig. 3 Tresca yield surface showing the directions of plastic strain increments at several positions

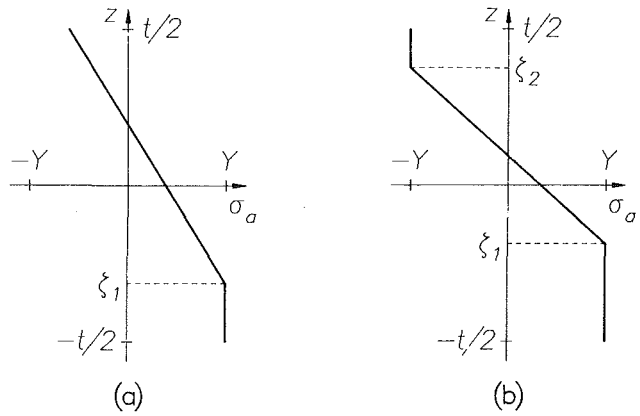


Fig. 4 Distribution of axial stress through the thickness when: (*a*) Only the inside is yielding; and (*b*) both sides are yielding

Plane Stress Solutions. Consider the problem where all yielding occurs under the condition $\sigma_1 = Y$, with $0 < \sigma_2 < Y$. Let e_{1c} and e_{2c} be the total current principal strains. For a loading history of the virgin material, the principal strains may be taken as the following functions of τ, which grows monotonically with time.

$$\begin{aligned} e_1 &= f(\tau)\ e_{1c} \\ e_2 &= g(\tau)\ e_{2c} \end{aligned} \tag{5}$$

If at some time the strains decrease, the material would be back in the elastic regime. This would not affect future yielding which is assumed to take place on the same side of the yield surface. Thus, it may be assumed with no loss of generality that f and g are continuous nondecreasing functions with

$$f(0) = g(0) = 0 \tag{6}$$

$$f(1) = g(1) = 1 \tag{7}$$

Yielding will initiate when $\sigma_1 = Y$, the yields stress, at some time, τ_y. Let $f(\tau_y) = f_y$ and $g(\tau_y) = g_y$. The elastic stress-strain law for plane stress (Timoshenko, 1970) gives

$$\sigma_1(\tau_y) = Y = \frac{E}{(1-\nu^2)}\left(f_y e_{1c} + \nu g_y e_{2c}\right) \tag{8}$$

$$\sigma_2(\tau_y) = \frac{E}{(1-\nu^2)}\left(g_y e_{2c} + \nu f_y e_{1c}\right) \tag{9}$$

The associated flow rule for the Tresca yield condition (Kachanov, 1974) gives the plastic strain increments during plastic flow as

$$\begin{aligned} de_1^p &= \ \ de_p \\ de_2^p &= \ \ 0 \\ de_3^p &= -de_p \end{aligned} \tag{10}$$

Expressing the total strain increments as the sum of the elastic and plastic parts

$$\begin{aligned} de_1 &= de_1^e + de_1^p = de_1^e + de_p \\ de_2 &= de_2^e \end{aligned} \tag{11}$$

The incremental stress-strain relations give

$$d\sigma_1 = \frac{E}{(1-\nu^2)}\left[de_1^e + \nu\ de_2^e\right] \tag{12}$$

$$= \frac{E}{(1-\nu^2)}\left[de_1 - de_p + \nu\ de_2\right]$$

$$d\sigma_2 = \frac{E}{(1-\nu^2)}\left[de_2 + \nu\left(de_1 - de_p\right)\right] \tag{13}$$

For perfectly plasticity, $d\sigma_1 = dY = 0$, thus the plastic strain increment can be found in terms of the two applied strains

$$de_p = de_1 + \nu de_2 = \left[f'(\tau)e_{1c} + \nu g'(\tau)e_{2c}\right]d\tau$$

Integrating this from the initial yield at τ_y to the current state at $\tau = 1$, and combining with equation (8) gives

$$\begin{aligned} e_p &= (1-f_y)e_{1c} + \nu(1-g_y)e_{2c} \\ &= e_{1c} + \nu e_{2c} - (1-\nu^2)\frac{Y}{E} \end{aligned} \tag{14}$$

Now the current stresses may be found as a function of the current total applied strains. Integrating equations (12), (13), with equations (8), (9) as initial values gives

$$\begin{aligned} \sigma_1 &= Y \\ \sigma_2 &= \nu Y + E e_{2c} \end{aligned} \tag{15}$$

For the case when all yielding is on the opposite side of the yield surface by $\sigma_1 = -Y$, the analogous relations are

$$\begin{aligned} \sigma_1 &= -Y \\ \sigma_2 &= -\nu Y + E e_{2c} \end{aligned} \tag{16}$$

When the yielding is contained on the side of the yield surface defined by $\sigma_1 - \sigma_2 = Y$, it can be shown (Brooks, 1982) that the stresses are

$$\begin{aligned} \sigma_1 &= \frac{E}{2(1-\nu)}\left(e_{1c} + e_{2c}\right) + \frac{Y}{2} \\ \sigma_2 &= \frac{E}{2(1-\nu)}\left(e_{1c} + e_{2c}\right) - \frac{Y}{2} \end{aligned} \tag{17}$$

Integration into Moment-Curvature Relations. The elastic solution already given for the clamped semi-infinite cylinder with internal pressure is well known (Kraus, 1967; Siede, 1975). Yielding begins on the inner diameter when the pressure is high enough to cause the axial stress to reach the yield stress. It was initially assumed, and later verified by the analysis, that as the pressure increases, yielding throughout the plastic zone for material near the inner diameter is with $\sigma_a = Y$ and that when the yielding begins on the outer surface it is under the condition $\sigma_a = -Y$.

The normalized moment, M, and axial force, N, are defined such that $M = \pm 1$ is the ultimate moment the cross section can carry if there is no axial force (i.e., $N = 0$) and $N = 1$ is the ultimate tensile load that can be applied with no moment

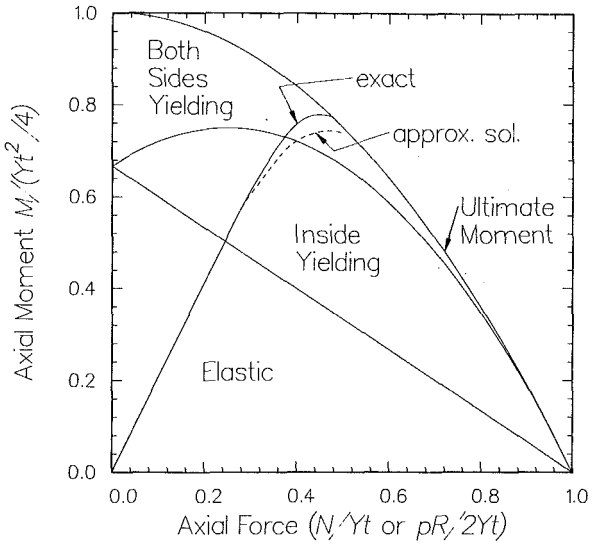


Fig. 5 The moment at the clamped end versus internal pressure. Also shown are boundaries between purely elastic, inside yielding, and both sides yielding regions.

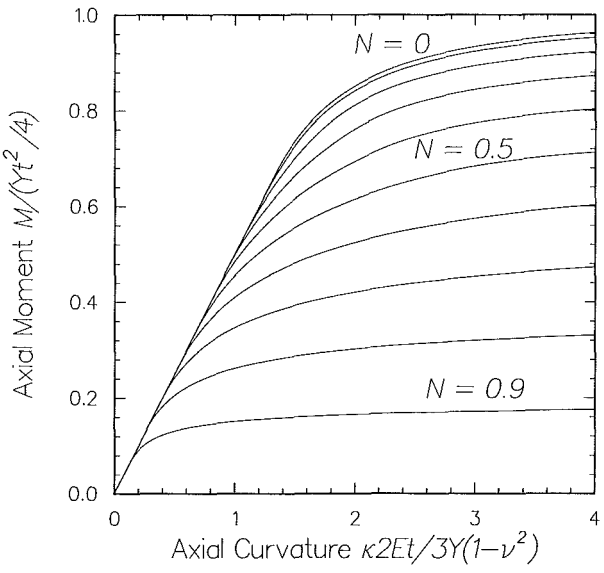


Fig. 6 Axial moment-curvature relation with constant axial force for several values of axial force

loading. Figure 4 shows the axial stress distribution through the shell thickness for loads above yield. There are three distributions of special interest. The combination of moment and force that causes each can be determined solely from consideration of the equilibrium of the cross section. For a given axial force some moment will cause the inside to just reach yield, as shown in Fig. 4(a) with $\zeta_1 = -t/2$.

$$M = -\frac{2}{3}(1-N) \tag{18}$$

As the moment is increased the outside will reach yield, as shown in Fig. 4(b) with $\zeta_2 = t/2$, when

$$M = -\frac{2}{3}(1-N)(1+2N) \tag{19}$$

Finally, the entire section becomes plastic at the ultimate moment as in Fig. 4(b) with $\zeta_1 = \zeta_2$.

$$M = N^2 - 1 \tag{20}$$

Equations (18) through (20) are plotted in Fig. 5 dividing the $N-M$ space into elastic, inside yielding, and both sides yielding regions. This figure is valid as long as there is no unloading of the yielding material and all of the yielding is under the condition $\sigma_a = \pm Y$.

The axial stress profile through the thickness after only the inside has yielded is shown in Fig. 4(a). The strains are still assumed to vary linearly through the thickness. For given midsurface strains and axial curvature (the hoop curvature is zero for cylinders), the strains are known everywhere. Integrating the stress-strain law, either the elastic law or the appropriate choice from equations (15) or (16), through the thickness gives the axial moment and force. The equations can be manipulated to give the axial curvature $\kappa[3Y(1-\nu^2)/2Et]$ in terms of the moment and force.

$$\kappa = \frac{-64(1-N_a)}{27}\left[2+\frac{M_a}{1-N_a}\right]^{-2} \tag{21}$$

Figure 4(b) shows the stress profile after yielding has initiated on both surfaces of the shell. The moment-curvature relationship for this case is

$$\kappa = \frac{-4}{3\sqrt{3}}\left[1-N_a^2+M_a\right]^{-1/2} \tag{22}$$

In each case the hoop stress resultant is

$$N_h = \nu N_a + \epsilon_h \tag{23}$$

The moment-curvature relationship is plotted in Fig. 6 with N_a varied as a parameter, but held constant for each curve.

Solution

In the elastic solution, the stresses peak sharply at the clamped end, so it was expected that the plasticity would be localized near the end. Following the approach of Steele (1968), the cylinder is divided into a semi-infinite region that is completely elastic and the region near the end where there has been plastic flow. Each section is analyzed independently and the displacements and stress resultants are matched at the intersection.

The axial stress resultant is known from static equilibrium (2). The moment at the intersection of the two regions is then calculated from equation (18) with the assumption that yielding at all axial positions starts on the inner diameter under the condition $\sigma_a = Y$. This assumption was confirmed by the analysis. The unknown quantities are Δ, the length of the plastic zone, and h_o, χ_o, and H_o, the radial deflection, rotation, and transverse shear at the end of the plastic zone.

Approximate Solution. The equilibrium equations for the cylindrical shell are

$$\frac{dH}{dx} = \frac{4}{\left[3(1-\nu^2)\right]^{1/2}}(N_h - p) \tag{24}$$

$$\frac{dM_a}{dx} = H \tag{25}$$

The analysis of the elastic-plastic problem can be simplfied by dropping the variation of transverse shear from the equilibrium equations in the plastic zone. As long as the plastic zone is small (as it must be, at least at the outset) this term should have a small effect. The moment in the plastic zone then is just a function of the moment at the intersection, M_o, which is known from equations (2) and (18), and the unknown transverse shear, H_o, given by

$$M_a(x) = M_o + H_o x \tag{26}$$

Initially all of the yielding in the plastic zone is on the inner diameter, so the moment-curvature relationship of equation (21) can be used throughout the plastic zone. Substituting

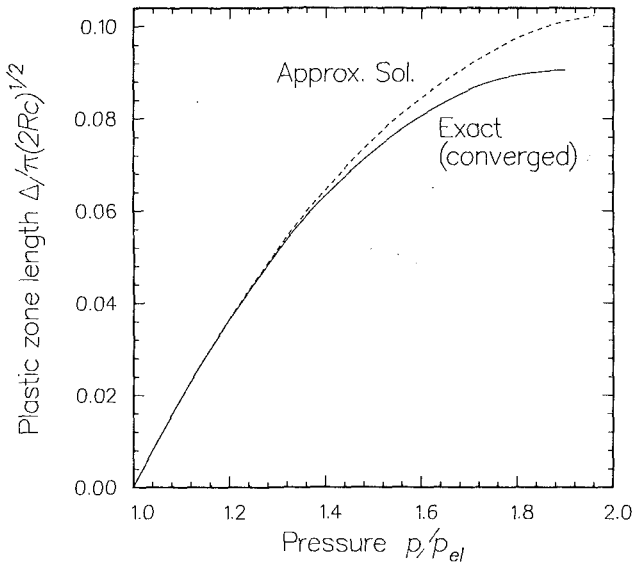


Fig. 7 Growth of the plastic zone length with pressure

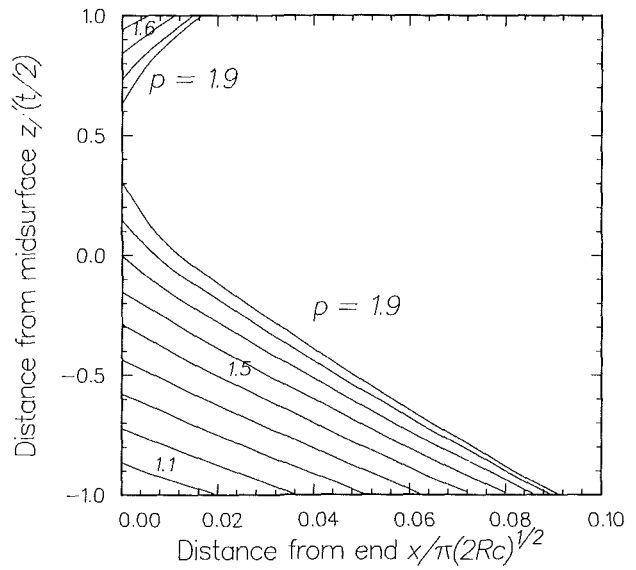


Fig. 8 Boundaries between the plastic and elastic regions through the thickness for several values of pressure

equation (26) into equation (21) gives the curvature in the plastic zone.

$$\kappa = \frac{-4(1-N_a)}{3}\left[1+\frac{3H_o x}{4(1-N_a)}\right]^{-2} \tag{27}$$

This equation may be integrated through the plastic zone, applying the clamped boundary conditions at the end, to give the radial deflection and the rotation at the intersection of the two regions.

$$\chi_0 = \frac{16(1-N)^2}{9H_0}\left[\frac{3H_0\Delta}{3H_0\Delta+4(1-N)}\right] \tag{28}$$

$$h_0 = \frac{64(1-N)^3}{27H_0^2}\left[\ell n\left[1+\frac{3H_0\Delta}{4(1-N)}\right]-\frac{3H_0\Delta}{3H_0\Delta+4(1-N)}\right] \tag{29}$$

At higher pressures, the outer diameter will also yield near the clamped end. The moment, M_2, required to cause the outer edge to yield at some given pressure level is given by equations (2) and (19). From equation (26), the position where this occurs is

$$\Delta_1 = \frac{M_2 - M_0}{H_0} \tag{30}$$

Substituting equation (26) into the moment-curvature equation (22) gives the curvature near the clamped end.

$$\kappa = \frac{-4}{3}\left[1+2N-3N^2+3H_0 x\right]^{-1/2} \quad \text{for } \Delta_1 < x < \Delta \tag{31}$$

For $0 < x < \Delta_1$, equation (27) is still valid. The rotation and deflection at the end of the plastic zone are

$$\chi_0 = \frac{16(1-N)^2}{2H_0}\left[\frac{3H_0\Delta_1}{3H_0\Delta_1+4(1-N)}\right]-\frac{8}{9H_0}\left[\Gamma_1-\Gamma\right] \tag{32}$$

$$h_0 = \frac{64(1-N)^3}{27H_0^2}\left[\ell n\left[1+\frac{3H_0\Delta_1}{4(1-N)}\right]-\frac{3H_0\Delta_1}{3H_0\Delta_1+4(1-N)}\right] + \frac{8}{9H_0}\left[\Delta\Gamma-\Delta_1\Gamma_1\right]+\frac{16}{81H_0^2}\left[\Gamma_1^3-\Gamma^3\right] \tag{33}$$

where

$$\Gamma = \left[1+2N-3N^2+3H_0\Delta\right]^{1/2}$$

$$\Gamma_1 = \left[1+2N-3N^2+3H_0\Delta_1\right]^{1/2} \tag{34}$$

The solution for the elastic region is obtained (using the complete equilibrium equations) with the boundary conditions that $M = M_0$ and $H = H_0$ at the intersection. This gives the rotation and deflection at the intersection as

$$\chi_0 = 2M_0 - H_0 \tag{35}$$

$$h_0 = -M_0 + H_0 + \frac{(2-\nu)p}{\left[3(1-\nu^2)\right]^{1/2}} \tag{36}$$

Setting the above equations equal to either equations (28)–(29) or (32)–(33), as appropriate, gives a pair of coupled nonlinear equations in the unknowns Δ and H_0. These equations were then solved by iteration with starting values obtained from the elastic solution.

Numerical Solution of Exact Equations. When the exact equilibrium equations were used it was not possible to integrate analytically the moment-curvature relationship through the plastic zone. Therefore, the equations were numerically integrated with assumed initial values at the end of the plastic zone and the results compared with the boundary conditions at the clamped end. The moment at the intersection of the elastic and plastic zones is known from equation (18). The method of solution is to assume this moment and an estimate of h_0 exist at the intersection of the regions. The elastic solution then gives H_0 and χ_0 by equations (35), (36). These are used as the initial values in the elastic-plastic equations which are then integrated until $\chi = 0$. If $h = 0$ at this point, then the problem is solved. Otherwise the initial estimate of h_0 is changed and the process is repeated until convergence.

Results

The results of the computations for both the approximate and the exact equations are plotted in Figs. 5 and 7 through 10. All results are for $\nu = 0.3$. It is seen that the two solutions agree well. As expected, the agreement is best during the early stages of yielding and there is some deviation at higher pressures.

The growth of the plastic zone as the pressure increases is given in Fig. 7. The maximum length is only about ten percent of the "elastic boundary layer." Thus, the plastic region is contained within a small distance from the clamped end as was initially expected from the elastic solution. This is why the approximate solution is fairly accurate even at relatively high pressures. Since the plastic zone is so short, the results in this paper may be applied to any cylinder which is long enough to be modeled as semi-infinite for the elastic solution.

Figure 8 shows the boundaries between the plastic and elastic parts of the cross section as if one made a cut through the thickness and marked the edge of the yielding area. The boundaries are shown for several values of pressure. The plastic area moves in from the clamped end and from the surfaces of the cylinder with increasing load. This verifies that there is no unloading in the plastic zone, which was an assumption of the analysis.

Figure 5 shows the variation of the moment at the clamped end (i.e., the maximum moment in the cylinder) with the pressure. Plotting the moment on this figure in this way allows one to see how close the section is to the ultimate load carrying capability when the calculations were stopped.

The moment in the approximate solution grows more slowly than in the exact solution. To see that this is reasonable, consider the term that was dropped from the equilibrium equation (24) in the approximate solution.

$$N_h - p = \nu N_a + \epsilon_h - p \simeq -p\left[1 - \frac{\nu}{2}\right] < 0$$

The hoop strain ϵ_h was taken as approximately zero in the above inequality because h and χ equal zero at the clamped end and the plastic zone is short. It would have to be quite large to change the sign of the dropped term. Thus, the dropped term increases the magnitude of the transverse shear, and thus that of the moment also, as one integrates away from the end of the elastic region.

The calculations were stopped just before the axial moment at the clamped end reached its maximum possible value. The axial resultant N_a is about one-half at this point. The hoop stress in the membrane solution away from the end is twice N_a or almost one. Since at $N = 1$ the whole cross section is yielding under pure tension, the pressure could not be increased further or the cylinder would burst in the membrane region, regardless of the state of stress at the clamped end.

Thus, even though there is a severe elastic stress concentration at the clamped end, the collapse pressure is that which causes yielding in the membrane solution. The limit analysis of Hodge (1963) gives the same result for the collapse pressure. Physically, this seems reasonable for a material that has no strain limit. The shell yields at the clamped end, but the deformation is constrained by the membrane region until yielding also occurs there. Actual failure could occur earlier if the bending strains at the end became larger than that permissible for the material.

The maximum strain in the cylinder is the axial strain on the inner surface at the clamped end. This is shown in Fig. 9 as a function of the applied pressure. At a pressure ninety-five percent above that which causes initial yield, the strain is 17.5 times the elastic limit strain (Y/E). Thus, many ductile materials would survive loading until general yielding in the membrane region. Because the maximum strain increases so slowly during early yielding, cylinders made of most common structural metals would survive pressures significantly greater than the elastic limit.

Figure 10 shows the stresses on the inner and outer surfaces of the cylinder at a pressure of 1.8 times the elastic limit pressure for the exact solution. The analogous curves for the approximate solution give essentially the same results (Brooks, 1982).

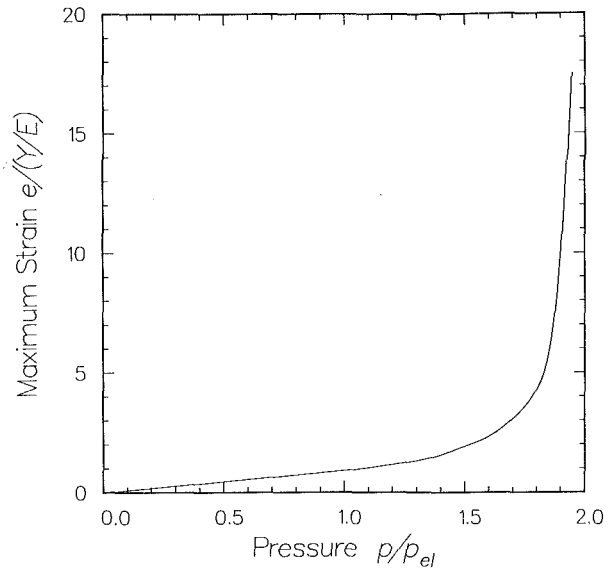


Fig. 9 Variation of the axial strain on the inner surface at the clamped end (maximum strain in cylinder) with pressure

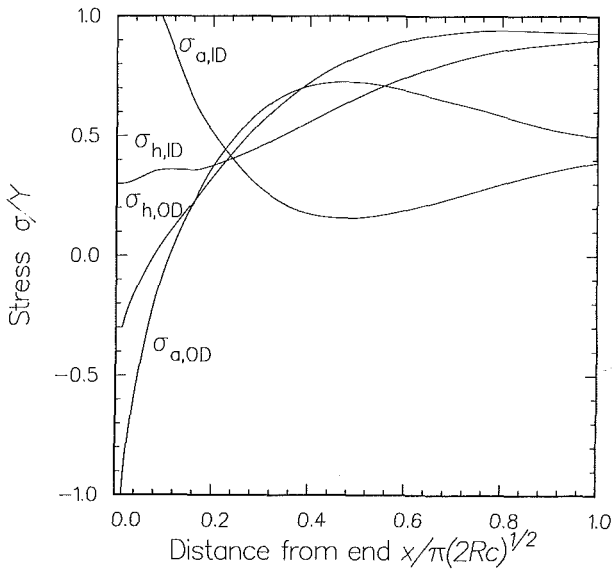


Fig. 10 Axial and hoop stress on the inner and outer diameters at a pressure of 1.8 times the elastic limit pressure

Conclusions

A very efficient method for elastic-plastic analysis of axisymmetric cylinders has been developed. Since the results are all given in dimensionless form they can be applied to any cylindrical shell.

The numerical solution of the equations can be considered exact, given the standard shell assumptions which were made. The method was so inexpensive that hundreds and even thousands of points were used in the numerical integration through the plastic zone. Because of this, these results could be useful to compare against other numerical methods. One complication would be that most general methods use the Mises yield condition rather than Tresca, as was used here.

An important assumption was that the plastic zone would remain small. This was based on the fact that the elastic "edge effect" solutions decay rapidly with distance. If the plastic zone becomes large, the method in its present form might be ineffective. This can be seen by considering the elastic complementary solution. It has four exponential solutions, two of

which decay and two which grow with axial distance. In an analytical solution the growing terms can be discarded, but when numerical integration is used all of the terms will be included and can be significant if the integration is over a large distance. The elastic-plastic solution could have similar problems, if the plastic zone were to become too large.

Although only cylinders are investigated here, it is not difficult to generalize the method to the general shell of revolution. The elastic solution for "steep" shells is similar in behavior to the cylinder. Thus, for shells where the plasticity occurs in the "steep" region the results should be very similar to those for the cylinder. The principal directions are still known so the Tresca yield condition can still easily be used. The only real complication is in the constitutive relations. The change of curvature in the circumferential direction, which is absent in the cylinder, has to be included.

There are several other items not considered here which could be important. Only the material nonlinearity is included, but geometric nonlinearity could be significant especially in the plastic zone where the curvature is changing rapidly. Shear deformation could also be important within this zone. Including the effect of material strain hardening would also be a logical extension of the present results.

Acknowledgments

The author would like to thank his Ph.D. advisor, Charles R. Steele, for his guidance during the time this research was being done. The support of the National Science Foundation, with a Graduate Fellowship while the author was a graduate student at Stanford University, and presently with grant MEA-8404055, has made this work possible.

References

Brooks, G. N., 1982, "Plastic Boundary Layers and Part-Through Cracks in Shells," Ph.D. Thesis, Stanford University.

Hodge, Philip G., Jr., 1963, *Limit Analysis of Rotationally Symmetric Plates and Shells*, Prentice-Hall, NJ.

Hodge, Philip G., 1959, Jr., *Plastic Analysis of Structures*, McGraw-Hill, NY.

Kachanov, L. M., 1974, *Fundamentals of the Theory of Plasticity*, Mir Publishers, Moscow.

Krauss, H., 1967, *Thin Elastic Shells*, Wiley, NY.

Lance, R. H., and Onat, E. T., 1962, "A Comparison of Experiments and Theory in the Plastic Bending of Circular Plates," *Journal of the Mechanics and Physics of Solids*, Vol. 10, pp. 301–311.

Leckie, F. A., 1965, "The Plastic Analysis of a Spherical Shell Subjected to a Radial Load Applied Through Rigid Boss," Brown University Report, NSF-GP-1115/23.

Reissner, E., 1950, "On Axisymmetric Deformations of Thin Shells of Revolution," *Proceedings of the Third Symposium on Applied Mathematics*, McGraw-Hill, NY, pp. 27–52.

Siede, P., 1975, *Small Elastic Deformations of Thin Shells*, Noordhoff International Publishing, Holland.

Steele, C. R., 1968, "Pressure Vessels with Edge Zone Yielding," *Int. J. Solids and Structures*, Vol. 4, pp. 777–785.

Steele, C. R., and Skogh, J., 1970, "Slope Discontinuities in Pressure Vessels," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, Vol. 92, Series E, pp. 587–595.

Timoshenko, S. P., 1970, *Theory of Elasticity*, McGraw-Hill, NY.

Analysis of Pipe Bends With Symmetrical Noncircular Cross Sections

J. F. Whatham
Australian Atomic Energy Commission,
Research Establishment,
Sutherland, NSW, 2232, Australia
Mem. ASME

A thin shell analysis is presented for pipe bends with symmetric noncircular cross sections under in-plane bending or internal pressure, using the static-geometric analogy and neglecting end effects. Any symmetric shape is possible but the analysis mainly concerns cross sections with double symmetry; an investigation of two-lobe (oval) and four-lobe cross sections demonstrates that pipe bend flexibility is almost inversely proportional to flank radius if the pipe wall is thin. Pressurizing a pipe bend of oval cross section produces a similar hoop stress distribution to that of a bending moment straightening the pipe.

Introduction

An exact thin shell analysis for circular pipe bends terminated by flanges or straight pipes and subjected to any end loading has been published (Whatham, 1986); we now consider the exact thin shell solution for pipe bends with symmetrical noncircular cross sections subjected to in-plane bending or internal pressure, neglecting end effects. Reference is made to the work of Clark et al. (1952) who obtained bending and pressurizing solutions for pipe bends with elliptical cross sections.

The first order linear shell theory of Novozhilov (1970) is employed but, as the loading generates no shear stresses, the equations revert to those of Love (1944). Use is made of Goldenweizer's static-geometric analogy and no approximations are made other than those inherent in the theory.

Governing equations are derived in terms of displacements and stress functions for in-plane bending and pressurization, although the pressurization is accompanied by an in-plane bending moment unless the cross section is circular; a reverse bending solution can be superimposed to obtain the effect of pressurization alone. The equations are solved by collocation at equal intervals around the pipe circumference and, as a test problem, the stresses are calculated in a typical pipe bend for different degrees of cross section ovality. Pipe bends with two-lobe (oval) and 4-lobe cross sections are then investigated for the relationship between flank radius and pipe bend flexibility.

Governing Equations

A segment of curved pipe with a symmetrical noncircular cross section is represented in Fig. 1 by its middle surface, acted upon by stress resultants and surface pressure to produce the displacements u, v, w, and a rotation ψ about the u axis; there are rotations about the other axes but these do not enter the analysis. The radius of curvature R is through the centroid of the cross section inside the pipe, and curvilinear coordinates for the middle surface are θ, η where

$$\theta = 2\pi s/c, \quad \eta = 2\pi \ell/c, \quad c = \text{mid-wall circumference}$$

From the Novozhilov equations we obtain:

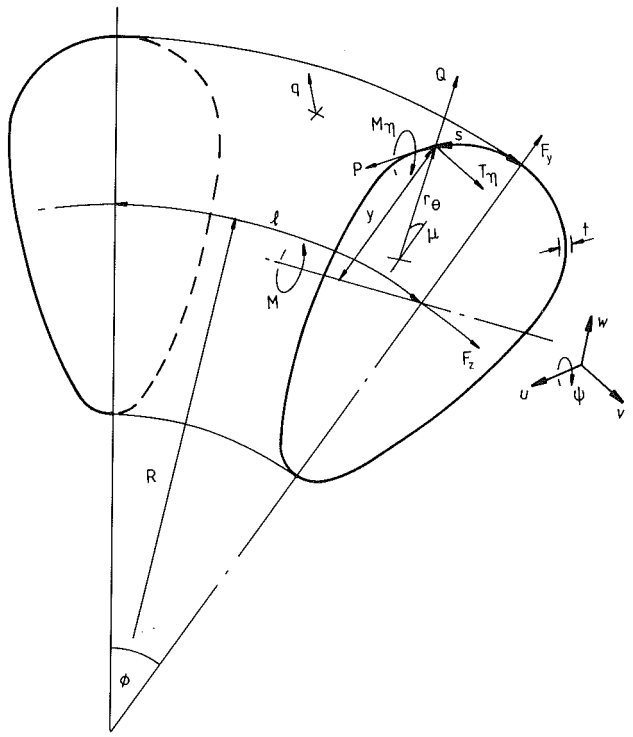


Fig. 1 Pipe bend middle surface

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August 5, 1986; final revision February 19, 1987.

(*i*) *Equilibrium Equations*

$$\frac{\partial}{\partial\theta}(\delta T_\theta^*)+g\frac{\partial}{\partial\theta}(\delta M_\theta^*)+\frac{\sin\mu}{\rho}(T_\eta^*+gM_\eta^*)+\frac{\partial P^*}{\partial\eta}=0$$

$$g\delta T_\theta^*-\frac{\partial^2}{\partial\theta^2}(\delta M_\theta^*)+\frac{\cos\mu}{\rho}T_\eta^*-\frac{\partial}{\partial\theta}\left(\frac{\sin\mu}{\rho}M_\eta^*\right)-\frac{\partial Q^*}{\partial\eta}=\delta q^* \quad (1)$$

where $T_\eta^*=T_\eta/Et$, $M_\eta^*=M_\eta/rEt$, $q^*=qr/Et$,

$$\rho=R/r,\delta=1+y/R,g=r/r_\theta=d\mu/d\theta,r=c/2\pi$$

(*ii*) *Stress-Strain Equations*

$$T_\theta^*=\frac{1}{1-\nu^2}(\epsilon_\theta+\nu\epsilon_\eta)\quad M_\theta^*=\frac{\gamma r}{12(1-\nu^2)}(\kappa_\theta+\nu\kappa_\eta)$$

$$T_\eta^*=\frac{1}{1-\nu^2}(\epsilon_\eta+\nu\epsilon_\theta)\quad M_\eta^*=\frac{\gamma r}{12(1-\nu^2)}(\kappa_\eta+\nu\kappa_\theta) \quad (2)$$

where $\gamma=(t/r)^2$

(*iii*) *Strain-Displacement Equations*

$$\epsilon_\theta=\frac{1}{r}\left(\frac{\partial u}{\partial\theta}+gw\right)\qquad \kappa_\theta=\frac{1}{r^2}\left(\frac{\partial}{\partial\theta}(gu)-\frac{\partial^2 w}{\partial\theta^2}\right) \quad (3)$$

$$\epsilon_\eta=\frac{1}{\delta r}\left(\frac{\partial v}{\partial\eta}-\frac{\sin\mu}{\rho}u+\frac{\cos\mu}{\rho}w\right)$$

$$\kappa_\eta=\frac{1}{\delta r}\left(\frac{\partial\psi}{\partial\eta}+\frac{\sin\mu}{\rho r}\left(\frac{\partial w}{\partial\theta}-gu\right)\right)$$

$$\lambda=\frac{g}{r}\frac{\partial v}{\partial\theta}-\frac{\partial\psi}{\partial\theta}\qquad \beta=\frac{1}{r}\frac{\partial^2 v}{\partial\theta^2}+g\psi$$

where $$\psi=\frac{1}{\delta r}\left(\frac{\cos\mu}{\rho}v-\frac{\partial w}{\partial\eta}\right)$$

Strains λ and β represent the twist and out-of-plane curvature given to the pipe wall if the pipe cross section warps, but for in-plane bending or pressurizing, without end effects, cross sections remain plane and these strains are zero, as are stress resultants P and Q. Incidentally, changes in curvature of the pipe wall are given by $\kappa_\eta-\epsilon_\eta/r_\eta$ and $\kappa_\theta-\epsilon_\theta/r_\theta$ rather than κ_η and κ_θ.

Now from Fig. 1

$$F_z=\int_o^c T_\eta ds$$

$$F_y=\int_o^c (Q\cos\mu-P\sin\mu)ds \quad (4)$$

$$M=-\int_o^c (M_\eta\cos\mu+yT_\eta)ds$$

whence by equations (1)

$$F_z=pA_i+\frac{\partial F_y}{\partial\phi} \quad (5)$$

where A_i is the inside cross section area and p is the pressure in the pipe.

$$A_i=\int_o^c y\cos\mu ds-ct/2+\pi t^2/4$$

$$q=p(1-t/2r_\theta)(1-t/2r_\eta) \quad (6)$$

$$r_\eta=\delta R/\cos\mu$$

Consider the stress resultants as the sum of the two components; for example,

$$T_\eta^*=T_{\eta a}^*+T_{\eta b}^*$$

Then from equations (2)

$$r\kappa_\eta+\nu r\kappa_\theta-\frac{12}{\gamma}(1-\nu^2)M_{\eta b}^*=\frac{12}{\gamma}(1-\nu^2)M_{\eta a}^*$$

$$\epsilon_\eta+\nu\epsilon_\theta-(1-\nu^2)T_{\eta b}^*=(1-\nu^2)T_{\eta a}^* \quad (7)$$

$$\epsilon_\theta-T_{\theta b}^*+\nu T_{\eta b}^*=T_{\theta a}^*-\nu T_{\eta a}^*$$

$$\frac{\gamma}{12}r\kappa_\theta-M_{\theta b}^*+\nu M_{\eta b}^*=M_{\theta a}^*-\nu M_{\eta a}^*$$

Components $T_{\eta a}^*$, etc., are stress resultants applied to the pipe bend to give the loading, whereas components $T_{\eta b}^*$, etc., are self-equilibrating.

Static-Geometric Analogy

Consider now the stress resultants generated by equation (3) by substituting stress functions U, W, V, Y for nondimensional displacements u/r, w/r, v/r, ψ in accordance with Goldenweizer's static-geometric analogy (Novozhilov, 1970).

$$M_\eta^*=-\frac{\partial U}{\partial\theta}-gW\qquad T_\eta^*=\frac{\partial}{\partial\theta}(gU)-\frac{\partial^2 W}{\partial\theta^2}$$

$$M_\theta^*=-\frac{1}{\delta}\left(\frac{\partial V}{\partial\eta}-\frac{\sin\mu}{\rho}U+\frac{\cos\mu}{\rho}W\right)$$

$$T_\theta^*=\frac{1}{\delta}\left(\frac{\partial Y}{\partial\eta}+\frac{\sin\mu}{\rho}\left(\frac{\partial W}{\partial\theta}-gU\right)\right) \quad (8)$$

$$P^*=g\frac{\partial V}{\partial\theta}-\frac{\partial Y}{\partial\theta}\qquad Q^*=\frac{\partial^2 V}{\partial\theta^2}+gY$$

where $$Y=\frac{1}{\delta}\left(\frac{\cos\mu}{\rho}V-\frac{\partial W}{\partial\eta}\right)$$

The stress resultants must be single valued, or continuous around the pipe cross section, but will satisfy equations (1) whatever the stress functions if pressure q is zero. For the forces and moment on the cross section, we find by substituting in equations (4) and integrating

$$F_z=rEt\left[gU-\frac{\partial W}{\partial\theta}\right]_o^{2\pi}$$

$$F_y=rEt\left[\frac{\partial V}{\partial\theta}\cos\mu+Y\sin\mu\right]_o^{2\pi} \quad (9)$$

$$M=r^2Et\left[U\cos\mu+W\sin\mu-\frac{y}{r}\left(gU-\frac{\partial W}{\partial\theta}\right)\right]_o^{2\pi}$$

If U, W, V, Y are single valued then F_y, F_z, M vanish and the stress resultants generated will be self-equilibrating.

The governing equations are now derived from equations (7) by writing the strains in terms of displacements and the self-equilibrating stress resultants in terms of single valued stress functions U_b, W_b, V_b, Y_b.

$$[A]\begin{Bmatrix}U_b\\W_b\\u/r\\w/r\end{Bmatrix}-\frac{\partial}{\partial\eta}\begin{Bmatrix}\psi\\v/r\\Y_b\\V_b\end{Bmatrix}=\begin{Bmatrix}-\frac{12}{\gamma}(1-\nu^2)\delta M_{\eta a}^*\\-(1-\nu^2)\delta T_{\eta a}^*\\\delta T_{\theta a}^*-\nu\delta T_{\eta a}^*\\-\delta M_{\theta a}^*+\nu\delta M_{\eta a}^*\end{Bmatrix} \quad (10)$$

where matrix A, given in Table 1, involves differentials in θ only.

Applied stress resultants $M_{\eta a}^*$, $T_{\eta a}^*$, $M_{\theta a}^*$, $T_{\theta a}^*$ are now required which give the loading on the pipe and also satisfy equations (1). This differs from the analysis for circular cross sections (Whatham, 1986) where the governing equations satisfied equilibrium and $T_{\eta a}^*$, P_a^* were only required to give the loading.

Table 1 Matrix A

U_b	W_b	u/r	w/r
$-\frac{12}{\gamma}(1-\nu^2)\delta\frac{\partial}{\partial\theta}$	$-\frac{12}{\gamma}(1-\nu^2)\delta g$	$\frac{\sin\mu}{\rho}g-\nu\delta\frac{\partial}{\partial\theta}(g\ )$	$-\frac{\sin\mu}{\rho}\frac{\partial}{\partial\theta}+\nu\delta\frac{\partial^2}{\partial\theta^2}$
$(1-\nu^2)\delta\frac{\partial}{\partial\theta}(g\ )$	$-(1-\nu^2)\delta\frac{\partial^2}{\partial\theta^2}$	$\frac{\sin\mu}{\rho}-\nu\delta\frac{\partial}{\partial\theta}$	$-\frac{\cos\mu}{\rho}-\nu\delta g$
$\frac{\sin\mu}{\rho}g+\nu\delta\frac{\partial}{\partial\theta}(g\ )$	$-\frac{\sin\mu}{\rho}\frac{\partial}{\partial\theta}-\nu\delta\frac{\partial^2}{\partial\theta^2}$	$\delta\frac{\partial}{\partial\theta}$	δg
$\frac{\sin\mu}{\rho}+\nu\delta\frac{\partial}{\partial\theta}$	$-\frac{\cos\mu}{\rho}+\nu\delta g$	$-\frac{\gamma}{12}\delta\frac{\partial}{\partial\theta}(g\ )$	$\frac{\gamma}{12}\delta\frac{\partial^2}{\partial\theta^2}$

Table 2 Oval cross section constants

b/a	$\pi a/c$	$A/\pi r^2$	B	C_1	$C_{1(\text{ellipse})}$
1	1	1	0	0	0
1.1	0.9511	0.9966	−0.0714	0.1679	0.1730
1.2	0.9046	0.9877	−0.1360	0.2862	0.3018
1.3	0.8608	0.9750	−0.1945	0.3705	0.3979
1.4	0.8198	0.9596	−0.2476	0.4308	0.4696
1.6	0.7457	0.9247	−0.3400	0.5038	0.5618
1.8	0.6815	0.8880	−0.4171	0.5381	0.6097
2.0	0.6260	0.8521	−0.4821	0.5508	0.6309

$$A=\int_o^c y\cos\mu ds,\quad A_i=A-\left(\frac{t}{r}-\frac{\gamma}{4}\right)\pi r^2$$

Furthermore, if the governing equations had been in terms of $M^*_{\eta b}$, $T^*_{\eta b}$, ϵ_θ, κ_θ satisfying equilibrium and compatibility, we would have had to include the last of equations (4) in terms of $M^*_{\eta b}$, $T^*_{\eta b}$ and equated to zero to obtain self-equilibrium and the same equation in terms of $-\epsilon_\theta$, κ_θ equated to zero to ensure the displacements were single valued.

Although end effects are neglected, an earlier paper (Whatham, 1981) described how the effects of a rigid flange or a tangent pipe could be included.

Applied Stress Resultants for In-Plane Bending

Stress functions can be used to generate these stress resultants because q is zero; if V_a, Y_a are zero then, by equations (5) and (9), F_y and F_z also vanish. The following multivalued stress functions were derived from equations (9) to give a bending moment M on the pipe and single valued stress resultants.

$$\begin{aligned} U_a &= M^*\theta\cos\mu \\ W_a &= M^*\theta\sin\mu \\ V_a &= Y_a = 0 \end{aligned} \tag{11}$$

where $M^*=M/2\pi r^2 Et$.

By equations (8)

$$\begin{aligned} T^*_{\eta a} &= -M^* g\cos\mu \\ M^*_{\eta a} &= -M^*\cos\mu \\ T^*_{\theta a} &= (M^*/\rho\delta)\sin^2\mu \\ M^*_{\theta a} &= 0 \end{aligned} \tag{12}$$

These stress resultants were substituted in equation (10) to obtain the solution for pure in-plane bending.

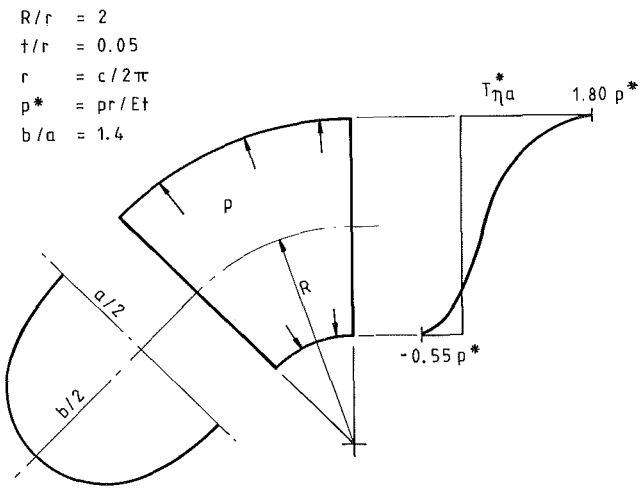


Fig. 2 $T^*_{\eta a}$ for pressurizing an oval pipe bend

Applied Stress Resultants for Pressurizing

Stress resultants could not be found which would simply pressurize a noncircular pipe bend; the stress resultants which were used pressurized the bend but at the same time imposed a bending moment on the cross section, requiring a bending moment applied in the opposite direction to neutralize it. Writing equations (1) in terms of applied stress resultants and neglecting $M^*_{\eta a}$, $M^*_{\theta a}$.

$$\begin{aligned} \frac{\partial}{\partial\theta}(\delta T^*_{\theta a})+\frac{\sin\mu}{\rho}T^*_{\eta a}&=0 \\ T^*_{\theta a}/r_\theta + T^*_{\eta a}/r_\eta &= q^*/r \end{aligned} \tag{13}$$

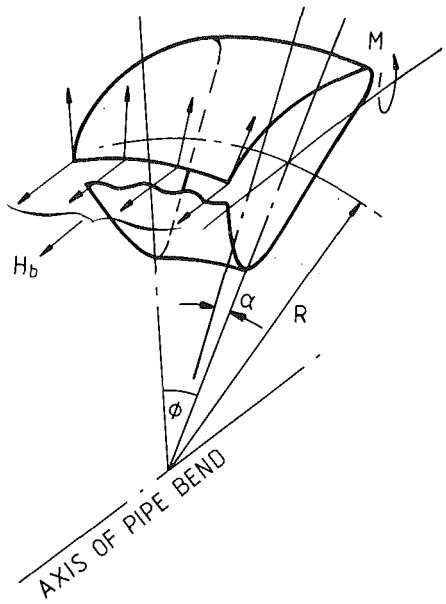


Fig. 3 Segment of middle surface

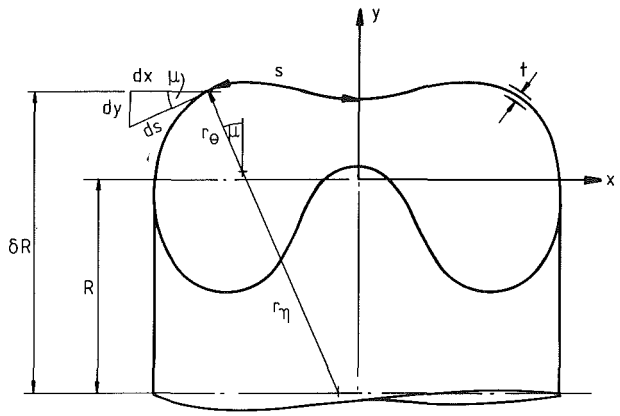


Fig. 4 Example of symmetrical cross section

As $t \to 0$ then $(q, T_{\eta a}, T_{\theta a}) \to (p, \bar{T}_{\eta a}, \bar{T}_{\theta a})$ and hence

$$\left[\delta \bar{T}^*_{\theta a} \cos\mu\right]^{\mu}_{\mu_o} = \frac{1}{2} p^* \rho \left[\delta^2\right]^{\mu}_{\mu_0} \tag{14}$$

where

$$p^* = pr/Et$$

If the cross section has double symmetry, then $\delta = 1$ at $\mu_o = \pi/2$ and

$$\bar{T}^*_{\theta a} = \frac{1}{2} p^* \rho (\delta^2 - 1)/\delta \cos\mu \tag{15}$$

$$\bar{T}^*_{\eta a} = \rho\delta(p^* - g\bar{T}^*_{\theta a})/\cos\mu$$

At the flanks ($\mu = \pm \pi/2$)

$$\bar{T}^*_{\theta a} = 2\bar{T}^*_{\eta a} = p^*/g$$

For a finite wall thickness, solving equations (13) with $T^*_{\theta a}$, $T^*_{\eta a}$, q^* replaced by $T^*_{\theta a} - \bar{T}^*_{\theta a}$, $T^*_{\eta a} - \bar{T}^*_{\eta a}$, $q^* - p^*$ gives

$$T^*_{\theta a} = \bar{T}^*_{\theta a} - \left(\frac{t}{2r} - \frac{\gamma}{8}\frac{r}{r_\eta}\right)p^*$$

$$T^*_{\eta a} = \bar{T}^*_{\eta a} - \left(\frac{t}{2r} - \frac{\gamma}{8}\frac{r}{r_\theta}\right)p^* \tag{16}$$

$$M^*_{\eta a} = M^*_{\theta a} = 0$$

The distribution of $T^*_{\eta a}$ in Fig. 2 indicates that a bending moment is produced when the pipe cross section is oval; suppose this moment is given by

$$M = -C_1 R\pi r^2 p \tag{17}$$

where coefficient C_1 must be determined. For cross sections with double symmetry

$$C_1 = \frac{1}{R\pi r^2 p}\int_o^c y\bar{T}_{\eta a} ds = 2\int_o^{2\pi} \frac{y}{c\cos\mu}\left(1 - \frac{gy}{r\cos\mu}\right)d\theta \tag{18}$$

whence C_1 may be determined by quadrature. As it is independent of t and R, the bending moment by equation (17) for a doubly symmetric shape is proportional to the bend radius, the pressure, and through πr^2 the area A enclosed by the mid-surface but is independent of wall thickness. Applying equation (18) to an elliptical cross section gives Lorenz's result (Clark et al., 1952)

$$C_{1_{(\text{ellipse})}} = \frac{A}{\pi r^2}\left(1 - \frac{a^2}{b^2}\right) \tag{19}$$

where b,a are the dimensions of the major and minor axes; Table 2 gives C_1 for a class of oval cross sections which will be described later and $C_{1_{(\text{ellipse})}}$ for the same circumference elliptical cross sections.

The stress resultants from equations (15) and (16) were substituted in equations (10) to obtain the solution for combined pressurizing and bending; for pressurizing alone the bending moment from equation (17) was applied to the pipe bend in the opposite direction and the solutions added.

Expressing the Variables and Solving

Stress functions U_b, W_b and dimensionless displacements u/r, w/r were expressed as Fourier series in θ with appropriate parity but V_b, Y_b, v/r, ψ could be expressed more simply. Rotation α in Fig. 3 is a measure of v/r and ψ because

$$v/r = -\alpha\rho\delta$$

$$\psi = -\alpha\cos\mu \tag{20}$$

and as a check, the warping strains λ, β are zero by equations (3).

There is a constant component H_b of force in the wall acting in a direction parallel to the pipe bend axis as shown in Fig. 3 and this component happens to be a measure of V_b and Y_b

$$V_b = H^*_b\rho\delta$$

$$Y_b = H^*_b\cos\mu \tag{21}$$

where

$$H^*_b = H_b/rEt$$

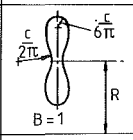
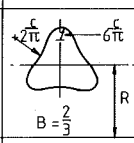
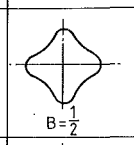
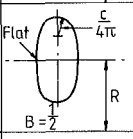
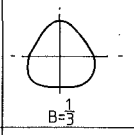
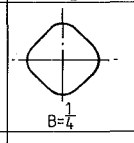
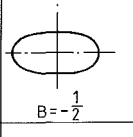
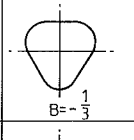
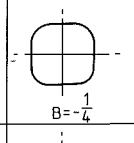
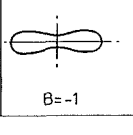
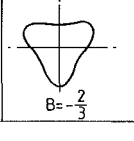
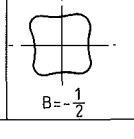
Stress resultants P_b, Q_b are zero by equations (8) and, in the case of in-plane bending, H_b is the total force because V_a, Y_a, and hence H_a, are zero. We note that α is zero in a pressurized torus but H_b is not.

If the Fourier series for the variables are truncated to N terms there are $4N+2$ coefficients to be determined in addition to ratios α/η and H^*_b/η – for in-plane bending or pressurizing, $\partial/\partial\eta$ in equation (10) is replaced by $1/\eta$. The four governing equations were expanded to the required $4N+4$ equations by collocation at the mid-points of $N+1$ equal intervals around half the circumference; these equations were then solved with $N = 32$ which gave converged solutions in all cases. For circular cross sections, the governing equations were also expanded by equating the coefficients of like terms (Fourier analysis) which gave the same solution for the same N.

Defining the Pipe Cross-Section

A randomly generated symmetrical middle surface is shown in Fig. 4, defined by angle μ as a function of distance s

Table 3 Multi-lobe cross sections

cos mθ at lobes	r_θ at lobes	r_θ between lobes	m = 2	m = 3	m = 4
1	$\frac{c}{6\pi}$	$-\frac{c}{2\pi}$			
	$\frac{c}{4\pi}$	∞			
-1	$\frac{c}{4\pi}$	∞			
	$\frac{c}{6\pi}$	$-\frac{c}{2\pi}$			

$$\mu = \theta + \mu_1 \sin\theta + \mu_2 \sin 2\theta + \ldots \tag{22}$$

where

$$\theta = 2\pi s/c$$

The function $g = d\mu/d\theta$ is easily calculated but the ratio $t/r_\theta = gt/r$ must nowhere exceed 0.3 for shell theory to apply (Whatham, 1986).

Cartesian coordinate y of the outline is required in order to calculate δ; taking the centroid of the cross section inside the pipe as origin, we see from Fig. 4

$$dx = \cos\mu ds, \text{ hence } x = -\int_o^s \cos\mu ds$$
$$dy = \sin\mu ds, \text{ hence } y = y_o - \int_o^s \sin\mu ds \tag{23}$$

enabling x,y to be determined by quadrature.

Coefficients μ_1, μ_2, etc., for a particular cross section are derived by curve fitting and, for continuity of x and y, it is necessary that

$$\int_o^c \sin\mu ds = r\int_o^{2\pi} \sin\mu d\theta = 0$$
$$\int_o^c \cos\mu ds = r\int_o^{2\pi} \cos\mu d\theta = 0 \tag{24}$$

The sin μ integral is zero because sin μ is an odd function of θ but the coefficients of equation (22) may have to be adjusted to eliminate the cos μ integral. For example, the cross section in Fig. 4 was generated by

$$\mu = \theta + 0.5\sin\theta - 0.9452\sin 2\theta + 0.3\sin 3\theta - 0.4\sin 4\theta \tag{25}$$

where the coefficient -0.9452 has been adjusted for that purpose.

We now consider the continuity of a class of multi-lobe cross sections shown in Table 3, defined by

$$\mu = \theta + B\sin m\theta \tag{26}$$

where m = number of lobes

For these sections

$$\int_o^{2\pi} \cos\mu d\theta = \int_o^{2\pi} \cos\theta\cos(B\sin m\theta)d\theta \quad (a) \tag{27}$$

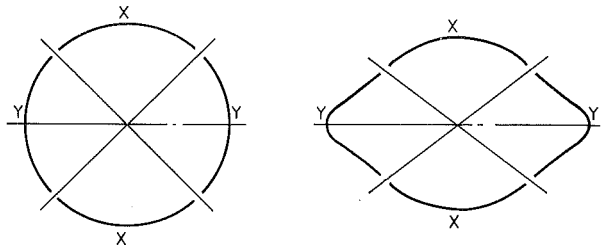


Fig. 5 Thin wall pipe bend distortion from in-plane bending

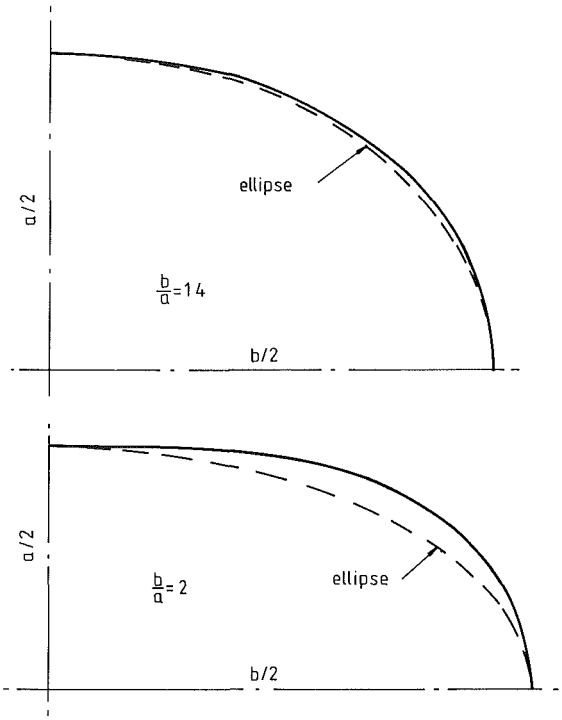


Fig. 6 Quadrants of oval cross sections

$$-\int_o^{2\pi} \sin\theta\sin(B\sin m\theta)d\theta \quad (b)$$

Integral (a) is zero because the integrand is an odd function of θ; to determine integral (b), formula 8.514-6 of Gradshteyn and Ryzhik (1980) yields

$$\sin(B\sin m\theta) = 2\sum_{k=o}^{\infty} J_{2k+1}(B)\sin(2k+1)m\theta \tag{28}$$

where $J_{2k+1}(B)$ are Bessel functions. Integral (b) then becomes

$$2\sum_{k=o}^{\infty} J_{2k+1}(B)\int_o^{2\pi} \sin\theta\sin(2k+1)m\theta d\theta \tag{29}$$

From orthogonality, integral (b) is zero if $m > 1$ but not when $m = 1$, unless $J_1(B) = 0$ which requires that

$$B = 3.8137, 7.0156, \text{ etc.}$$

Such shapes are absurd because the contour crosses itself, so we must have $m > 1$ and the circumference is then continuous regardless of B.

In-plane bending distorts circular pipe bends with thick walls ($t/r = 0.1$) to an oval shape but thin walls ($t/r = 0.01$) distort as shown in Fig. 5 where segments X remain almost stress free and the pipe bend deflection depends on segments Y. Two and four-lobe cross sections were used to investigate the relationship between pipe bend flexibility and flank radius r_f, the cross sections being defined by equation (26) with

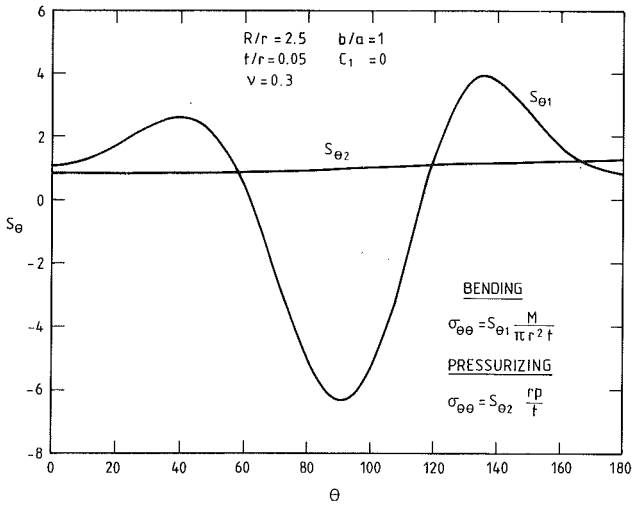


Fig. 7 Hoop stress on outside surface—circular cross section

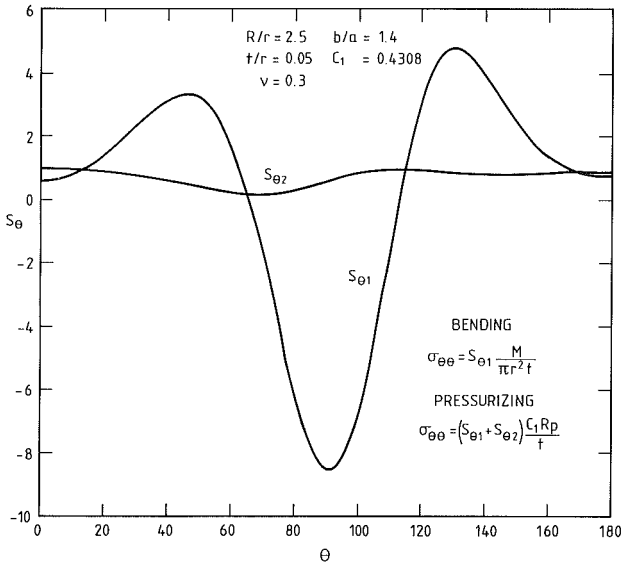


Fig. 8 Hoop stress on outside surface—oval diameter ratio 1.4

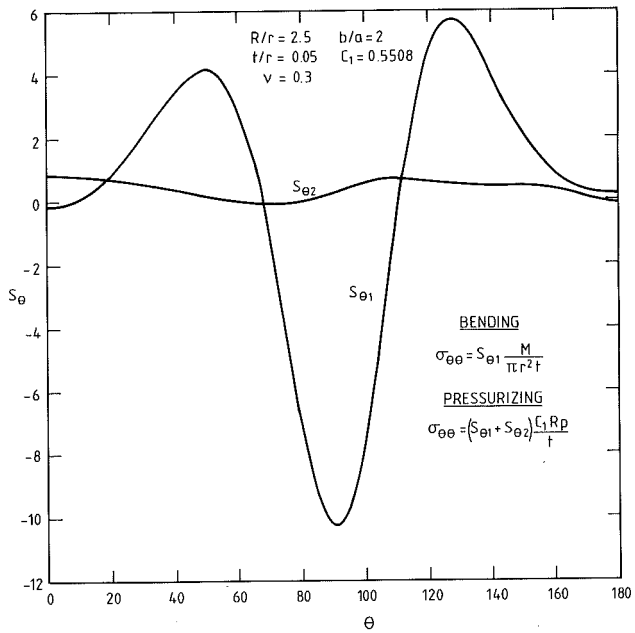


Fig. 9 Hoop stress on outside surface—oval diameter ratio 2.0

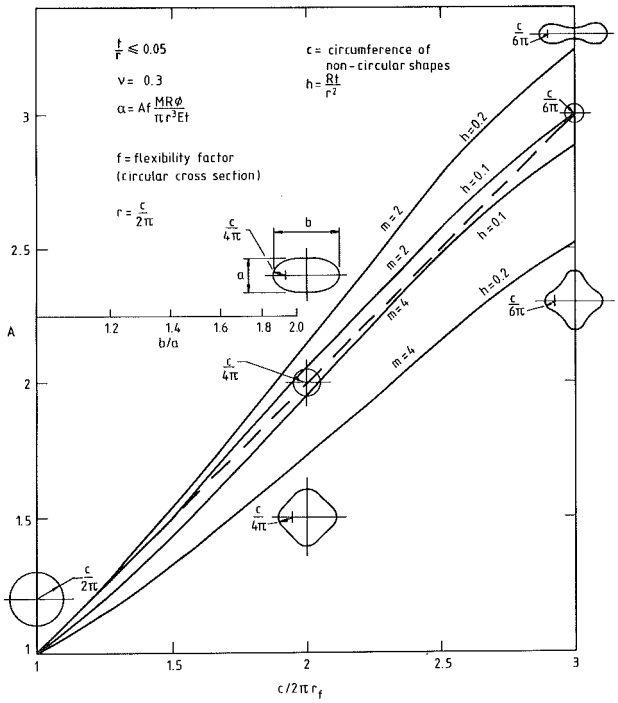


Fig. 10 Pipe bend flexibility versus flank curvature

$$m=2,\ B=\frac{1}{2}(1-c/2\pi r_f)$$
$$m=4,\ B=-\frac{1}{4}(1-c/2\pi r_f) \tag{30}$$

Two-lobe oval cross sections were then investigated further and the relationship between B and major-to-minor axis ratio b/a is given in Table 2; quadrants of two cross sections in Fig. 6 show that, if b/a is less than 1.4, they approximate the elliptical cross sections analyzed by Clark et al. (1952).

Results

Figure 7 shows the outside surface hoop stress distributions from in-plane bending or internally pressurizing a pipe bend of circular cross-section. The stresses cannot be added to obtain the combined effect of bending and pressurizing because there is a nonlinearity; pressurizing maintains circularity of the cross section and counteracts the ovalizing effect of the bending moment whether the bend angle is being opened or closed. But if the pipe cross section is noncircular, then pressurizing and bending solutions can be added.

Figures 8 and 9 show the hoop stress on pipe bends of oval cross section from a bending moment which opens the bend and also the hoop stress from the combined effects of internal pressure and a bending moment $C_1R\pi r^2p$ which closes the bend. The stresses from the combined loading are relatively small and hence, by superposition, the stresses from pressurizing alone approximate those from a bending moment which opens the bend, tending to straighten the pipe. This supports the observation of Clark et al. (1952) regarding pipe bends of elliptical cross section.

Factors enabling the deflections under bending moment or pressure to be calculated have been published for a range of oval pipe bends in which $R/r=1.25-20$, $t/r=0.01-0.1$ and $b/a=1-2$ (Whatham, 1983).

Results of the investigation of pipe bends with two and four-lobe cross sections plotted in Fig. 10 show that, for cross sections with double symmetry, flexibility is almost inversely proportional to flank radius r_f if the pipe wall is thin

($t/r \leq 0.05$) and the bend characteristic $h = 4\pi^2 Rt/c^2$ is less than 0.1. This is also the asymptotic result of Clark et al. (1952) for elliptic pipe bends written as

$$\alpha = \frac{\sqrt{3(1-\nu^2)}M\phi}{\pi r_f t^2 E} \tag{31}$$

Conclusions

Equations have been presented for an exact thin shell analysis of curved pipes with symmetrical noncircular cross sections when subjected to in-plane bending or pressurization, employing Goldenweizer's static-geometric analogy. There is no restriction on the symmetrical cross section shape, provided the wall thickness-to-radius ratio does not exceed 0.3. End effects were neglected but an earlier paper (Whatham, 1981) described how end effects from, for example, a rigid flange or a tangent pipe could be included.

Hoop stress distributions showed that, for oval cross sections, pressurizing has approximately the same effect as a bending moment which tends to straighten the pipe. Pressure and bending moment solutions could be superimposed unless the cross section was circular. Calculations for two and four-lobe cross sections showed that the flexibility of pipe bends with double symmetry is almost inversely proportional to flank radius if the pipe wall is thin and the bend radius moderate.

The work has resulted in the publication of factors giving the deflection under bending moment or pressure of a range of pipe bends with oval cross sections.

Acknowledgment

The author acknowledges the advice and encouragement of Professor J. J. Thompson of the School of Nuclear Engineering, University of New South Wales.

References

Clark, R. A., Gilroy, T. I., and Reissner, E., 1952, "Stresses and Deformations of Toroidal Shells of Elliptical Cross Section," ASME JOURNAL OF APPLIED MECHANICS, Vol. 19, pp. 37–48.

Gradshteyn, I. S., and Ryzhik, I. M., 1980, *Table of Integrals, Series, and Products*, Academic Press, New York.

Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, 4th Ed., Dover Publications, New York.

Novozhilov, V. V., 1970, *Thin Shell Theory*, Wolters-Noordhoff, Gröningen.

Whatham, J. F., 1981, "Thin Shell Analysis of Non-Circular Pipe Bends," *Journal of Nuclear Engineering and Design*, Vol. 67, pp. 287–296.

Whatham, J. F., 1983, "Results of Pipe Bend Analysis, Part V: Flexibility and Pressure Deflection Factors of Oval Pipe Bends," Australian Atomic Energy Commission Report E555.

Whatham, J. F., 1986, "Pipe Bend Analysis by Thin Shell Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 173–180.

D. S. Cairns
Research Assistant.

P. A. Lagace
Associate Professor of
Aeronautics and Astronautics.

Technology Laboratory
for Advanced Composites,
Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
Cambridge, MA 02139

Thick Composite Plates Subjected to Lateral Loading

An analytical solution for an orthotropic plate subjected to general lateral loading is presented. The solution uses a stress function approach to obtain the localized stresses and strains due to the loading by an axisymmetric rigid sphere. Plots of load versus local indentation are compared with experimental test data previously reported in the literature. The analysis agrees well with the experimental data and could be used in conjunction with failure criteria to predict damage initiation in such a localized region.

1 Introduction

An important factor in the design of composite structures is their response to impact events. Events such as tool drop and runway kickup are examples of impact on aircraft structures. During impact, localized stresses are generated which can cause premature failure. Accurate prediction of these stresses is important in advanced composite materials as a first step in assessing structural sensitivity due to impact. Advanced composites, due to their relatively low strength in nonfiber-reinforced directions (e.g., out-of-plane), exhibit sensitivity to the lateral loading which occurs during an impact event. This results in delamination (separation of plies) and transverse cracking (in the direction perpendicular to fibers in an individual ply) as well as fiber breakage.

Current research has been aimed at separating the local contact problem from the dynamic problem such that analysis or test data from statical indentation laws may be used in the analysis of dynamical transients (Lal, 1983; Sun and Chattopadhyay, 1975). The solution to the contact problem for an infinite half-plane (Conway, 1956) has been employed with some success by Greszczuk (1975) but is limited to the infinite boundaries in the analysis. Recently, Tan and Sun (1985) conducted static indentation experiments to determine the contact laws to be implemented into subsequent finite element analyses. The present solution is analytic and is used to develop the static indentation laws based on ply properties and, more importantly, the localized deformation such that failure criteria may be applied.

2 Axisymmetric Elasticity Solution

To solve for the localized loading in a homogeneous, orthotropic plate, of thickness h and radius R_d, the axisymmetric boundary value problem shown in Fig. 1 is solved. While the assumption of through-the thickness homogeneity may be severe, subsequent comparison with test data will demonstrate its utility for laminates constructed from thin plies. The plate is supported along its edges and a moment M_r is applied via superposition along its perimeter to approximate the far-field loading outside of the local perturbations due to the axisymmetric load distribution caused by contact. For small strains under the assumption of axisymmetry, the following kinematic relationships apply:

$$\epsilon_{rr} = \frac{\partial u_r}{\partial r} \qquad \epsilon_{zz} = \frac{\partial w}{\partial z}$$
$$\epsilon_{\theta\theta} = \frac{u_r}{r} \qquad \gamma_{rz} = \frac{\partial u_r}{\partial z} + \frac{\partial w}{\partial r} \tag{1}$$

Here, u_r, is the radial displacement and w is the displacement in the z direction. The assumption of axisymmetry provides that $\partial/\partial\theta$ is identically zero. The compatibility conditions associated with the above assumptions are:

$$\epsilon_{rr} - \frac{\partial}{\partial r}(r\epsilon_{\theta\theta}) = 0$$
$$\frac{\partial^2 \epsilon_{rr}}{\partial z^2} + \frac{\partial^2 \epsilon_{zz}}{\partial r^2} - \frac{\partial^2 \gamma_{rz}}{\partial r \partial z} = 0 \tag{2}$$

The constitutive properties for a cylindrically orthotropic material can be represented as:

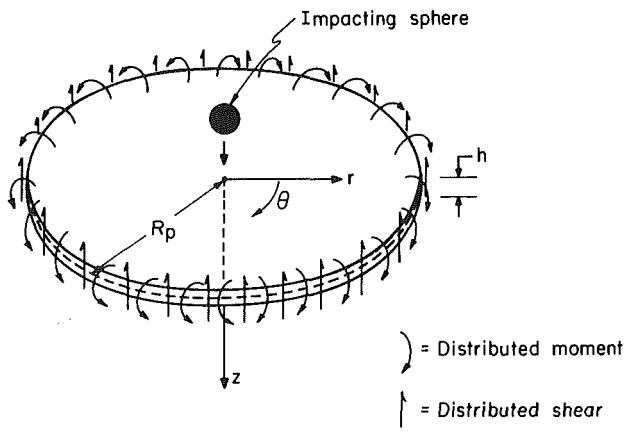


Fig. 1 Problem schematic

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 22, 1986; final revision, September 17, 1986.

$$\{\epsilon_{rr}\epsilon_{\theta\theta}\epsilon_{zz}\} = [a_{ij}]\ \{\sigma_r\sigma_\theta\sigma_z\}$$
$$\gamma_{rz} = a_{44}\ \sigma_{rz} \tag{3}$$

where a_{ij} is the engineering compliance matrix. For the present problem, the material is furthermore assumed to be transversely isotropic in the $r-\theta$ plane. In terms of the engineering constants, the constitutive properties are thus:

$$\begin{aligned} a_{11} &= a_{22} = \frac{1}{E_{rr}} = \frac{1}{E_{\theta\theta}} \qquad a_{33} = \frac{1}{E_{zz}} \\ a_{12} &= -\frac{\nu_{r\theta}}{E_{rr}} = -\frac{\nu_{\theta r}}{E_{\theta\theta}} \qquad a_{44} = \frac{1}{G_{rz}} = \frac{1}{G_{\theta z}} \\ a_{13} &= -\frac{\nu_{rz}}{E_{rr}} = -\frac{\nu_{\theta z}}{E_{\theta\theta}} \end{aligned} \tag{4}$$

The governing stress equilibrium equations for axisymmetric problems in the absence of body forces may be written as:

$$\begin{aligned} \frac{\partial\sigma_{rr}}{\partial r} + \frac{\partial\sigma_{rz}}{\partial z} + \frac{\sigma_{rr}-\sigma_{\theta\theta}}{r} = 0 \\ \frac{\partial\sigma_{rz}}{\partial r} + \frac{\partial\sigma_{zz}}{\partial z} + \frac{\sigma_{rz}}{r} = 0 \end{aligned} \tag{5}$$

Combining equations (2) and (3) to represent the compatibiity in terms of the stress components, the compatibility conditions become:

$$\begin{aligned} &a_{11}\sigma_{rr} + a_{12}\sigma_{\theta\theta} + a_{13}\sigma_{zz} - \frac{\partial}{\partial r} \\ &\qquad [r(a_{12}\sigma_{rr} + a_{11}\sigma_{\theta\theta} + a_{13}\sigma_{zz})] = 0 \\ &\frac{\partial^2}{\partial z^2}\ [a_{11}\sigma_{rr} + a_{12}\sigma_{\theta\theta} + a_{13}\sigma_{zz}] + \frac{\partial^2}{\partial r^2} \\ &\qquad [a_{13}\ (\sigma_{rr} + \sigma_{\theta\theta}) + a_{33}\sigma_{zz}]\ - a_{44}\ \frac{\partial^2\sigma_{rz}}{\partial r\partial z} = 0 \end{aligned} \tag{6}$$

In the absence of body forces (or if body forces are derivable from a potential) a stress function $\psi\ (r, z)$ (Lekhnitskii, 1963) is introduced satisfying the first equilibrium equation of equation (5) and equation (6) such that:

$$\begin{aligned} \sigma_{rr} &= -\frac{\partial}{\partial z}\left(\frac{\partial^2\psi}{\partial r^2} + \frac{b}{r}\frac{\partial\psi}{\partial r} + a\frac{\partial^2\psi}{\partial z^2}\right) \\ \sigma_{\theta\theta} &= -\frac{\partial}{\partial z}\left(b\frac{\partial^2\psi}{\partial r^2} + \frac{1}{r}\frac{\partial\psi}{\partial r} + a\frac{\partial^2\psi}{\partial z^2}\right) \\ \sigma_{zz} &= \frac{\partial}{\partial z}\left(c\frac{\partial^2\psi}{\partial r^2} + \frac{c}{r}\frac{\partial\psi}{\partial r} + d\frac{\partial^2\psi}{\partial z^2}\right) \\ \sigma_{rz} &= \frac{\partial}{\partial r}\left(\frac{\partial^2\psi}{\partial r^2} + \frac{1}{r}\frac{\partial\psi}{\partial r} + a\frac{\partial^2\psi}{\partial z^2}\right) \end{aligned} \tag{7}$$

where

$$\begin{aligned} a &= \frac{a_{13}(a_{11}-a_{12})}{\Delta} \\ b &= \frac{a_{13}(a_{13}+a_{44})-a_{12}a_{33}}{\Delta} \\ c &= \frac{a_{13}(a_{11}-a_{12})+a_{11}a_{44}}{\Delta} \\ d &= \frac{a_{11}^2-a_{12}^2}{\Delta} \\ \Delta &= a_{11}a_{33}-a_{13}^2 \end{aligned} \tag{8}$$

Equations (5) and (7) are combined to obtain the governing equation in ψ:

$$\begin{aligned} &\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)\left(\frac{\partial^2\psi}{\partial r^2} + \frac{1}{r}\frac{\partial\psi}{\partial r} + a\frac{\partial^2\psi}{\partial z^2}\right) + \\ &\qquad \frac{\partial^2}{\partial r^2}\left(c\frac{\partial^2\psi}{\partial r^2} + \frac{c}{r}\frac{\partial\psi}{\partial r} + d\frac{\partial^2\psi}{\partial z^2}\right) = 0 \end{aligned} \tag{9}$$

Equation (9) may be further written in the form:

$$\nabla_1^2\ \nabla_2^2\ \psi = 0 \tag{10}$$

where

$$\nabla_i^2 = \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{s_i^2}\frac{\partial^2}{\partial z^2}\right) \tag{11}$$

and

$$s_i = \left[\frac{a+c \pm [(a+c)^2-4d]^{1/2}}{2d}\right]^{1/2} \tag{12}$$

A solution for ψ is sought in the form of equation (13) which is separable in r and z:

$$\psi = \sum_{m=1}^{\infty} f_m(z)\ g_m(r) \tag{13}$$

The edge boundary conditions (traction free) for a supported plate may be approximated by harmonics of a Bessel function of the first kind:

$$g_m(r) = J_o(\omega_m r) \tag{14a}$$

with

$$\omega_m = \frac{\mu_m}{R_d} \tag{14b}$$

and

$$J_o\ (\mu_m) = 0 \tag{14c}$$

This satisfies the condition that the solution of ψ be regular at r equal to zero. Substitution into equation (10) results in an ordinary differential equation in z. The solution of this equation may be satisfied by equation (15):

$$f_m(z) = A_m e^{s_1\omega_m z} + B_m e^{s_2\omega_m z} + C_m e^{s_1\omega_m z} + D_m e^{-s_2\omega_m z} \tag{15}$$

for real distinct roots s_1 and s_2. The constants A_m, B_m, C_m, and D_m are evaluated from the upper and lower surface stresses as specified in the boundary conditions.

3 Boundary Conditions

The actual boundary conditions for the plate away from the point of load application are neither simply-supported nor clamped since the edge is elastic. The area of interest, however, is the very localized region near the point of contact. Thus, as an approximation to the edge conditions, a constant moment is imposed on the plate to properly adjust the bending stresses σ_{rr} and $\sigma_{\theta\theta}$. The magnitude of this bending moment is calculated using the procedure outlined by Love (1927). A moment is calculated at a distance away from the point of load application. This moment is then invoked on the plate edge and the results of this solution are added to the current simply-supported solution via superposition. This method is validated by the fact that the standard plate solution is recovered in a short distance away from the local perturbations as will be subsequently shown.

On the top surface of the plate, the loading $p(r)$ is known and may be expanded as a Fourier-Bessel series of the form:

$$p(r) = \sum_{m=1}^{\infty} \beta_m J_o\ (\omega_m r) \tag{16a}$$

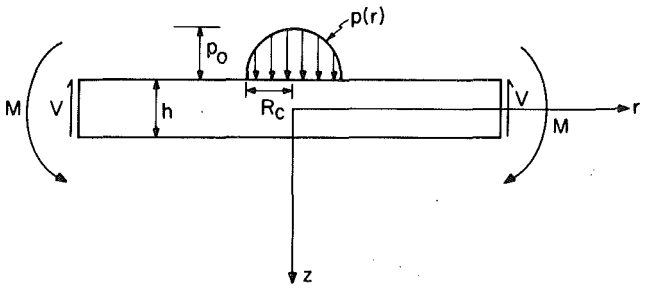


Fig. 2 Hertzian contact loading

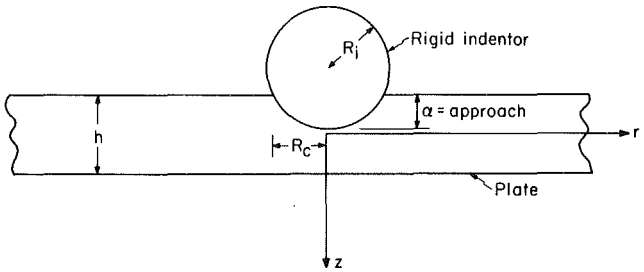


Fig. 3 Rigid indentor contact schematic

where

$$\beta_m = \frac{2}{J_1^2(\mu_m) R_p^2} \int_o^{R_p} p(r) r J_o(\omega_m r) dr \tag{16b}$$

The boundary conditions at the top surface (z equal to $-h/2$) are:

$$\begin{aligned} \sigma_{zz} &= -p(r) \\ \sigma_{rz} &= 0 \end{aligned} \tag{17}$$

On the bottom surface of the plate, (z equal to $h/2$) the boundary conditions are as follows:

$$\begin{aligned} \sigma_{zz} &= 0 \\ \sigma_{rz} &= 0 \end{aligned} \tag{18}$$

Invoking these boundary conditions in equation (15) results in the expressions for the constants A_m, B_m, C_m, and D_m contained in Appendix A.

4 The Hertzian Assumption

As an approximation, the load distribution is assumed to be Hertzian in nature. This assumption is valid based on recent experimental data generated by Tan and Sun (1985) and supported analytically by Sankar (1985) for the small indentation to indentor-radius ratios considered here. The loading assumption for a rigid indentor is shown in Fig. 2. The loading distribution (Hertzian) is of the form:

$$p(r) = p_o \left(1 - \left(\frac{r}{R_c}\right)^2\right)^{.5} \tag{19}$$

where p_o is the maximum load intensity and R_c is the radius of contact. The Hertzian loading is expanded as a Fourier-Bessel series of the type as shown in equation (16). The β_m are obtained by using equations (16) and (19) with the integration in equation (16) extending up to the radius of contact, R_c.

For the contact problem, the indentor is assumed to be rigid such that the indentation is an explicit function of R_c. This assumption is valid for metal indentors where the stiffness of the indentor is much greater than the E_{zz} of a composite material. (A typical value of E_{zz} is approximately 10.6 GPa.)

In order to determine the unknown constant p_o, the strain is integrated with respect to z to obtain the displacement w as:

$$w = -a_{13}\left((b+1)\left[\frac{\partial^2\psi}{\partial r^2} + \frac{1}{r}\frac{\partial\psi}{\partial r}\right] + 2a\frac{\partial^2\psi}{\partial z^2}\right) + a_{33}\left(c\frac{\partial^2\psi}{\partial r^2} + \frac{c}{r}\frac{\partial\psi}{\partial r} + d\frac{\partial^2\psi}{\partial z^2}\right) \tag{20}$$

Finally, the approach α, is defined as the maximum indentor penetration as illustrated in Fig. 3 based on an indentor radius R_i. From geometry and the rigid indentor assumption the approach α is calculated according to:

$$\alpha = R_i - (R_i^2 - R_c^2)^{.5} \tag{21}$$

as well as the fact that

$$\alpha = w(0, -h/2) - w(0, h/2) \tag{22}$$

The solution procedure is as follows. A unit p_o is invoked and the approach α is evaluated at the center of the plate (r equal to 0) as the difference between the deflections on the top surface and the bottom surface. The unit amplitude p_o is linearly scaled until the approach as determined in equation (22) is equal to that as determined in equation (21). The total load P is then determined by integration of the loading function, expressed in equation (19), over the contact surface to obtain:

$$P = \frac{2\pi R_c^2 p_o}{3} \tag{23}$$

The solution is now complete. For a given radius of indentation, the axisymmetric stresses and strains in an orthotropic plate may be determined based on the assumptions that the indentor is rigid and the constant stresses are Hertzian in nature. The solution is illustrated in the following numerical example.

5 Analytical/Experimental Comparisons

A $[0/+45/0/-45/0]_{2s}$ laminate of graphite/epoxy was analyzed to compare with test data generated on the same layup by Tan and Sun (1985). The assumed ply data based on Hercules AS1/3501-6 graphite/epoxy prepreg system (Lagace, 1982) is:

$$\begin{aligned} E_{11} &= 130.8 \text{ GPa} \\ E_{22} &= E_{33} = 10.6 \text{ GPa} \\ G_{12} &= G_{13} = 6.0 \text{ GPa} \\ \nu_{12} &= \nu_{13} = 0.28 \\ \nu_{23} &= 0.34 \\ G_{23} &= 3.9 \text{ GPa} \\ \text{Thickness} &= 0.134 \text{ mm} \end{aligned}$$

To determine the three-dimensional constitutive properties in an average sense, the three-dimensional stiffness matrix based on the above ply properties was rotated in the x-y plane, summed and averaged through the thickness. As the laminate is balanced and symmetric, the following average engineering properties based on the above procedure are:

$$\begin{aligned} E_{xx} &= 87.8 \text{ GPa} \\ E_{yy} &= 21.4 \text{ GPa} \\ E_{zz} &= 11.9 \text{ GPa} \\ G_{xz} &= 4.25 \text{ GPa} \\ G_{yz} &= 5.42 \text{ GPa} \\ \nu_{xy} &= 0.60 \\ \nu_{xz} &= 0.20 \\ \nu_{yz} &= 0.30 \end{aligned}$$

Note that the in-plane shear modulus, G_{xy}, does not enter into the problem from the condition of axisymmetry. The above plate is not transversely isotropic but using the properties generated along both axes (x and y), the analysis was conducted using x as the major axis and y as the minor axis. These properties were used in an axisymmetric fashion via equation (4). Under the transverse isotropy assumed in the present analysis, the r, θ constitutive properties correspond to the x and y direction depending on the in-plane (major or minor) axis used in the analysis. The current analysis was found to be

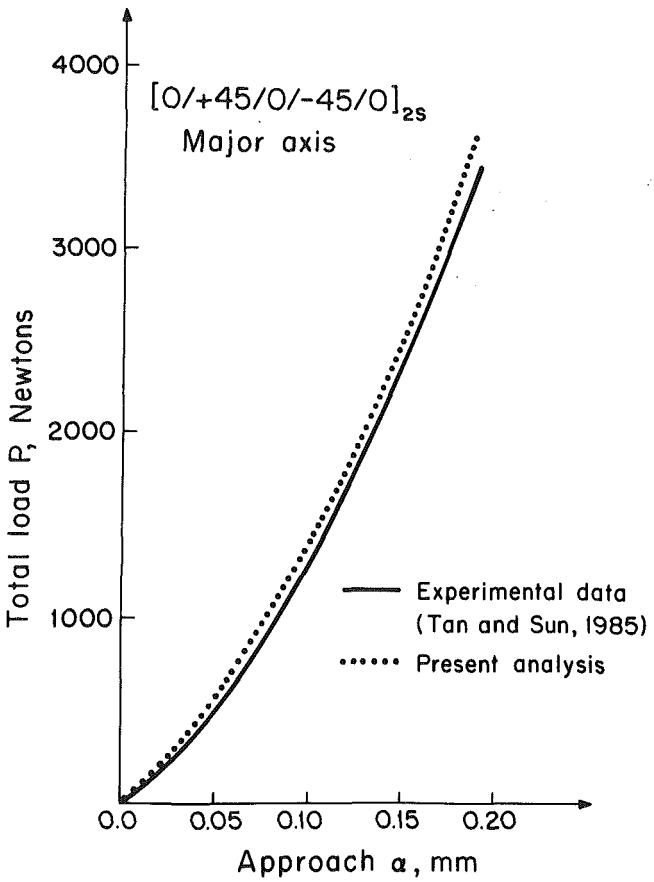


Fig. 4 Major axis contact load versus deflection

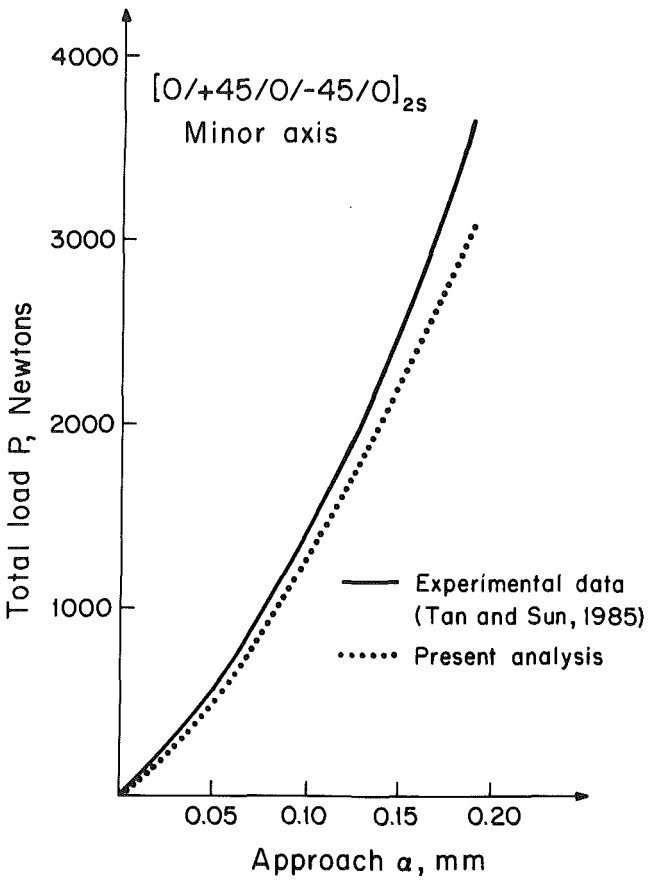


Fig. 5 Minor axis contact load versus deflection

much more dependent on the through-the-thickness properties than the in-plane properties. Hence, the analysis was controlled by the engineering constants E_{zz}, G_{xz}, and G_{yz}.

Using the technique outlined above, the case of a rigid 12.7 mm diameter sphere was used for comparison to data obtained in tests under the same conditions (Tan and Sun, 1985). Two types of local indentation tests were conducted on laminated beams with either direction (major or minor) along the beam axis. An aspect ratio of 0.05 (plate thickness/plate radius) was used for the analysis since the classical plate solution (Love, 1927) is recovered outside of these boundaries and only the very local deformations are of interest in the present problem.

Solutions for the major axis loading are compared with the experimental data generated by Tan and Sun (1985) in Fig. 4. Here, the nondimensional depth parameter ζ is defined as:

$$\zeta = (z + h/2)/h$$

Fifteen harmonics of the loading function $p(r)$ were used (equation (16)) such that oscillations of less than two percent in the loading, as expanded as a Fourier-Bessel series, were present. Excellent agreement is found between the analysis and the actual test data as the error is less than 1.5 percent. The agreement for the minor axis solution shown in Fig. 5 is not as close due to the geometric nonlinearities noted by Tan and Sun (1985). The low flexural modulus in the y direction caused more of the indentor to come in contact with the plate effectively reducing p_o and increasing R_c as a function of the load. The current model is incapable of handling such nonlinearities and is thus in error by approximately 8 percent from the test data. These results may, however, be adequate as input to other analyses depending upon the particular application.

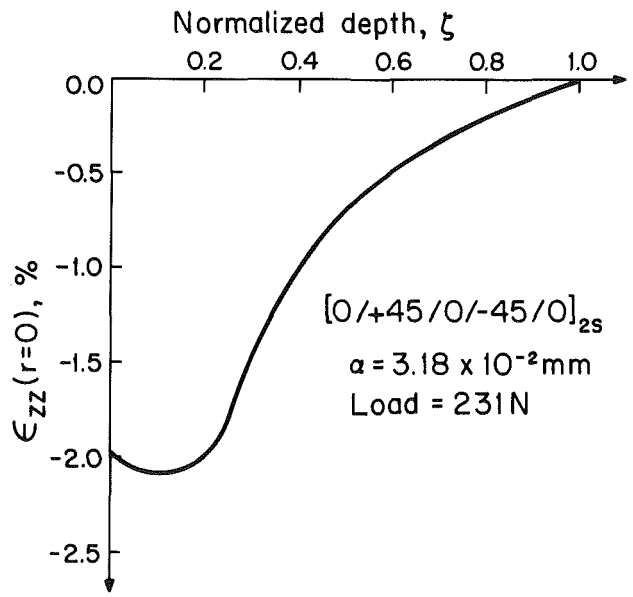


Fig. 6 Contact compressive strains (at r equal to 0) as a function of normalized depth

6 Contact Strains

The strains generated due to the contact problem are of particular interest in developing failure criteria. These strains are used to predict incipient damage due to the localized loading. The strains may be calculated via equations (3) and (7) and are presented for the previous example for an approach, α, of 0.032 mm.

In Fig. 6, the through-the-thickness strains ϵ_{zz} are plotted as

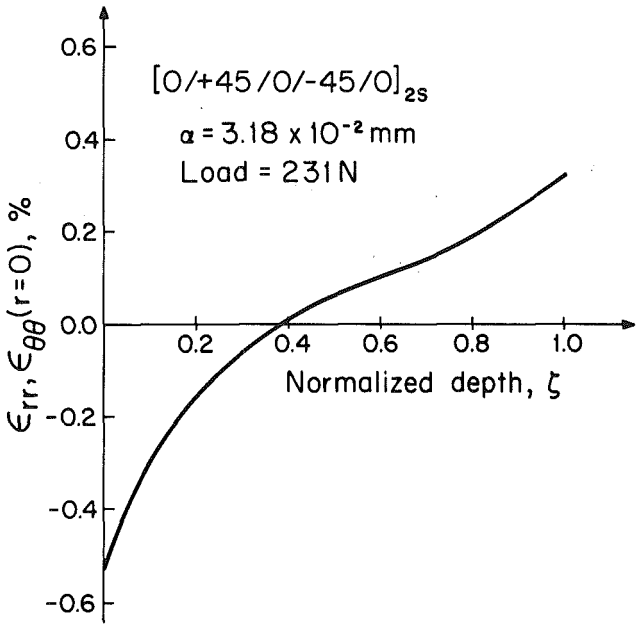


Fig. 7 Bending strains (at r equal to zero) as a function of normalized depth

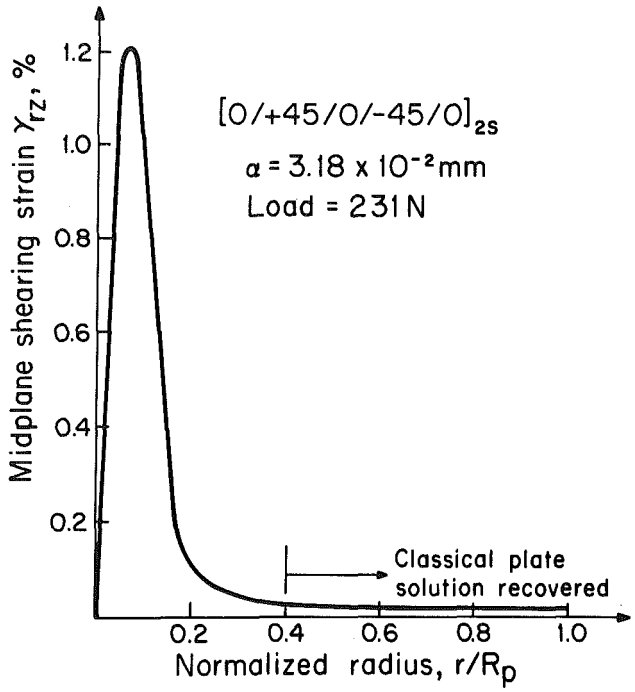


Fig. 8 Midplane shearing strain as a function of normalized radius

a function of depth at the center of the plate for a loading of 231 N at the center of the plate. The strains begin at a finite value and decay to zero at the bottom surface (t equal to $h/2$). Figure 7 is a plot of the bending strains at the center of the plate ($\sigma_{\theta\theta}$ is equal to σ_{rr} at r equal 0). The strains are compressive at the top surface and become tensile through the thickness. The neutral axis is not at the midplane of the laminate. It is also interesting to note that the bending strains are nonlinear through the thickness in the elasticity solution.

A plot of the shearing strains at the laminate midplane is shown in Fig. 8 as a function of radius. The shearing stress σ_{rz} is zero at r equal to zero as required by equations (7) and (14). This figure illustrates how quickly the local stresses in the contact region die out away from the point of load application. In fact, at r/a equal to 0.4, the standard plate solution is recovered and hence classical methods are valid outside of this region.

7 Additional Loadings

Not all loadings can be adequately represented by a Fourier-Bessel series and hence accurate solutions cannot be directly obtained. However, if the point load solution can be obtained, superposition may be used to obtain solutions for more general loadings.

The point load solution can be obtained by considering a loading of intensity p over radius c as shown in fig. 9, then the total load P is $p\pi c^2$. With the use of equation (16b) and taking the limit as the radius of contact c approaches zero, the coefficients β_m become:

$$\beta_m = \frac{P}{\pi J_1^2(u_m) R_p^2} \tag{25}$$

The same boundary conditions as employed previously are applicable such that the simultaneous equations for the unknown coefficients A_m, B_m, C_m, and D_m as presented in Appendix A may be solved. To obtain more general loadings this solution may be directly integrated using superposition for any loading function. This solution may also be used for sharp indentors which are not spherical as presumed in the preceding analyses.

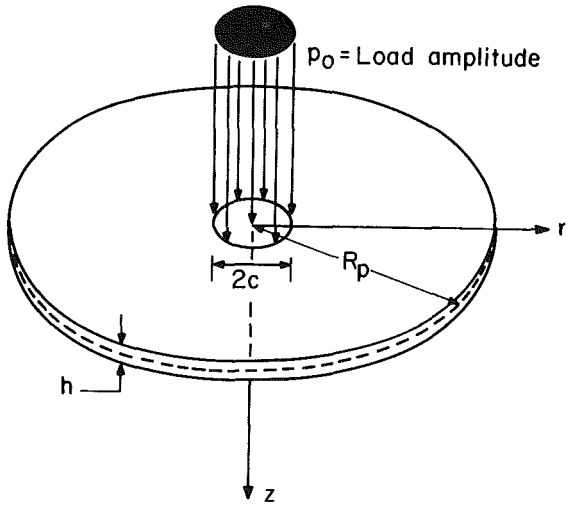


Fig. 9 Constant load over a finite area

8 Summary

A relatively straightforward and efficient approach has been presented to solve for the deflection and strains in a composite laminate subjected to lateral loading. Even with the assumption of homogeneity, the analysis agrees well with test data and can be used to predict localized strains which can then be utilized with appropriate failure criteria to predict damage initiation due to contact loading. The model cannot account for all of the geometric and material nonlinearities experienced during large deformations and loadings but does provide useful data to predict damage thresholds and damage extent due to contact or impact type problems. The present method offers advantages over finite element analyses in the linear range since it is a continuous and analytic solution. It is thus particularly useful in parametric-type studies in contact problems of orthotropic materials.

Acknowledgments

The authors wish to acknowledge the thoughtful comments of Dr. M. G. Worster of the mathematics department at the Massachusetts Institute of Technology. This work was supported by a joint Federal Aviation Administration/Navy program under contract no. N00019-85-C-0090.

References

Conway, H. D., 1956, *Zeitschrift fuer Angewandte Mathematik und Physik*, Vol. 7, pp. 460–465.

Greszczuk, L. B., 1975, "Response of Isotropic and Composite Materials to Particle Impact," *Foreign Object Impact Damage to Composites, ASTM STP 568*, American Society for Testing and Materials, pp. 183–208.

Lagace, P. A., 1982, "Static Tensile Fracture of Graphite/Epoxy," TELAC Report 82-4, Massachusetts Institute of Technology, Cambridge, MA.

Lal, K. M., 1983, "Low Velocity Transverse Impact Behavior of 8-Ply Graphite-Epoxy Laminates," *Journal of Reinforced Plastics and Composites*, Vol. 2, pp. 226-238.

Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Body*, Holden-Day, San Francisco, CA, pp. 350–351.

Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th Edition, Cambridge University Press, New York, p. 274, pp. 475–485.

Sankar, B. V., 1985, "Contact Law for Transversely Isotropic Materials," *Proceedings of 26th AIAA/ASME/ASCE/AHS Structures, Structural Dynamics and Materials Conference*, Orlando, FL, pp. 516–521.

Sun, C. T., and Chattopadhyay, S., 1975, "Dynamic Response of Anisotropic Laminates Under Initial Stress to Impact of a Mass," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 693–698.

Tan, T. M., and Sun, C. T., 1985, "Use of Statistical Indentation Laws for the Impact Analysis of Composite Plates, ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 6–12.

APPENDIX A

Solution for the Constants A_m, B_m, C_m, and D_m

The constants A_m, B_m, C_m, and D_m (in equation (15)) may be obtained by placing the expressions for ψ in the equations for σ_{zz} and σ_{rz} (equation (7)) at the mth harmonic.

$$\begin{aligned}\sigma_{zz} = \omega_m^3 \,[&(-cs_1 + ds_1^3)A_m e^{s_1\omega_m z} \\ &+ (-cs_2 + ds_2^3)\, B_m e^{s_2\omega_m z} + (ds_1 - ds_1^3)\, c_m e^{-s_1\omega_m z} \\ &+ (cs_2 - ds_2^3)\, D_m e^{-s_2\omega_m z}]\, J_o\,(\omega_m r)\end{aligned} \tag{A.1}$$

$$\begin{aligned}\sigma_{rz} = \omega_m^3 \,[&(1 + as_1^2)\, A_m e^{s_1\omega_m z} + (1 + as_2^2)\, B_m e^{s_2\omega_m z} \\ &+ (1 + as_1^2) c_m e^{-s_1\omega_m z} + (1 + as_2^2)\, D_m e \\ &^{-s_2\omega_m z}]\, J_1\,(\omega_m r)\end{aligned}$$

The boundary conditions expressed in equations (17) and (18) are invoked at each surface.

With the β_m known for any given loading, the coefficients A_m, B_m, C_m, and D_m are obtained by solving the simultaneous equations at each harmonic m:

$$\omega_m^3 \begin{bmatrix} (-cs_1 + ds_1^3)e^{-\omega_m s_1 h/2} & (-cs_2 + ds_2^3)e^{-\omega_m s_1 h/2} \\ (-cs_1 + ds_1^3)e^{-\omega_m s_i h/2} & (-cs_2 + ds_2^3)e^{\omega_m s_2 h/2} \\ (1 + as_1^2)e^{-\omega_m s_1 h/2} & (1 + as_2^3)e^{-\omega_m s_1 h/2} \\ (1 + as_1^2)e^{-\omega_m s_1 h/2} & (1 + as_2^2)e^{\omega_m s_2 h/2} \end{bmatrix}$$

$$\begin{bmatrix} (cs_1 - ds_2^3)e^{-\omega_m s_2 h/2} & (cs_2 - ds_2^3)e^{\omega_m s_2 h/2} \\ (cs_1 - ds_2^3)e^{-\omega_m s_1 h/2} & (cs_2 - ds_2^3)e^{-\omega_m s_2 h/2} \\ (1 + as_1^2)e^{\omega_m s_1 h/2} & (1 + as_2^2)e^{\omega_m s_2 h/2} \\ (1 + as_1^2)e^{\omega_m s_1 h/2} & (1 + as_2^2)e^{-\omega_m s_2 h/2} \end{bmatrix}$$

$$\begin{Bmatrix} A_m \\ B_m \\ C_m \\ D_m \end{Bmatrix} = \begin{Bmatrix} \beta_m \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{A.2}$$

Buckling of a Rectangular Frame Revisited

P. Seide
Department of Civil Engineering,
University of Southern California,
Los Angeles, CA 90089-0242

An investigation of the buckling under uniform beam load of a rectangular frame with columns restrained by linear rotational springs indicates that for certain ranges of bending stiffness ratio, length-height ratio, and support rotational stiffness parameter, the antisymmetrical bifurcation mode of buckling does not exist and buckling occurs at a symmetrical deformation limit load. The ranges of parameters for which this phenomenon may be important are studied.

Introduction

It is an accepted fact in Horne and Merchant (1965) that single story frames having members with primary bending moments will buckle in an antisymmetric mode at a load less than the symmetric limit load. The basis for this belief is found in the investigations of Chilver (1956), Chwalla (1938), and Lee (1963) of the behavior of frames with columns simply supported at the base. During the course of an investigation of the problem for clamped frames, results at variance with the above behavior were obtained. For the particular values of frame geometric and bending stiffness parameters investigated, no antisymmetric bifurcation point was found. Failure occurred when the symmetric load-deformation curve reached a limit point. It was decided then to redo the problem for a frame restrained at its base by linear springs (Fig. 1) and to investigate the variation of the critical load and mode of buckling with the various parameters. The derivation of the pertinent equations and a discussion of the calculated results are given herein.

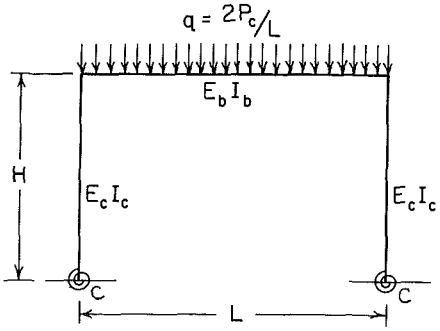


Fig. 1 Rectangular frame with elastic torsional restraint

Derivation

1 General Stiffness Relationships. The end load-end deformation relations for an axially loaded member of uniform bending stiffness can be expressed in the following form (see Fig. 2 and Horne and Merchant, 1965)

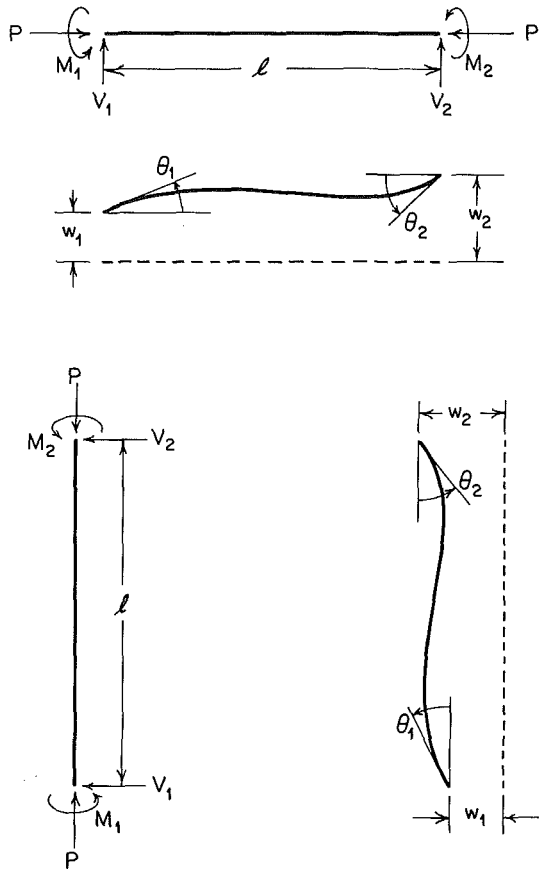


Fig. 2 Positive directions of forces and moments: (*a*) horizontal beam; (*b*) vertical column

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October 22, 1985; final revision, February 23, 1987.

$$\begin{Bmatrix} V_1 \\ M_1 \\ V_2 \\ M_2 \end{Bmatrix} = \begin{bmatrix} \frac{EI}{\ell^3}\phi_1(j) & \frac{EI}{\ell^2}\phi_2(j) & \frac{-EI}{\ell^3}\phi_1(j) & \frac{EI}{\ell^2}\phi_2(j) \\ \frac{EI}{\ell^2}\phi_2(j) & \frac{EI}{\ell}\phi_3(j) & -\frac{EI}{\ell^2}\phi_2(j) & \frac{EI}{\ell}\phi_4(j) \\ \frac{-EI}{\ell^3}\phi_1(j) & \frac{-EI}{\ell^2}\phi_2(j) & \frac{EI}{\ell^2}\phi_1(j) & -\frac{EI}{\ell^2}\phi_2(j) \\ \frac{EI}{\ell^2}\phi_2(j) & \frac{EI}{\ell}\phi_4(j) & \frac{-EI}{\ell^2}\phi_2(j) & \frac{EI}{\ell}\phi_3(j) \end{bmatrix} \begin{Bmatrix} W_1 \\ \theta_1 \\ W_2 \\ \theta_2 \end{Bmatrix} \quad (1)$$

where

$$j = \frac{P\ell^2}{\pi^2 EI} \quad (2a)$$

$$\phi_1(j) = \frac{(\pi\sqrt{j})^3}{\Delta} = \phi_3^2(j) - \phi_4^2(j) = 2\phi_2(j) - \pi^2 j \quad (2b)$$

$$\phi_2(j) = \frac{(\pi\sqrt{j})^2 \tan\frac{\pi}{2}\sqrt{j}}{\Delta} = \phi_3(j) + \phi_4(j) \quad (2c)$$

$$\phi_3(j) = \frac{\pi\sqrt{j}(1 - \pi\sqrt{j}\cot\pi\sqrt{j})}{\Delta} \quad (2d)$$

$$\phi_4(j) = \frac{\pi\sqrt{j}(\pi\sqrt{j}\csc\pi\sqrt{j} - 1)}{\Delta} \quad (2e)$$

$$\Delta = 2\left(\tan\frac{\pi}{2}\sqrt{j} - \frac{\pi}{2}\sqrt{j}\right). \quad (2f)$$

If the member is restrained by a linear rotational spring of stiffness C at end 1 so that

$$W_1 = 0 \quad (3a)$$

$$M_1 = -C\theta_1 \quad (3b)$$

the pertinent and load-end deformation relationships may be determined as

$$\begin{Bmatrix} V_2 \\ M_2 \end{Bmatrix} = \frac{1}{1 + \frac{EI}{\ell C}\phi_3(j)} \begin{bmatrix} \frac{EI}{\ell^3}\left\{\phi_1(j) + \frac{EI}{\ell C}[\phi_1(j_1)\phi_3(j) - \phi_2^2(j)]\right\} & -\frac{EI}{\ell^2}\left[\phi_2(j) + \frac{EI}{\ell C}\phi_1(j)\right] \\ -\frac{EI}{\ell^2}\left[\phi_2(j) + \frac{EI}{\ell C}\phi_1(j)\right] & \frac{EI}{\ell}\left\{\phi_3(j) + \frac{EI}{\ell C}\phi_1(j)\right] \end{bmatrix} \begin{Bmatrix} W_2 \\ \theta_2 \end{Bmatrix} \quad (4)$$

The cases of frames simply supported or clamped at the base are obtained by putting C equal to zero or infinity in the equations.

Finally, if an axially compressed member is subjected to uniform transverse load q, the fixed-end reactions are given by

$$\begin{Bmatrix} V_{1f} \\ M_{1f} \\ V_{2f} \\ M_{2f} \end{Bmatrix} = \begin{Bmatrix} q\frac{\ell}{2} \\ m_1(j)q\ell^2 \\ q\frac{\ell}{2} \\ -m_1(j)q\ell^2 \end{Bmatrix} \quad (5)$$

where

$$m_1(j) = \frac{1}{2\phi_2(j)}. \quad (6)$$

2 Symmetric Frame Deformations. In analyzing the symmetric deformation of the frame, let the magnitude of the uniform load on the horizontal member be taken as $2P_c/L$ so that the load in each vertical column is P_c. Since the distortion is assumed symmetrical about the center line, the equilibrium of only one of the joints need be considered. The end displacement of the elastically restrained vertical column on the left at the upper joint is a rotation $-\theta$ (Fig. 3), so that the moment in the column at this joint is given by equations (4) as

$$M_{c2} = -\frac{\phi_3(j_c) + \frac{E_cI_c}{HC}\phi_1(j_c)}{1 + \frac{E_cI_c}{HC}\phi_3(j_c)}\frac{E_cI_c}{H}\theta \quad (7)$$

with

$$j_c = \frac{P_cH^2}{\pi^2 E_cI_c} \quad (8)$$

The compressive axial load in the horizontal beam is unknown and will be denoted by P_b. The moment in the beam at the left joint is the sum of the fixed-end moment due to the uniform load and the moment due to end rotation such that

$$W_1 = W_2 = 0 \quad (9a)$$

$$\theta_1 = -\theta_2 = -\theta. \quad (9b)$$

Thus equations (1) and (5) yield

$$M_{b1} = \frac{P_cL}{\phi_2(j_b)} - \frac{E_bI_b}{L}[\phi_3(j_b) - \phi_4(j_b)]\theta \quad (10)$$

with

$$j_b = \frac{P_bL^2}{\pi^2 E_bI_b}. \quad (11)$$

The total moment applied to the joint is given by

$$M = M_{c2} + M_{b1} = \frac{E_cI_c}{H}\left(\left(-\left\{\frac{\phi_3(j_c) + \frac{E_cI_c}{HC}\phi_1(j_c)}{1 + \frac{E_cI_c}{HC}\phi_3(j_c)} + \frac{E_bI_b/L}{E_cI_c/H}[\phi_3(j_b) - \phi_4(j_b)]\right\}\theta + \frac{\pi^2 j_c}{\phi_2(j_b)}\frac{L}{H}\right)\right) = 0 \quad (12)$$

from which the angle of rotation may be obtained.

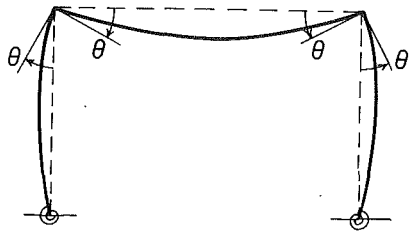


Fig. 3 Symmetrical deformed shape

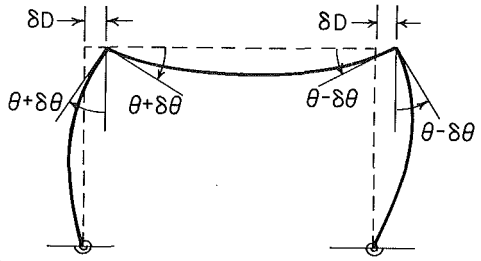


Fig. 4 Deformed shape with antisymmetric perturbations

The axial load in the beam is equal in magnitude but opposite in direction to the shear force in the column at the joint. With the beam axial load considered positive in compression, the use of the first of equations (4) and equation (11) and the substitution of the expression for θ given by equation (12) yields the following relation between j_b and j_c

$$j_b = j_c \frac{\dfrac{E_cI_c}{E_bI_b}\left(\dfrac{L}{H}\right)^3 \dfrac{\phi_2(j_c)+\dfrac{E_cI_c}{HC}\phi_1(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)} \dfrac{1}{\phi_2(j_b)}}{\dfrac{\phi_3(j_c)+\dfrac{E_cI_c}{HC}\phi_1(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)} + \dfrac{E_bI_b/L}{E_cI_c/H}[\phi_3(j_b)-\phi_4(j_b)]} \tag{13}$$

3 Infinitesimal Antisymmetric Deformations of the Frame. In order to investigate the possibility of a bifurcation of the symmetric deformation state, assume that for a given distributed load additional infinitesimal antisymmetrical deformations of the frame occur, i.e., a sideways motion $-\delta D$ and equal joint rotations $-\delta\theta$ (Fig. 5). The column axial load parameters change by equal and opposite amounts since the vertical load on the beam remains constant. The use of equations (4) then yields the shear in each column at the upper joints, correct to first order terms, as

$$\left.\begin{matrix} V_{c2}^L \\ V_{c2}^R \end{matrix}\right\} = \frac{E_cI_c/H^2}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}\left(\left(\pm\left[\phi_2(j_c)+\frac{E_cI_c}{HC}\phi_1(j_c)\right]\theta\right.\right.$$
$$-\left\{\phi_1(j_c)+\frac{E_cI_c}{HC}[\phi_1(j_c)\phi_3(j_c)-\phi_2^2(j_c)]\right\}\frac{\delta D}{H}$$
$$+\left[\phi_2(j_c)+\frac{E_cI_c}{HC}\phi_1(j_c)\right]\delta\theta+\left\{\phi_2'(j_c)+\frac{E_cI_c}{HC}\phi_1'(j_c)\right.$$
$$\left.\left.-\frac{\dfrac{E_cI_c}{HC}\phi_3'(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}\left[\phi_2(j_c)+\frac{E_cI_c}{HC}\phi_1(j_c)\right]\right\}\theta\delta j_c\right)\right) \tag{14}$$

where the upper sign refers to the left column, the lower to the right column, and primes indicate differentiation with respect to the argument. The column shear forces change by equal amounts indicating that the beam axial load is unchanged.

The joint moment in the columns and in the beam are given by equations (4) and equations (1) correct to first order terms as

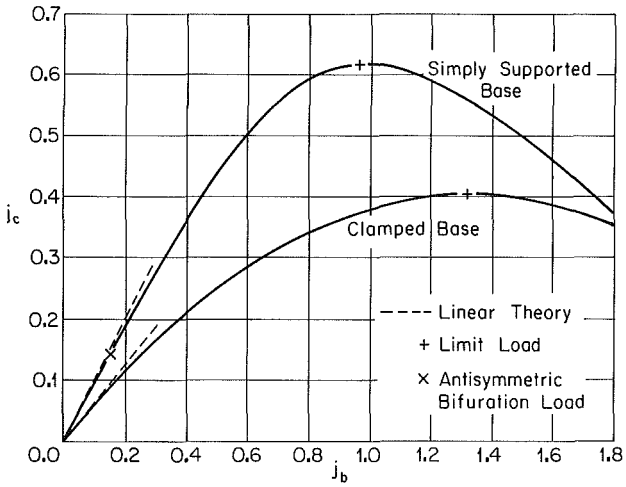


Fig. 5 Variation of column axial load with beam axial load ($E_bI_b/E_cI_c = 1$, $L/H = 2$)

$$\left.\begin{matrix} M_{c2}^L \\ M_{c2}^R \end{matrix}\right\} = \frac{E_cI_c/H}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}\left(\left(\mp\left[\phi_3(j_c)+\frac{E_cI_c}{HC}\phi_1(j_c)\right]\theta\right.\right.$$
$$+\left[\phi_2(j_c)+\frac{E_cI_c}{HC}\phi_1(j_c)\right]\frac{\delta D}{H}-\left[\phi_3(j_c)+\frac{E_cI_c}{HC}\phi_1(j_c)\right]\delta\theta$$
$$-\left\{\phi_3'(j_c)+\frac{E_cI_c}{HC}\phi_1'(j_c)\right.$$
$$\left.\left.-\frac{\dfrac{E_cI_c}{HC}\phi_3'(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}\left[\phi_3(j_c)+\frac{E_cI_c}{HC}\phi_1(j_c)\right]\right\}\theta\delta j_c\right)\right) \tag{15}$$

$$\left.\begin{matrix} M_{b1} \\ M_{b2} \end{matrix}\right\} = \mp\left\{\frac{E_bI_b}{L}[\phi_3(j_b)-\phi_4(j_b)]\theta\right.$$
$$\left.-\frac{P_cL}{\phi_2(j_b)}\right\}-\frac{E_bI_b}{L}\phi_2(j_b)\delta\theta. \tag{16}$$

Finally, the change in axial compressive load in the columns is equal to the change of end shear in the beam. Then

$$\delta j_c = -\frac{2}{\pi^2}\frac{E_bI_b}{E_cI_c}\left(\frac{H}{L}\right)^2\phi_2(j_b)\delta\theta. \tag{17}$$

The infinitesimal horizontal load required to be applied to the frame to provide the infinitesimal antisymmetrical displacement is the sum of the column shear forces. Thus

$$\delta V = -(V_{c2}^L + V_{c2}^R) \tag{18a}$$

while the infinitesimal joint moments required to produce the infinitesimal antisymmetric deformations are given by virtue of equation (12), as

$$\delta M = -(M_{c2}^L + M_{b1}) = -(M_{c2}^R + M_{b2}). \tag{18b}$$

The use of equations (14) to (17) yield the infinitesimal antisymmetric load-deformation relations as

$$\begin{Bmatrix} \dfrac{H^2\delta V}{2E_cI_c} \\ \\ \dfrac{H\delta M}{E_cI_c} \end{Bmatrix} =$$

$$\begin{bmatrix} \dfrac{\phi_1(j_c)+\dfrac{E_cI_c}{HC}[\phi_1(j_c)\phi_3(j_c)-\phi_2^2(j_c)]}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)} & -\left\{\dfrac{\phi_2(j_c)+\dfrac{E_cI_c}{HC}\phi_1(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}-\dfrac{2}{\pi^2}\dfrac{E_bI_b/L^2}{E_cI_c/H^2}\left[\dfrac{\phi_2(j_c)+\dfrac{E_cI_c}{HC}\phi_1(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}\right]'\phi_2(j_b)\theta\right\} \\ \\ -\dfrac{\phi_2(j_c)+\dfrac{E_cI_c}{HC}\phi_1(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)} & \dfrac{\phi_3(j_c)+\dfrac{E_cI_c}{HC}\phi_1(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}+\dfrac{E_bI_b/L}{E_cI_c/H}\left\{1-\dfrac{2}{\pi^2}\dfrac{H}{L}\left[\dfrac{\phi_3(j_c)+\dfrac{E_cI_c}{HC}\phi_1(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}\right]'\theta\right\}\phi_2(j_b) \end{bmatrix}$$

$$\begin{Bmatrix} \delta D/H \\ \\ \delta\theta \end{Bmatrix} \tag{19}$$

where primes indicate differentiation with respect to the argument j_c and θ is given by equation (12). Buckling in an antisymmetric mode is possible if the determinant of the stiffness matrix of equations (19) vanishes, for then the frame may be in equilibrium with small antisymmetric displacements without the imposition of antisymmetric loads.

4 Buckling of Frames With Loads Applied at the Joints. Results for a frame elastically restrained at the base and with the vertical load applied only at the joints may be obtained from the previous equations. In this case j_b, θ, and the fixed-end beam moments vanish prior to buckling. The criterion for buckling in a symmetric mode is obtained from equation (12) by setting the infinitesimal symmetric joint stiffness $dM/d\theta$ equal to zero to yield

$$\frac{\phi_3(j_c)+\dfrac{E_cI_c}{HC}\phi_1(j_c)}{1+\dfrac{E_cI_c}{HC}\phi_3(j_c)}+2\frac{E_bI_b/L}{E_cI_c/H}=0. \tag{20}$$

For buckling in an antisymmetric mode setting the determinant of the resulting stiffness matrix of equation (19) equal to zero yields after much manipulation

$$\frac{\cot\pi\sqrt{j_c}-\dfrac{E_cI_c}{HC}\pi\sqrt{j_c}}{\pi\sqrt{j_c}\left(1+\dfrac{E_cI_c}{HC}\pi\sqrt{j_c}\cot\pi\sqrt{j_c}\right)}+6\frac{E_bI_b/L}{E_cI_c/H}=0 \tag{21}$$

equations (20) and (21) have been obtained previously in a somewhat different form by Galambos (1960) and by Appeltauer and Barth (1961).

Results and Discussion

Calculations were first carried out for simply supported and clamped frames having the relative stiffnesses and dimensions

$$\frac{E_bI_b}{E_cI_c}=1$$

$$\frac{L}{H}=2.$$

These are the values treated by Horne and Merchant (1965). Equation (13) was solved, with the use of computer code DRTMI of Anon (1970), for the lowest values of j_c corresponding to a range of values of j_b. This scheme was adopted since for both types of support the values of j_c increase to a maximum as j_b increases and then decrease, as shown in Fig. 5. The maximum values of j_c represent critical loads for symmetric deformation since the frame cannot be in equilibrium for larger values of loading. The maximum values of j_c obtained are 0.6174 for the simply supported frame and 0.4044 for the clamped frame. The anomaly of the ostensibly stiffer structure having the lower symmetric critical load may be attributed to the clamped frame having a much larger beam axial load.

Another interesting phenomenon is revealed by an examination of the antisymmetric bifurcation point. For each set of values j_b and j_c, corresponding values of the determinant of the stiffness matrix of equation (19) were calculated. For the simply supported frame the determinant vanishes at a value of j_c of 0.1409 whereas for the clamped frame there is no value of j_c for which the determinant vanishes. Thus the symmetric limit load is the buckling load of the clamped frame while the simply supported frame buckles in an antisymmetric mode.

The effect of finite rotational restraint on the critical loads of the same frame is shown in Fig. 6. Calculations of the symmetric limit load and the antisymmetric bifurcation load were made and indicate that the symmetric limit load decreases with increasing support stiffness whereas the antisymmetric bifurcation load increases. The two loads coincide for a value of support spring flexibility parameter E_cI_c/HC of about 0.18, below which the antisymmetric bifurcation point ceases to exist. The frame thus buckles in an antisymmetric mode with in-

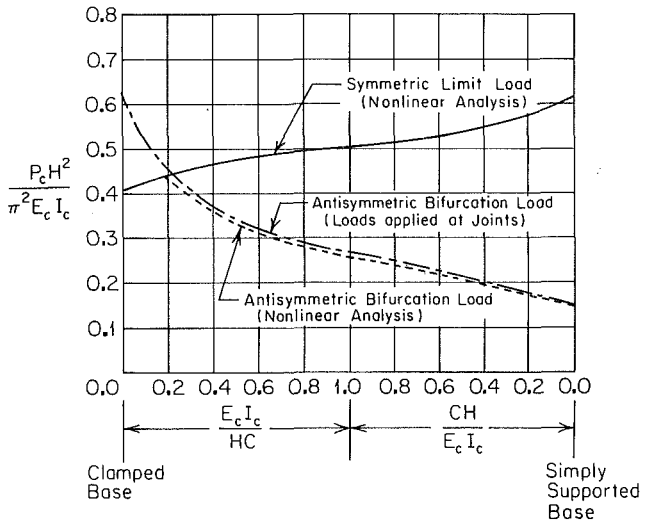


Fig. 6 Variation of critical loads with support torsional stiffness ($E_b I_b/E_c I_c = 1$, $L/H = 2$)

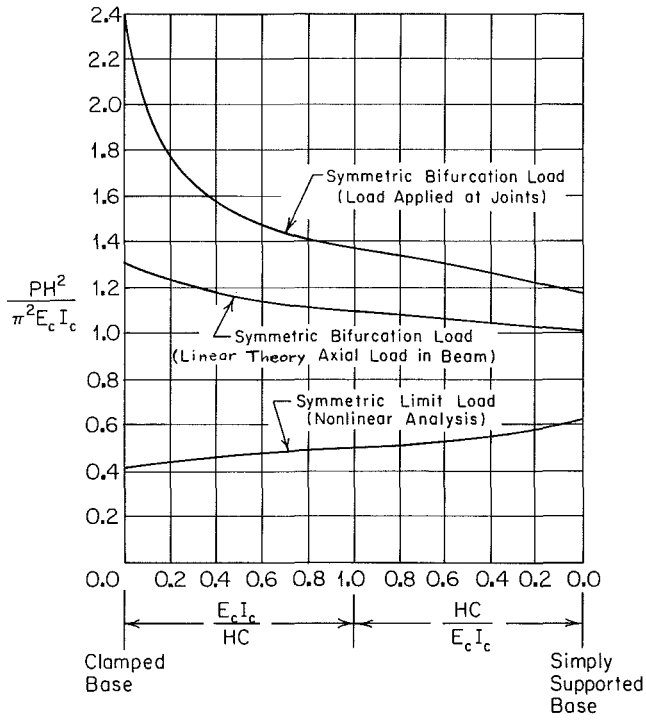


Fig. 7 Comparison of various methods of calculation of symmetric critical load ($E_b I_b/E_c I_c = 1$, $L/H = 2$)

creasing load as the support rotational stiffness increases until a critical value of stiffness is reached. For larger values of support stiffness buckling occurs in a symmetric mode with *decreasing* critical load.

Also shown in Fig. 6 is the antisymmetric buckling load with the load applied at the joints (equation (21)) which is reasonably close to the values obtained from the nonlinear analysis. This is not the case with the symmetric buckling load with the load applied at the joints. The results shown in Fig. 7 indicate that the discrepancy between the values obtained from equation (20) and the symmetric limit loads given by the nonlinear analysis increases drastically as the support rotational stiffness increases. For the simply supported frame the ratio of the two loads is 1.90 whereas for the clamped frame the ratio is 5.75. The buckling loads obtained using the values

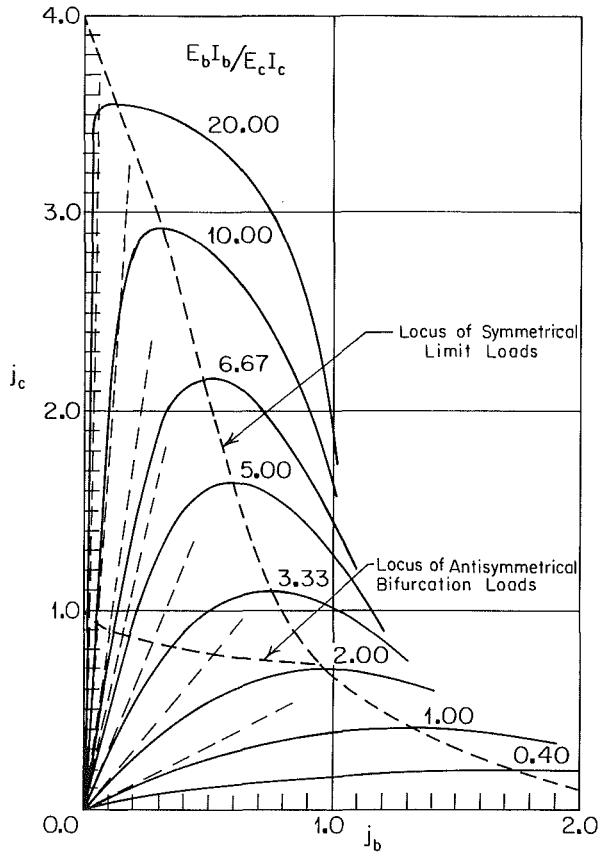


Fig. 8 Variation of column and beam axial loads and antisymmetric bifurcation load with bending stiffness ratio ($L/H = 2$, $C = \infty$)

of beam axial load given by linear theory and assuming the beam and column loads to be applied at the joints are shown in the figure by the middle curve. These values are closer to the results of the nonlinear analysis but are still considerably in error and fail to predict the decrease of critical symmetric load with increasing support rotational stiffness.

The foregoing investigation has indicated a change of buckling mode from antisymmetric bifurcation to a symmetric limit point with increasing support torsional stiffness for a particular class of frames. To determine whether this is always the case, a series of calculations were carried out for clamped frames with a value of L/H of 2 and varying values of bending stiffness ratio. The results of the solution of equation (13) are shown in Fig. 8 together with the values for which the determinant of the stiffness matrix of equation (19) vanishes. The results indicate that the symmetric limit load is critical for values of bending stiffness ratio $E_b I_b/E_c I_c$ less than about 2.1, while the antisymmetric bifurcation load is critical for larger values. Thus the range of frame parameters where a change of buckling mode occurs is limited. The effect of various approximations is shown in Fig. 9, which indicates, as before, that the antisymmetric bifurcation load, where it exists, is affected very little by prebuckling deformations whereas the symmetric limit load can be calculated accurately only by means of a nonlinear analysis.

The complete range of parameters for which the symmetric limit load is critical is indicated by the results of Fig. 10 where symmetric limit loads and antisymmetric bifurcation loads are shown for clamped frames having various bending stiffness and dimension ratios. If the symmetric limit load predominates for a clamped frame it may be critical for an elastically supported frame and should be investigated. If, however, antisymmetric buckling predominates for a clamped frame it will be predominant for an elastically supported

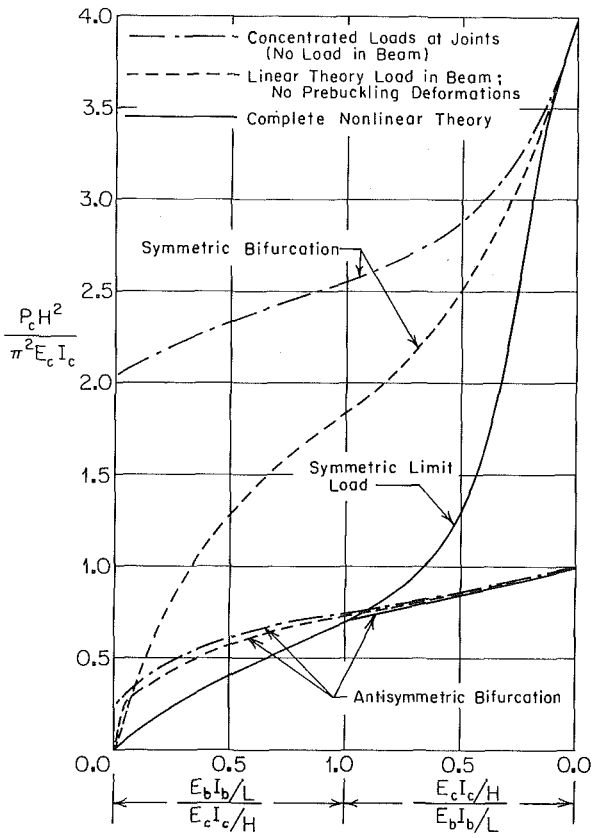


Fig. 9 Comparison of various methods of calculation of symmetric and antisymmetric critical loads ($L/H = 2$, $C = \infty$)

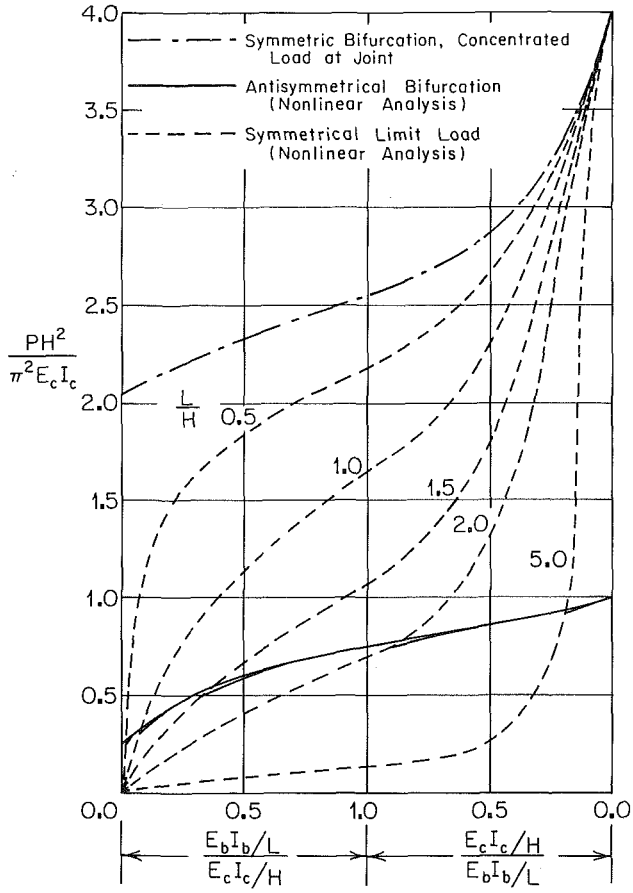


Fig. 10 Symmetric and antisymmetric critical loads for clamped frames

frame and can be calculated with reasonable accuracy by neglecting prebuckling deformations. It can be seen, however, that generally symmetrical buckling is important only for frames for which the beam bending stiffness is less than the column bending stiffness. The maximum stiffness ratio $E_b I_b / E_c I_c$ for which symmetrical buckling predominates decreases with decreasing length-height ratio from a value of about 1.0 for L/H of 5 to a value of about 0.1 for L/H of 1. It should also be noted that the neglect of support fixity by the assumption of simple support is always conservative. For simply supported frames, antisymmetrical buckling predominates and is satisfactorily calculated by assuming beam and column loads to be given by the results of linear theory.

References

Anon., 1970, Programmer's Manual (GH20-0205-4), System/360 Scientific Subrouting Package (360A-CM-03X).

Appeltauer, J. W., and Barta, T. A., 1961, Discussion of "Influence of Partial Base Fixity on Frame Stability," by T. V. Galambos, *ASCE J. Struct. Div.*, Vol. 87, pp. 29–38.

Chilver, A. H., 1956, "Buckling of a Simple Portal Frame," *J. Mech. Phys. Solids*, Vol. 5, pp. 18–25.

Chwalla, E., 1938, "Die Stabilitat Lotrecht Belasteter Rechteckrahmen," *Der Bauingenieur*, Vol. 19, p. 69.

Galambos, T. V., 1960, "Influence of Partial Base Fixity of Frame Stability," *ASCE J. Struct. Div.*, Vol. 86, pp. 85–108.

Horne, M. R., and Merchant, W., 1965, *Instability of Frames*, Pergamon Press, pp. 115–122.

Lu, L.-W., 1963, "Stability of Frames Under Primary Bending Moments," *ASCE J. Struct. Div.*, Vol. 89, pp. 35–62.

A Small Strain and Moderate Rotation Theory of Elastic Anisotropic Plates

J. N. Reddy
Clifton C. Garvin Professor,
Department of Engineering Science and Mechanics,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061

A general nonlinear theory for the dynamics of elastic anisotropic plates that accounts for transverse shear strains and moderate rotations is presented. The theory contains, as special cases, the von Kármán classical plate theory, the first-order shear deformation theory (i.e., the Reissner-Mindlin plate theory) and the third-order shear deformation plate theory. The theory is characterized, even for isotropic plates, by strong coupling between various equations of motion.

1 Introduction

In geometrically nonlinear theories of elastic anisotropic plates one often assumes that the strains and rotations about the normal to the midplane are infinitesimal and retains the products and squares of the derivatives of the transverse deflection in the strain-displacement equations (the von Kármán assumption; see Medwadowski, 1958; Ebcioglu, 1964; Chia, 1980; Reddy, 1983 and 1984a). The full geometric nonlinearity (implied by the strain-displacement equations of nonlinear elasticity) in shell theories was considered by Naghdi (1972), Librescu (1975), Yokoo and Matsunaga (1974), Habip (1966), and Pietraszkiewicz (1979), among others. Consideration of full geometric nonlinearity not only results in complex equations, but is not warranted in most practical problems. On the other hand, the von Kármán nonlinear theory does not account for all moderate rotation terms that could be of significance in the analysis (especially in stability problems) of plates. The small strain and moderate rotation concept was used in the classical theory of plates and shells (in which the transverse strains are neglected) by Sanders (1963), Koiter (1966), Reissner (1958), Pietraszkiewicz (1980, 1984), and Schmidt (1984) and in refined plate and shell theories by Wempner (1973), Naghdi and Vangsarnpigoon (1983), and Librescu and Schmidt (1986). Additional large rotation theories can be found in the works of Ranjan and Steele (1980), Kayuk and Sakhastskii (1985), Nolte et al. (1986), and Iura (1986).

The present study deals with a new higher-order theory of anisotropic plates that accounts for transverse shear strains and moderate rotations. The theory is a generalization of the classical plate theory, the first-order shear deformation (i.e., Reissner-Mindlin) plate theory, the refined theory of Kromm (1953, 1955), and the third-order shear deformation theories of Reddy (1984a, b). The theory is based on an assumed displacement field and orders of magnitudes of linear strains and rotations. The associated strain-displacement equations are presented and the equations of motion are derived using the principle of virtual work. Specialization of the equations of motion for various well-known existing theories is demonstrated. A detailed account of the strain-displacement equations, constitutive equations and the equations of motion of the first-order shear deformation theory with moderate rotations is presented. The latter should aid the development of the finite element models of the theory.

2 Notation and Basic Assumptions

Points of a three-dimensional continuum V are denoted by their orthogonal curvilinear coordinates $\mathbf{x}=(x^1,x^2,x^3)$. Covariant and contravariant base vectors at points of the continuum are denoted by $\mathbf{g}_i$ and $\mathbf{g}^i$, respectively. Latin indices are assumed to have values 1, 2, 3, and the Greek indices have values 1, 2. The plate continuum in the undeformed configuration is defined by the Cartesian product of points in the midplane Ω and the normal $[-h/2, h/2]$:

$$V=\Omega\times\left[-\frac{h}{2},\frac{h}{2}\right]$$

where h denotes the constant thickness of the plate. Let x^α denote the curvilinear inplane coordinates and x^3 be the normal to Ω. The metric tensor components of Ω are denoted by

$$g_{\alpha\beta}=\mathbf{g}_\alpha\cdot\mathbf{g}_\beta,\ g^{\alpha\beta}=\mathbf{g}^\alpha\cdot\mathbf{g}^\beta,\ g^{33}=g_{33}=1$$
$$\mathbf{g}_\alpha=\frac{\partial\mathbf{r}}{\partial x^\alpha},\ \mathbf{g}_\alpha\cdot\mathbf{g}^\beta=\delta_\alpha^\beta,\ \mathbf{g}_3=\mathbf{n} \tag{1}$$

where $\mathbf{r}$ is the position vector of a particle (x^α, x^3) at time t, δ_α^β is the Kronecker delta, and $\mathbf{n}$ is the unit normal to the boundary of Ω.

The displacement vector of a point in the plate at time t is of the form

$$\mathbf{u}=u^\alpha\mathbf{g}_\alpha+u^3\mathbf{n}=u_\alpha\mathbf{g}^\alpha+u_3\mathbf{n} \tag{2}$$

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 12, 1986; final revision, December 24, 1986.

Paper No. 87-WA/APM-5.

where the Einstein summation convention on repeated subscripts is assumed. The covariant components of the Green-Lagrange strain tensor are given by

$$\epsilon_{ij} = \frac{1}{2}(u_{i|j} + u_{j|i} + u_{m|i}u^{m}{}_{|j}) \tag{3}$$

where a vertical line denotes covariant differentiation (see Reddy and Rasmussen, 1982). The strain components ϵ_{ij} can be expressed in terms of the linearized strains e_{ij} and rotations ω_{ij} as

$$\epsilon_{ij} = e_{ij} + \frac{1}{2}e_{mi}e^{m}_{j} + \frac{1}{2}(e_{mi}\omega^{m}_{j} + e_{mj}\omega^{m}_{i}) + \frac{1}{2}\omega_{mi}\omega^{m}_{j} \tag{4}$$

where

$$e_{ij} = \frac{1}{2}(u_{i|j} + u_{j|i}), \quad \omega_{ij} = \frac{1}{2}(u_{i|j} - u_{j|i}) \tag{5}$$

We now assume that the strains ϵ_{ij} and rotations ω_{ij} are of the following magnitude (cf Librescu and Schmidt, 1986):

$$\epsilon_{ij} = 0(\theta^2), \quad \omega_{\alpha\beta} = 0(\theta^2), \quad \omega_{\alpha 3} = 0(\theta), \quad \theta << 1 \tag{6}$$

Equation (6) implies that the strains and the rotations about the normal to the midplane are small, and that the rotations of a normal to the midplane are moderate. Such assumptions are justified in view of the large inplane rigidity and transverse flexibility of composite laminates.

Neglecting terms of order (θ^4) and higher in the strain-displacement equations (4), we obtain (cf Librescu and Schmidt, 1986):

$$\begin{aligned}
\epsilon_{\alpha\beta} &= e_{\alpha\beta} + \frac{1}{2}\underline{(e_{3\alpha}\omega^{3}_{\beta} + e_{3\beta}\omega^{3}_{\alpha})} + \frac{1}{2}\omega_{3\alpha}\omega^{3}_{\beta} \\
\epsilon_{\alpha 3} &= e_{\alpha 3} + \frac{1}{2}\underline{(e_{\lambda\alpha}\omega^{\lambda}_{3} + e_{33}\omega^{3}_{\alpha})} + \frac{1}{2}\underline{\omega_{\lambda\alpha}\omega^{\lambda}_{3}} \\
\epsilon_{33} &= e_{33} + \underline{e_{\lambda 3}\omega^{\lambda}_{3}} + \frac{1}{2}\omega_{\lambda 3}\omega^{\lambda}_{3}
\end{aligned} \tag{7}$$

where the underlined terms are of order (θ^3).

3 Displacements and Strains

The present theory is based on the following assumed variation of the displacement components across the plate thickness:

$$\begin{aligned}
u_{\alpha}(x^{\beta}, x^{3}, t) &= u^{o}_{\alpha}(x^{\beta}, t) - x^{3}u^{o}_{3|\alpha} + f(x^{3})u^{1}_{\alpha}(x^{\beta}, t) \\
u_{3}(x^{\beta}, x^{3}, t) &= u^{o}_{3}(x^{\beta}, t) + \hat{u}^{o}_{3}(x^{\beta}, t)
\end{aligned} \tag{8}$$

where f is a specified function of the thickness coordinate x^3. Note that the transverse deflection is assumed to be independent of x^3 and consists of two parts, one due to bending and the other due to transverse shear. The separation of the transverse displacement into two parts allows the representation of nonvanishing shear strain at clamped edges. When $\hat{u}^{0}_{3}$ is zero, then the linear portion of ϵ_{13} vanishes at a clamped edge because $u^{o}_{3,\alpha} = u^{1}_{\alpha} = 0$ there (cf Huffington, Jr., 1963; Krishna Murty, 1986). The particular form of displacement field is assumed in order to include the displacement fields of the classical plate theory (set $\hat{u}^{o}_{3} = 0$ and $u^{1}_{\alpha} = 0$), the first-order shear deformation theory (set $u^{o}_{3} = 0$ and $f(x^3) = x^3[1 - 4/3(x^3/h)^2]$), among others. Note that in all derivations presented here no special form of $f(x^3)$ is used, to keep the generality of the theory.

For the displacement field in equation (8), the strains for the moderate rotation theory become (consistent with the assumptions in equation (6)),

$$\begin{aligned}
\epsilon_{\alpha\beta} &= \epsilon^{o}_{\alpha\beta} + x^{3}\epsilon^{1}_{\alpha\beta} + f\kappa_{\alpha\beta} \\
\epsilon_{\alpha 3} &= \epsilon^{o}_{\alpha 3} + g\hat{\epsilon}^{o}_{\alpha 3} + x^{3}\kappa^{o}_{\alpha 3} + gx^{3}\kappa^{1}_{\alpha 3} + f\epsilon^{1}_{\alpha 3} + fg\hat{\epsilon}^{1}_{\alpha 3} \\
\epsilon_{33} &= \epsilon^{o}_{33} + g\tilde{\epsilon}^{o}_{33} + g^{2}\hat{\epsilon}^{o}_{33}
\end{aligned} \tag{9}$$

where $g = df/dx^3$, and

$$\begin{aligned}
\epsilon^{o}_{\alpha\beta} &= \frac{1}{2}(u^{o}_{\alpha|\beta} + u^{o}_{\beta|\alpha}) + \frac{1}{2}(u^{o}_{3|\alpha} + \hat{u}^{o}_{3|\alpha})(u^{o}_{3|\beta} + \hat{u}^{o}_{3|\beta}) \\
\epsilon^{1}_{\alpha\beta} &= -u^{o}_{3|\alpha\beta} \\
\kappa_{\alpha\beta} &= \frac{1}{2}(u^{1}_{\alpha|\beta} + u^{1}_{\beta|\alpha}) \\
\epsilon^{o}_{\alpha 3} &= \frac{1}{2}(\hat{u}^{o}_{3|\alpha} - u^{o}_{\beta|\alpha}u^{o}_{3|\beta}) \\
\hat{\epsilon}^{o}_{\alpha 3} &= \frac{1}{2}(u^{1}_{\alpha} + u^{o}_{\beta|\alpha}u^{1}_{\beta}) \\
\kappa^{o}_{\alpha 3} &= \frac{1}{2}u^{o}_{3|\lambda\alpha}u^{o}_{3|\lambda} \\
\kappa^{1}_{\alpha 3} &= -\frac{1}{2}u^{o}_{3|\lambda\alpha}u^{1}_{\lambda} \\
\epsilon^{1}_{\alpha 3} &= -\frac{1}{2}u^{1}_{\lambda|\alpha}u^{o}_{3|\lambda} \\
\hat{\epsilon}^{1}_{\alpha 3} &= \frac{1}{2}u^{1}_{\lambda|\alpha}u^{1}_{\lambda} \\
\epsilon^{o}_{33} &= \frac{1}{2}u^{o}_{3|\alpha}u^{o}_{3|\alpha} \\
\tilde{\epsilon}^{o}_{33} &= -u^{o}_{3|\alpha}u^{1}_{\alpha} \\
\hat{\epsilon}^{o}_{33} &= \frac{1}{2}u^{1}_{\alpha}u^{1}_{\alpha}
\end{aligned} \tag{10}$$

4 Equations of Motion

The dynamic version of the principle of virtual displacements is used to derive variationally consistent equations of motion associated with the displacement field in equation (8). The principle can be stated (see Reddy, 1984c), in the absence of body forces and prescribed tractions, as

$$0 = \int_{o}^{T}\left[\int_{V}(\sigma^{ij}\delta\epsilon_{ij})dV + \int_{\Omega}q\delta u_{3}dA - \int_{V}\rho(\dot{u}_{i}\delta\dot{u}_{i})dV\right]dt \tag{11}$$

where σ^{ij} denote the contravariant components of the symmetric stress tensor, $q = q(x^{\alpha})$ is the distributed transverse force per unit area, and ρ is the density of the material of the plate. The superposed dot denotes the time derivative, $\dot{u} \equiv \partial u/\partial t$. Using the strains in equation (9), equation (11) can be written as

$$\begin{aligned}
0 = \int_{o}^{T}\Bigg[\int_{\Omega}\Bigg\{\int_{-h/2}^{h/2}[&\sigma^{\alpha\beta}(\delta\epsilon^{o}_{\alpha\beta} + x^{3}\delta\epsilon^{1}_{\alpha\beta} + f\delta\kappa_{\alpha\beta}) + 2\sigma^{\alpha 3}(\delta\epsilon^{o}_{\alpha 3} \\
&+ g\delta\hat{\epsilon}^{o}_{\alpha 3} + x^{3}\delta\kappa^{o}_{\alpha 3} + x^{3}g\delta\kappa^{1}_{\alpha 3} + f\delta\epsilon^{1}_{\alpha 3} + fg\delta\hat{\epsilon}^{1}_{\alpha 3}) + \sigma^{33}(\delta\epsilon^{o}_{33} + g\delta\tilde{\epsilon}^{o}_{33} \\
&+ g^{2}\delta\hat{\epsilon}^{o}_{33}) - \rho\dot{u}^{o}_{\alpha}\delta\dot{u}^{o}_{\alpha} - (x^{3})^{2}\dot{u}^{o}_{3|\alpha}\delta\dot{u}^{o}_{3|\alpha} - \rho f^{2}\dot{u}^{1}_{\alpha}\delta\dot{u}^{1}_{\alpha} + \rho x^{3}(\dot{u}^{o}_{\alpha}\delta\dot{u}^{o}_{3|\alpha} \\
&+ \delta\dot{u}^{o}_{\alpha}\dot{u}^{o}_{3|\alpha}) - f\rho(\dot{u}^{o}_{\alpha}\delta\dot{u}^{1}_{\alpha} + \delta\dot{u}^{o}_{\alpha}\dot{u}^{1}_{\alpha}) + \rho f x^{3}(\dot{u}^{o}_{3|\alpha}\delta\dot{u}^{1}_{\alpha} + \delta\dot{u}^{o}_{3|\alpha}\dot{u}^{1}_{\alpha}) \\
&- \rho(\dot{u}^{o}_{3} + \dot{\hat{u}}^{o}_{3})(\delta\dot{u}^{o}_{3} + \delta\dot{\hat{u}}^{o}_{3})]dx^{3} + q(\delta u^{o}_{3} + \delta\hat{u}^{o}_{3})\Bigg\}dA\Bigg]dt
\end{aligned} \tag{12}$$

Following the standard procedure in the development of plate theories, we introduce stress resultants, couples, and analogous higher-order quantities,

$$(N^{\alpha\beta}, M^{\alpha\beta}, P^{\alpha\beta}) = \int_{-h/2}^{h/2}\sigma^{\alpha\beta}(1, x^{3}, f)dx^{3}$$

$$(Q^\alpha,\hat{Q}^\alpha,R^\alpha,\hat{R}^\alpha,S^\alpha,\hat{S}^\alpha)=\int_{-h/2}^{h/2}\sigma^{\alpha 3}(1,g,x^3,x^3 g,f,fg)dx^3$$

$$(N^3,\tilde{N}^3,\hat{N}^3)=\int_{-h/2}^{h/2}\sigma^{33}(1,g,g^2)dx^3 \tag{13}$$

Note that the introduction of the moments, couples, and higher-order quantities is necessary to reduce the three-dimensional theory to a two-dimensional one. Substituting the relations (13) into equation (12), we obtain

$$\begin{aligned}0=\int_o^T\Big\{\int_\Omega [&N^{\alpha\beta}\delta\epsilon^o_{\alpha\beta}+M^{\alpha\beta}\delta\epsilon^1_{\alpha\beta}+P^{\alpha\beta}\delta\kappa_{\alpha\beta}\\&+2(Q^\alpha\delta\epsilon^o_{\alpha 3}+\hat{Q}^\alpha\delta\hat{\epsilon}^o_{\alpha 3}+R^\alpha\delta\kappa^o_{\alpha 3}+\hat{R}^\alpha\delta\kappa^1_{\alpha 3}+S^\alpha\delta\epsilon^1_{\alpha 3}+\hat{S}^\alpha\delta\hat{\epsilon}^1_{\alpha 3})\\&+N^3\delta\epsilon^o_{33}+\tilde{N}^3\delta\tilde{\epsilon}^o_{33}+\hat{N}^3\delta\hat{\epsilon}^o_{33}+q(\delta u^o_3+\delta\hat{u}^o_3)\\&-I_o[\dot{u}^o_\alpha\delta\dot{u}^o_\alpha+(\dot{u}^o_3+\dot{\hat{u}}^o_3)(\delta\dot{u}^o_3+\delta\dot{\hat{u}}^o_3)]+I_1(\dot{u}^o_\alpha\delta\dot{u}^o_{3\,|\alpha}+\delta\dot{u}^o_\alpha\dot{u}^o_{3\,|\alpha})\\&-I_1^f(\dot{u}^o_\alpha\delta\dot{u}^1_\alpha+\delta\dot{u}^o_\alpha\dot{u}^1_\alpha)+\hat{I}_2(\dot{u}^o_{3\,|\alpha}\delta\dot{u}^1_\alpha+\delta\dot{u}^o_{3\,|\alpha}\dot{u}^1_\alpha)\\&-I_2\dot{u}^o_{3\,|\alpha}\delta\dot{u}^o_{3\,|\alpha}-I_2^f\dot{u}^1_\alpha\delta\dot{u}^1_\alpha]dA\Big\}dt\end{aligned} \tag{14}$$

where I's denote the inertias,

$$I_o=\int_{-h/2}^{h/2}\rho dx^3$$

$$I_1=\int_{-h/2}^{h/2}\rho x^3 dx^3$$

$$I_1^f=\int_{-h/2}^{h/2}\rho f dx^3$$

$$I_2=\int_{-h/2}^{h/2}\rho (x^3)^2 dx^3$$

$$\hat{I}_2=\int_{-h/2}^{h/2}\rho x^3 f dx^3$$

$$I_2^f=\int_{-h/2}^{h/2}\rho f^2 dx^3 \tag{15}$$

The equations of motion of the theory are obtained by substituting equation (10) for the strains in terms of the displacements (u^o_α, u^o_3, $\hat{u}^o_3$, u^1_α), integrating by parts to transfer differentiation from the displacements to the stress resultants and couples, collecting the coefficients of the various virtual displacements, and invoking the fundamental lemma of the calculus of variations. We obtain the following six equations:

$$\begin{aligned}&\delta u^o_\alpha\colon\ N^{\alpha\beta}\,|_\beta-\underline{(Q^\beta u^o_{3\,|\alpha})\,|_\beta+(\hat{Q}^\beta u^1_\alpha)\,|_\beta}=I_o\ddot{u}^o_\alpha-I_1\ddot{u}^o_{3\,|\alpha}+I_1^f\ddot{u}^1_\alpha\\&\delta u^o_3\colon\ M^{\alpha\beta}\,|_{\alpha\beta}+[N^{\alpha\beta}(u^o_{3\,|\beta}+\hat{u}^o_{3\,|\beta})]\,|_\alpha-\underline{(Q_\alpha u^o_{\beta\,|\alpha})\,|_\beta}\\&\underline{-(R^\alpha\,|_\alpha u^o_{3\,|\alpha})\,|_\beta+(\hat{R}^\alpha u^1_\beta)\,|_{\beta\alpha}-(S^\alpha u^1_{\beta\,|\alpha}+(N^3 u^o_{3\,|\alpha})\,|_\alpha}\\&\underline{-(\tilde{N}^\alpha u^1_\alpha)\,|_\alpha}=q+I_o(\ddot{u}^o_3+\ddot{\hat{u}}^o_3)+I_1\ddot{u}^o_{\alpha\,|\alpha}+\hat{I}_2\ddot{u}^1_{\alpha\,|\alpha}-I_2\ddot{u}^o_{3\,|\alpha\alpha}\\&\delta\hat{u}^o_3\colon\ [N^{\alpha\beta}(u^o_{3\,|\beta}+\hat{u}^o_{3\,|\beta})]\,|_\alpha+Q^\alpha\,|_\alpha=q+I_o(\ddot{u}^o_3+\ddot{\hat{u}}^o_3)\\&\delta u^1_\alpha\colon\ P^{\alpha\beta}\,|_\beta-\hat{Q}^\beta(\delta_{\alpha\beta}+\underline{u^o_{\alpha\,|\beta}})+\underline{\hat{R}^\beta u^1_{3\,|\alpha\beta}+\hat{S}^\beta\,|_\beta u^1_\alpha}\\&\underline{-(S^\beta u^o_{3\,|\alpha})\,|_\beta-\hat{N}^3 u^1_\alpha+\tilde{N}^3 u^o_{3\,|\alpha}}=I_1^f\ddot{u}^o_\alpha-\hat{I}_2\ddot{u}^o_{3\,|\alpha}+I_2^f\ddot{u}^1_\alpha\end{aligned} \tag{16}$$

where the underlined terms are due entirely to the inclusion of moderate rotations (i.e., over and above the von Karman nonlinear terms).

Equations (16) can be specialized to the three different theories discussed earlier. The equations are summarized below:

(*i*) *Classical Plate Theory* ($\hat{u}^o_3=0$, $u^1_\alpha=0$)

$$N^{\alpha\beta}\,|_\beta-(Q^\beta u^o_{3\,|\alpha})\,|_\beta=I_o\ddot{u}^o_\alpha$$

$$M^{\alpha\beta}\,|_{\alpha\beta}+(N^{\alpha\beta}u^o_{3\,|\beta})\,|_\alpha-(Q^\alpha u^o_{\beta\,|\alpha})\,|_\beta-(R^\alpha\,|_\alpha u^o_{3\,|\beta})\,|_\beta+(N^3 u^o_{3\,|\alpha})\,|_\alpha=q+I_o\ddot{u}^o_3+I_1\ddot{u}^o_{\alpha\,|\alpha}-I_2\ddot{u}^o_{3\,|\alpha\alpha} \tag{17}$$

(*ii*) *First-Order Shear Deformation Plate Theory* ($u^o_3=0$, $f=x^3$)

$$N^{\alpha\beta}\,|_\beta+(Q^\beta u^1_\alpha)\,|_\beta=I_o\ddot{u}^o_\alpha+I_1\ddot{u}^1_\alpha$$

$$Q^\alpha\,|_\alpha+N^{\alpha\beta}\hat{u}^o_{3\,|\beta})\,|_\alpha=q+I_o\ddot{\hat{u}}^o_3$$

$$M^{\alpha\beta}\,|_\beta-Q^\beta(\delta_{\alpha\beta}+u^o_{\alpha\,|\beta})-N^3 u^1_\alpha+R^\beta\,|_\beta u^1_\alpha=I_1\ddot{u}^o_\alpha+I_2\ddot{u}^1_\alpha \tag{18}$$

(*iii*) *Third-Order Shear Deformation Plate Theory*

$$\left(u^o_3=0,\ f=x^3\left[1-\frac{4}{3}(x^3\,|h)^2\right]\right)$$

$$N^{\alpha\beta}\,|_\beta+(\hat{Q}^\beta u^1_\alpha)\,|_\beta=I_o\ddot{u}^o_\alpha+I_1^f\ddot{u}^1_\alpha$$

$$Q^\alpha\,|_\alpha+(N^{\alpha\beta}\hat{u}^o_{3\,|\beta})\,|_\alpha=q+I_o\ddot{\hat{u}}^o_3$$

$$P^{\alpha\beta}\,|_\beta-\hat{Q}^\beta(\delta_{\alpha\beta}+u^o_{\alpha\,|\beta})+\hat{S}^\beta\,|_\beta u^1_\alpha-\hat{N}^3 u^1_\alpha=I_1^f\ddot{u}^o_\alpha+I_2^f\ddot{u}^1_\alpha \tag{19}$$

Note that several other theories can be obtained from equation (16) as special cases. For example, the refined theory of Kromm (1953, 1955) can be obtained by setting $\hat{u}^o_3=0$.

5 The First-Order Theory With Moderate Rotations

In view of the extensive use of the first-order shear deformation theory (most often referred to as the Reissner-Mindlin plate theory or simply the Mindlin plate theory) in practice, the detailed equations of the theory with moderate rotation effects are presented here. Librescu and Schmidt (1986) presented such a theory, but the governing equations given there were in terms of the resultants instead of displacements. It is informative to note that the substitution of the displacements and their gradients for the resultants into the equations given by Librescu and Schmidt (1986) will give rise to many additional terms which, by the assumption in equation (16), should be zero. To make use of the relations (16) these additional terms must be expressed in terms of the strains ϵ_{ij} and rotations ω_{ij}.

Displacement Field

$$\begin{aligned}u_\alpha(x^\beta,x^3,t)&=u^o_\alpha(x^\beta,t)+x^3u^1_\alpha(x^\beta,t)\\u_3(x^\beta,x^3,t)&=u^o_3(x^\beta,t)\end{aligned} \tag{20}$$

Strain-Displacement Equations

$$\begin{aligned}\epsilon_{\alpha\beta}&=\epsilon^o_{\alpha\beta}+x^3\epsilon^1_{\alpha\beta}\\\epsilon_{3\alpha}&=\epsilon^o_{3\alpha}+x^3\epsilon^1_{3\alpha}\\\epsilon_{33}&=\epsilon^o_{33}\end{aligned} \tag{21}$$

with

$$\epsilon^o_{\alpha\beta}=\frac{1}{2}(u^o_{\alpha\,|\beta}+u^o_{\beta\,|\alpha})+\frac{1}{2}u^o_{3\,|\alpha}u^o_{3\,|\beta}$$

$$\epsilon^1_{\alpha\beta}=\frac{1}{2}(u^1_{\alpha\,|\beta}+u^1_{\beta\,|\alpha})$$

$$\epsilon^o_{3\alpha}=\frac{1}{2}(u^o_{3\,|\alpha}+u^1_\alpha)+\frac{1}{2}u^1_\beta u^o_{\beta\,|\alpha}$$

$$\epsilon^1_{3\alpha}=\frac{1}{2}u^1_\beta u^1_{\beta\,|\alpha}$$

$$\epsilon^o_{33}=\frac{1}{2}u^1_\alpha u^1_\alpha \tag{22}$$

Equations of Motion

$$N^{\alpha\beta}\,|_\beta + (Q^\beta u^1_{\alpha|\beta})\,|_\beta = I_o \ddot{u}^o_\alpha + I_1 \ddot{u}^1_\alpha$$

$$M^{\alpha\beta}\,|_\beta - Q^\beta\,(\delta_{\alpha\beta} + u^o_{\alpha|\beta}) - N^3 u^1_\alpha + R^\beta\,|_\beta u^1_\alpha = I_1 \ddot{u}^o_\alpha + I_2 \ddot{u}^1_\alpha$$

$$(N^{\alpha\beta} u^o_{3|\beta})\,|_\beta + Q^\alpha\,|_\alpha = q + I_o \ddot{u}^o_3 \tag{23}$$

where

$$(N^{\alpha\beta}, M^{\alpha\beta}) = \int_{-h/2}^{h/2} \sigma^{\alpha\beta}(1, x^3)\,dx^3$$

$$(Q^\alpha, R^\alpha) = \int_{-h/2}^{h/2} \sigma^{\alpha 3}(1, x^3)\,dx^3$$

$$N^3 = \int_{-h/2}^{h/2} \sigma^{33}\,dx^3 \tag{24}$$

Constitutive Equations of a Laminate

$$\begin{Bmatrix} N^1 \\ N^2 \\ N^6 \\ N^3 \\ M^1 \\ M^2 \\ M^3 \end{Bmatrix} = \begin{bmatrix} A^{11} & A^{12} & A^{16} & A^{13} & B^{11} & B^{12} & B^{16} \\ A^{12} & A^{22} & A^{26} & A^{23} & B^{12} & B^{22} & B^{26} \\ A^{16} & A^{26} & A^{66} & A^{36} & B^{16} & B^{26} & B^{66} \\ A^{13} & A^{23} & A^{36} & A^{33} & B^{13} & B^{23} & B^{36} \\ B^{11} & B^{12} & B^{16} & B^{13} & D^{11} & D^{12} & D^{16} \\ B^{12} & B^{22} & B^{26} & B^{23} & D^{12} & D^{22} & D^{26} \\ B^{16} & B^{26} & B^{66} & B^{36} & D^{16} & D^{26} & D^{66} \end{bmatrix} \begin{Bmatrix} \epsilon^o_{11} \\ \epsilon^o_{22} \\ 2\epsilon^o_{12} \\ \epsilon^o_{33} \\ \epsilon^1_{11} \\ \epsilon^1_{22} \\ 2\epsilon^1_{12} \end{Bmatrix}$$

$$\begin{Bmatrix} Q^2 \\ Q^1 \\ R^2 \\ R^1 \end{Bmatrix} = \begin{bmatrix} A^{44} & A^{45} & B^{44} & B^{45} \\ A^{45} & A^{55} & B^{45} & B^{55} \\ B^{44} & B^{45} & D^{44} & D^{45} \\ B^{45} & B^{55} & D^{45} & D^{55} \end{bmatrix} \begin{Bmatrix} 2\epsilon^o_{23} \\ 2\epsilon^o_{13} \\ 2\epsilon^1_{23} \\ 2\epsilon^1_{13} \end{Bmatrix} \tag{25}$$

where $N^{11} = N^1$, $N^{22} = N^2$, $N^{12} = N^6$, etc., and

$$(A^{ij}, B^{ij}, D^{ij}) = \sum_{k=1}^{N} \int_{x^3_{k-1}}^{x^3_k} Q^{ij}_{(k)}(1, x^3, (x^3)^2)\,dx^3$$

$$(i, j = 1,2,3,6) \tag{26}$$

N being the number of layers in the laminate, $Q^{ij}_{(k)}$ are the elastic coefficients of kth layer in the laminate coordinate system, and (x^3_{k-1}, x^3_k) are the x^3 coordinates of the bottom and top of the kth layer. The coefficients A^{ij}, B^{ij}, and D^{ij} for $i, j = 4, 5$ are defined in equation (26) except that $Q^{ij}_{(k)}$ are multiplied by appropriate shear correction factors. Inspection of the equations of motion (23) reveal that the moderate rotation terms couple all three equations of motion. Recall that the von Karman nonlinearity is present only in the third equation.

6 Closure

A higher-order shear deformation theory for the dynamics of general anisotropic plate that accounts for moderate rotations is developed. The theory contains, as special cases, the von Kármán analogs of the classical plate theory, the Reissner-Mindlin plate theory and a third-order shear deformation theory. All equations of motion of the theory are strongly coupled. The detailed equations of the first-order, moderate rotation, shear deformation theory of laminated composite plates are outlined. A close examination of the virtual work statement shows that the displacement finite element model of the first-order theory still requires C^o elements for the approximation of the five displacements $(u^o_\alpha, u^1_\alpha, u^o_3)$. The development of finite element models of the moderate rotation theories presented here awaits attention.

Acknowledgments

The support of this research by the Mathematical Sciences Division of the Army Research Office through Grant DAAG-29-85-K-0007 is gratefully acknowledged. It is also a pleasure to acknowledge the support of the German Academic Exchange Service and the Alexander von Humboldt Foundation Research Grants.

References

Chia, C. Y., 1980, *Nonlinear Analysis of Plates*, McGraw-Hill, New York.

Ebcioglu, I. K., 1964, "A Large Deflection Theory of Anisotropic Plates," *Ingenieur-Archiv*, Vol. 33, pp. 396-403.

Habip, L. M., 1966, "Theory of Elastic Plates in the Reference State," *Int. J. Solids & Structures*, Vol. 2, pp. 157-166.

Huffington, Jr., N. J., 1963, "Response of Elastic Columns to Axial Pulse Loading," *AIAA Journal*, Vol. 1, No. 9, pp. 2099-2104.

Iura, M., 1986, "A Generalized Variational Principle for Thin Elastic Shells with Finite Rotation," *International Journal of Solids and Structures*, Vol. 22, No. 2, pp. 141-154.

Kayuk, Ya. F., and Sakhatskii, V. G., 1985, "Nonlinear Shell Theory on the Basis of the Concept of Finite Rotation," *Soviet Applied Mechanics*, Vol. 21, No. 4, pp. 366-373.

Koiter, W. T., 1966, "On the Nonlinear Theory of Thin Elastic Shells," *Proc. Kon. Ned. Ak. Wet.*, Series B., Vol. 69, pp. 1-54.

Krishna Murty, A. V., 1986, "Toward a Consistent Plate Theory," *AIAA Journal*, Vol. 24, pp. 1047-1048.

Kromm, A., 1953, "Verallgemeinerte Theorie der Plattenstatik," *Ing.-Arch.*, Vol. 21, pp. 266-286.

Kromm, A., 1955, "Über die Randquerkräfte bei Gestützten Platten," *Z. angew. Math. Mech.*, Vol. 35, pp. 231-242.

Librescu, L., 1975, *Elastostatics and Kinematics of Anisotropic and Heterogeneous Shell-Type Structures*, Noordhoff, Leyden, The Netherlands.

Librescu, L., and Schmidt, R., 1986, "Higher-Order Moderate Rotation Theories for Elastic Anisotropic Plates," *Finite Rotations in Structural Mechanics, Proc. of Euromech Symposium, No. 197, Jablonna, Poland, 1985*, Springer-Verlag, Berlin, pp. 158-174.

Medwadowski, S. J., 1958, "A Refined Theory of Elastic, Orthotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 25, pp. 437-443.

Naghdi, P. M., 1972, "The Theory of Shells and Plates," in *Handbuch der Physik*, Vol. VIa/2, Flugge, S., ed., Springer-Verlag, New York, pp. 425-640.

Naghdi, P. M., and Vongsarnpigoon, L., 1983, "A Theory of Shells with Small Strains Accompanied by Moderate Rotations," *Arch. for Rational Mechanics and Analysis*, Vol. 83, pp. 245-283.

Nolte, L. P., Makowski, J., and Stumpff, H., 1986, "On the Derivation and Comparative Analysis of Large Rotation Shell Theories," *Ingenieur-Archiv*, Vol. 56, No. 2, pp. 145-160.

Pietraszkiewicz, W., 1979, *Finite Rotations and Lagrangian Description in the Non-Linear Theory of Shells*, Polish Academy of Sciences, Institute of Fluid-Flow Machinery, Polish Scientific Publishers, Warszawa, Poznan.

Pietraszkiewicz, W., 1980, "Finite Rotations in the Non-Linear Theory of Thin Shells," in *Thin Shell Theory, New Trends and Applications*, Olszak, W., ed., Springer-Verlag, New York, pp. 153-208.

Pietraszkiewicz, W., 1984, "Lagrangian Description and Incremental Formulation in the Nonlinear Theory of Thin Shells," *Int. J. Non-Linear Mechanics*, Vol. 19, pp. 115-140.

Ranjan, G. V., and Steele, C. R., 1980, "Nonlinear Corrections for Edge Bending of Shells," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, No. 4, pp. 861-865.

Reddy, J. N., and Rasmussen, M. L., 1982, *Advanced Engineering Analysis*, Wiley, New York.

Reddy, J. N., 1983, "Geometrically Nonlinear Transient Analysis of Laminated Composite Plates," *AIAA Journal*, Vol. 21, No. 4, pp. 621-629.

Reddy, J. N., 1984a, "Refined Nonlinear Theory of Plates with Transverse Shear Deformation," *Int. J. Solids and Structures*, Vol. 20, No. 9/10, pp. 881-896.

Reddy, J. N., 1984b, "A Simple Higher-Order Theory for Laminated Composite Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 745-752.

Reddy, J. N., 1984c, *Energy and Variational Methods in Applied Mechanics*, Wiley, New York.

Reissner, E., 1958, "Rotationally Symmetric Problems in the Theory of Thin Elastic Shells," *Proc. 3rd U.S. Nat. Congr. of Appl. Mech.*, pp. 51-69.

Sanders, J. L., 1963, "Nonlinear Theories for Thin Shells," *Quart. Appl. Math.*, Vol. 21, pp. 21-36.

Schmidt, R., 1984, "On Geometrically Nonlinear Theories for Thin Elastic Shells," in *Flexible Shells, Theory and Applications, Proc. Euromech Colloqium, No. 165, Munich, 1983*, Axelrad, E. L., and Emmerling, F. A., eds., Springer-Verlag, Heidelberg, pp. 76-90.

Wempner, G. A., 1973, *Mechanics of Solids with Applications to Thin Bodies*, McGraw-Hill, New York.

Yokoo, Y., and Matsunaga, H., 1974, "A General Theory of Elastic Shells," *Int. J. Solids and Structures*, Vol. 10, pp. 261-274.

Nearly Circular Connections of Elastic Half Spaces

Huajian Gao

James R. Rice

Division of Applied Sciences,
Harvard University,
Cambridge, MA 02138

In this paper we solve the elasticity problem of two elastic half spaces that are joined together over a region that does not differ much from a circle, i.e., the problem of an external planar crack leaving a nearly circular uncracked connection. The method we use is based on the perturbation technique developed by Rice (1985) for solving the elastic field of a crack whose front deviates slightly from some reference geometry. Quantities such as crack opening displacement and stress intensity factor are derived in detail to the first order of accuracy in the deviation of the shape of the connection from a circle. In addition, some results such as the crack face weight functions and Green's functions for a perfectly circular connection are also discussed under various boundary conditions at infinity. The formulae derived are used to study the configurational stability problem for quasistatic growth of an external circular crack. The results, derived when the crack front is perturbed from circular in a harmonic wave form and is subjected to axisymmetric loading, suggest that a perturbation of wavenumber higher than one is configurationally stable under all boundary conditions at infinity. The perturbation with wavenumber equal to one, which corresponds to a translational shift of the geometric center of the circular connection, turns out to be configurationally stable if any rotation in the remote field is suppressed and configurationally unstable if there is no such restraint.

Introduction

Rice (1985) developed a method of solving the elasticity problem of a planar crack whose front differs slightly in location from that of some reference geometry. It has been applied to cases such as semi-infinite planar cracks with slightly nonstraight fronts (Rice, 1985; Gao and Rice, 1986) and internal somewhat circular cracks (Gao and Rice, 1987). The latter work (Gao and Rice, 1987) has shown that the perturbation method is not only convenient but also remarkably accurate in determining crack opening displacement and stress intensity factors for crack configurations that differ moderately from a circular reference geometry. The internal circular crack problem was addressed much earlier in a perturbation sense by Panasyuk (1962), and Gao and Rice (1987) compare their approach to his. Rice's perturbation method can be carried out immediately for a tensile crack if the solution for the stress intensity factor distribution is known along the reference crack front due to a pair of concentrated wedging forces acting to open the crack at an arbitrary location on its surfaces. Such a point force solution, sometimes called the crack face weight function after Bueckner (1970, 1973) and Rice (1972), was derived by Stallybrass (1981) for an external circular crack, i.e., a circular connection between elastic half-spaces under a traction free boundary condition at infinity. Following Stallybrass's work we are also able to clarify ambiguities in some previously proposed solutions in the literature (e.g., Kassir and Sih, 1975; Tada et al., 1973).

In this paper we therefore solve for the crack opening displacement and tensile mode stress intensity factor for a slightly noncircular connection. The notation δ (F) is used in what follows to denote the variation in some field variable F from its form for the reference circular crack to that for the perturbed crack shape.

Consider two isotropic, homogeneous three-dimensional elastic semi-infinite solids joined over some slightly noncircular connection of bounding contour c. A Cartesian coordinate system x, y, z is attached so that the joining planes lie on $y = 0$ and the origin of the coordinate system is assumed to coincide with the center of some convenient reference circle. This configuration forms an external crack with its front c described by some function $a(s)$ where $a(s)$ is the distance from the origin of the coordinate system to the position s along the crack front; $a(s)$ is nearly constant, and is constant on the reference circle. The crack system is subjected to some distribution of fixed forces that induce "*M*ode 1" tension along the crack front. We may note that in this case when the crack grows into the connecting ligament, $a(s)$ decreases. Therefore, we represent the crack growth from the reference circular shape to the actual shape by $-\delta a(s)$. In this circumstance it can be shown, following Rice (1985), that the variation in opening displacement $\Delta u(x,z)$ between upper and lower crack surfaces at location x, z, when the crack front is

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13-18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March 6, 1987.

Paper No. 87-WA/APM-14.

altered from the reference circular front by $-\delta a(s)$ in presence of the fixed load system, is

$$\delta[\Delta u(x,z)] = -\frac{2(1-\nu^2)}{E} \oint_c K^0(s) k(s;x,z) \delta a(s) ds \tag{1}$$

to first order in $\delta a(s)$. Here $K^0(s)$ is the Mode 1 intensity factor induced along the reference crack front by the fixed load system and $k(s;x,z)$ is the intensity factor that would be induced at arc length position s along the reference front by a pair of unit wedging forces opening the crack at location at x, z; $k(s; x, z)$ can be called the Mode 1 crack face weight function. This weight function is discussed in detail in Appendix A. Here the intensity factor K is defined so that $K/\sqrt{2\pi\epsilon}$ is the asymptotic form of the tensile stress at small perpendicular distance ϵ from the crack front on the prolongation of the crack plane within the connection.

Crack Opening Displacement

We choose the reference crack as a perfectly circular connection of radius a and adopt polar coordinates for convenience so that $s = a\theta'$ in equation (1). Here the polar coordinate angle θ' is measured from the positive x axis, increasing towards the positive z axis. To emphasize dependence on the reference circular radius a, we introduce the notations $K^0(s) = K^0[\theta;a]$ and $k(s; x, z) = k(\theta'; r, \theta; a)$ for the intensity factors induced at θ' along the reference crack front, respectively, by the given load system and by a pair of unit wedging point forces at polar position r, θ. Then equation (1) becomes

$$\delta[\Delta u(r,\theta)] = -\frac{2(1-\nu^2)}{E} \int_0^{2\pi} \times K^0[\theta'; a] k(\theta'; r, \theta; a)\, a\, \delta a(\theta')\, d\theta' \tag{2}$$

where $\delta a(\theta') = a(\theta') - a$. Also, we introduce the notation $\Delta u\ (r, \theta) = \Delta u^0\ [r, \theta; a]$ to describe the opening of a perfectly circular connection of radius a under the given loadings.

We can also derive $K^0\ [\theta'; a]$ by the law of superposition when some distributed load $p(r; \theta)$ is acting on the external crack faces

$$K^0[\theta'; a] = \int_0^{2\pi} \int_a^{\infty} p(\rho, \phi)\, k(\theta'; \rho, \phi; a)\, \rho d\rho d\phi \tag{3}$$

The problem of general tensile loading can also be described in this way when $p(r, \theta)$ is equated to the tensile stress which the general loading would induce at r, θ in the absence of the external crack.

To find the opening displacement field for a perfectly circular connection, we impose a uniform crack growth, i.e., $\delta a(\theta') = \delta a$ in equation (2). Then dividing both sides of equation (2) by δa and lettering $\delta a \to 0$, we get

$$\frac{\partial \Delta u^0[r, \theta; a]}{\partial a} = -2\frac{(1-\nu^2)}{E} \int_0^{2\pi} \times K^0[\theta'; a] k(\theta'; r, \theta; a) a\, d\theta' \tag{4}$$

Noting that $\Delta u^0\ [r, \theta; a] = 0$ when $a \geq r$ (only crack faces open), we integrate over the crack size variable a' and get

$$\Delta u^0[r, \theta; a] = 2\frac{(1-\nu^2)}{E} \int_0^{2\pi} \int_a^r \times K^0[\theta'; a']\, k(\theta'; r, \theta; a')\, a'\, da'\, d\theta' \tag{5}$$

Substituting equation (3) into (5), we get the following general crack opening displacement for external circular cracks,

$$\Delta u^0[r, \theta; a] = 2\frac{(1-\nu^2)}{E} \int_0^{2\pi} \int_a^r \int_0^{2\pi} \int_{a'}^{\infty} k(\theta'; \rho, \phi; a') \times k(\theta'; r, \theta; a')\, a' p(\rho, \phi)\, \rho d\rho d\phi\, da' d\theta' \tag{6}$$

If we switch the order of integration with respect to a', θ' and ρ, ϕ, we therefore could rewrite equation (6) as

$$\Delta u^0\ [r, \theta; a] = \int_0^{2\pi} \int_a^{\infty} D(r, \theta; \rho, \phi)\, p(\rho, \phi)\, \rho d\rho d\phi \tag{7}$$

where

$$D(r, \theta; \rho, \phi) = 2\frac{(1-\nu^2)}{E} \int_a^{\min(r,\rho)} \int_0^{2\pi} \times k(\theta'; \rho, \phi; a')\, k(\theta'; r, \theta; a')\, a'\, d\theta'\, da' \tag{8}$$

is clearly identified as the crack face Green's function for an external circular crack, and it is further discussed in detail in Appendix B and also in Appendix D.

Equation (6), or equation (7) combined with equation (8), gives us the formula to determine the crack opening displacement for a perfectly circular connection. The integrals in those equations can be carried out once the loading system $p(r, \theta)$ and the crack face weight function $k(\theta'; r, \theta; a)$ is known. The function $k(\theta'; r, \theta; a)$ is discussed in Appendix A and presented under various boundary conditions at infinity. The most general form of $k(\theta'; r, \theta; a)$ is given by equation (A-9) of Appendix A under traction free, completely unrestrained displacement conditions at infinity. For convenience we present it here too:

$$k(\theta'; r, \theta; a) = \frac{1}{(\pi a)^{3/2}} \left\{ \left[\cos^{-1}\left(\frac{a}{r}\right) + \frac{a\sqrt{r^2-a^2}}{a^2+r^2-2ar\cos(\theta'-\theta)} \right] + 3\left[\frac{r}{a}\cos^{-1}\left(\frac{a}{r}\right) + \left(1-\frac{a^2}{r^2}\right)^{1/2} \right] \cos(\theta'-\theta) \right\} \tag{9}$$

When the shape of the connection is slightly noncircular, it is convenient for purposes of calculating the opening $\Delta u\ (r, \theta)$ along the ray at any particular angle θ to take the radius of the reference circular crack front to be a circle of radius equal to $a(\theta)$. We then are able to let r approach simultaneously both the reference front and the actual perturbed front. This procedure, as described in earlier papers (Rice, 1985, Gao and Rice, 1986, 1987), is necessary to retain the correct asymptotic behavior near the crack front as is crucial for the calculation of the stress intensity factor along the perturbed crack front. Then equation (2) becomes,

$$\delta\ [\Delta u(r, \theta)] = 2\frac{(1-\nu^2)}{E} \int_0^{2\pi} K^0[\theta'; a(\theta)]\ k(\theta'; r, \theta; a(\theta)) \times [\, a(\theta) - a(\theta')]\, a(\theta)\ \ d\theta' \tag{10}$$

Equation (10) plus equation (6) then gives the total opening displacement as

$$\begin{aligned} \Delta u(r, \theta) &= \Delta u^0[r, \theta; a(\theta)] + \delta\ [\Delta u(r, \theta)] \\ &= 2\frac{(1-\nu^2)}{E} \int_0^{2\pi} \left\{ \int_{a(\theta)}^r K^0[\theta'; a']\ k(\theta'; r, \theta; a')\, a'\, da' \right. \\ &\quad \left. + K^0[\theta'; a(\theta)] k(\theta'; r, \theta; a(\theta))\ [a(\theta) - a(\theta')] a(\theta) \right\} d\theta' \\ &\simeq 2\frac{(1-\nu^2)}{E} \int_0^{2\pi} \int_{a(\theta')}^r \times K^0[\theta'; a']\ k(\theta'; r, \theta; a')\, a'\, da'\, d\theta' \end{aligned} \tag{11}$$

where the last $\simeq$ means equal to first order of accuracy in $a(\theta') \doteq a(\theta)$. Equation (11) can be used to evaluate the opening displacement for a slightly noncircular connection if one is given the shape of that connection (i.e., the function $a(\theta')$).

Stress Intensity Factors

Stress intensity factors can be extracted from the near tip behavior of the crack opening displacement, as indicated by equation (*B*-13) of Appendix B. The same relation holds between $\delta\,[\Delta u(r,\theta)]$ and $\delta K(\theta)$, the first order variation of intensity factor as

$$\delta[\Delta u(r,\theta)] = \frac{8(1-\nu^2)}{E} \times \left\{ \delta K(\theta) \sqrt{\frac{a(\theta)-r}{2\pi}} + \mathrm{O}\,[(a(\theta)-r)^{3/2}] \right\} \tag{12}$$

where again the variations are from the reference circular front, of radius equal to $a(\theta)$, to the perturbed front. Substituting crack face weight function (9) into equation (10) and letting $r \to a(\theta)$, we get the following asymptotic formula

$$\delta[\Delta u(r,\theta)] = \frac{2(1-\nu^2)}{\pi E} \sqrt{\frac{r-a(\theta)}{\pi a(\theta)}} \int_0^{2\pi} \times K^0[\theta';\, a(\theta)] \left\{ \sqrt{\frac{2}{a(\theta)}} + \frac{a(\theta)\sqrt{2a(\theta)}}{a(\theta)^2 4\,\sin^2\,[(\theta-\theta')/2]} + \frac{6\sqrt{2}\,\cos\,(\theta'-\theta)}{\sqrt{a(\theta)}} \right\} [a(\theta)-a(\theta')]\,d\theta' \tag{13}$$

Comparing equation (13) with (12), we see immediately that the variation in stress intensity factor is

$$\delta K(\theta) = K(\theta) - K^0[\theta;\, a(\theta)] = \frac{1}{2\pi}\, PV \int_0^{2\pi} K^0[\theta';\, a(\theta)]\,[1-a(\theta')/a(\theta)] \times \left\{ 1 + \frac{1}{4\,\sin^2\,[(\theta'-\theta)/2]} + 6\,\cos\,(\theta'-\theta) \right\} d\theta' \tag{14}$$

Here PV denotes principal value. Equation (14) gives the formula to evaluate the stress intensity factor when the shape of the connection, i.e., $a(\theta')$, and the loading configuration, i.e., $K^0\,[\theta';\, a(\theta)]$, are known.

In fact, equation (14) is correct only when we do not have a displacement-restraint type of boundary condition at infinity, i.e., when the crack system is subjected only to fixed forces. Similar to the discussion in Appendix A, we treat some typical displacement boundary conditions at infinity in the following.

***(i)* "Clamped" at Infinity, i.e., Fixed Against Any Displacement.** In this case, the crack face weight function should be $k_d(\theta';\, r,\, \theta;\, a)$ of equation (*A*-6) of Appendix A. Following the similar steps leading to equation (14), we have

$$\delta K(\theta) = \frac{1}{8\pi}\, PV \int_0^{2\pi} \frac{K^0[\theta';\, a(\theta)]1-a(\theta')/a(\theta)]}{\sin^2[(\theta'-\theta)/2]}\, d\theta' \tag{15}$$

***(ii)* Free Vertical Motion But Fixed Against Rotation.** In this case, the crack face weight function should be $k_v(\theta';\, r,\theta;a)$ of equation (*A*-7) of Appendix A. Similarly we have

$$\delta K(\theta) = \frac{1}{2\pi}\, PV \int_0^{2\pi} K^0\,[\theta';a(\theta)]\,[1-a(\theta')/a(\theta)] \times \left\{ 1 + \frac{1}{4\,\sin^2[(\theta'-\theta)/2]} \right\} d\theta' \tag{16}$$

***(iii)* Free Rotation But Fixed Against Vertical Displacement Along *y* Axis.** In this case, the crack face weight function should be $k_r(\theta';\, r,\, \theta;\, a)$ of equation (*A*-8). Therefore,

$$\delta K(\theta) = \frac{1}{2\pi}\, PV \int_0^{2\pi} K^0\,[\theta';a(\theta)]\,[1-a(\theta')/a(\theta)] \times \left\{ \frac{1}{4\,\sin^2[(\theta'-\theta)/2]} + 6\,\cos\,(\theta'-\theta) \right\} d\theta' \tag{17}$$

From now on, for conciseness we will refer to the above different cases of boundary conditions at infinity by their case number, e.g., case (*i*) represents fixed displacement at infinity, and the case of equations (9), (13) and (14), for which there is no restraint against displacement at infinity, will be called case (*iv*).

Growth Mode of an External Circular Crack

The previous elastic analysis of somewhat circular connections may be used to study the configurational stability of the fracturing process of a bonded circular area between two large elastic solids, at least when this occurs quasistatically (e.g., by fatigue load cycling or sustained load corrosion) under elastic fracture mechanics conditions. We study the configurational stability of the mode of growth as a concentric circle of diminishing radius for an external, initially circular, crack under some spatially fixed axisymmetric loading system. Since any somewhat noncircular crack growth profile could be represented in terms of a Fourier series, it will be sufficient to consider the following perturbation of the front in a harmonic wave form:

$$a(\theta) = a_0 - Re\,[Ae^{in\theta}] \tag{18}$$

where a_0 is a real constant, n is an integer, A is a constant (possibly complex) and $|A|/a_0 << 1$. We assume that the quasistatic growth rate of the crack increases with the intensity factor at the same location along the front. Then a small harmonic perturbation of wave number n can be said to be configurationally unstable (increase in amplitude $|A|$) during subcritical crack growth if the intensity factor $K(\theta)$ is decreased from $K^0[\theta;\, a_0]$ when $a(\theta)$ exceeds a_0 and increased when $a(\theta)$ is less than a_0, and configurationally stable if the opposite is true. That is, crack growth is likely to amplify the forms of those unstable wave configurations, if any exist. Of course, the growth or decay of the harmonic perturbations is understood to be superposed on the uniform axially symmetric diminuation of a_0 in describing the total crack growth.

Since the applied loading is now considered axially symmetric relative to the reference crack center, $K^0[\theta';\, a] = K^0[a]$, i.e., it is independent of angle. Substituting equation (18) into equations (14), (15), (16), and (17), carrying out the integrations, and expanding $K^0[a]$ to the linear term in a Taylor series about a_0, we have to the first order in $|A|$,

$$K(\theta) = K^0[a_0] - \left\{ \frac{dK^0[a_0]}{da_0} + \frac{n_1}{2a_0}\, K^0[a_0] \right\} Re\,[Ae^{in\theta}] \tag{19}$$

where for case (*i*), $n_1 = n$; for case (*ii*), $n_1 = n + 2$; for case (*iii*),

$$n_1 = \begin{cases} -5 & n=1 \\ n & \text{otherwise} \end{cases}$$

and for case (*iv*),

$$n_1 = \begin{cases} -3 & n=1 \\ n+2 & \text{otherwise} \end{cases} \tag{20}$$

Clearly if the sum within the curly brackets in equation (19) is positive, any perturbation from circular of the corresponding wavenumber would be diminishing, i.e., configurationally stable since K attains the smallest value at the places where the

crack has grown most, i.e., where $Ae^{in\theta} = |A|$. For convenience we name this sum by $H(n_1)$ to emphasize its dependence on the number n_1 (which further relates to wave number n) so that the critical, neutrally stable situation can be said to be reached at a number n_1^c (not necessarily integer) satisfying $H(n_1^c) = 0$. It is easy to see that when $n_1 > n_1^c$ the quantity within the curly bracket in equation (19) becomes positive and it becomes negative if the opposite is true. Therefore, $n_1 < n_1^c$ must be satisfied for a configurationally unstable wavy mode perturbation. It can also be noticed that the translational shift mode, i.e., $n = 1$ is most likely to be unstable for cases (*iii*) and (*iv*) since $n_1 < 0$ in those cases, and higher modes ($n > 1$) are more likely to be unstable for cases (*i*) and (*iii*) since $n_1 = n$ in those cases. Hence it might be suitable to conclude here that case (*ii*) when points at infinity can only move freely in the vertical direction and are fixed against rotation is the most stable crack system while case (*iii*) when points at infinity can only rotate freely about a fixed point on the central axis y is the most unstable system, especially for the translational mode $n = 1$.

Remotely Applied Centered Force; Imposed Remote Displacement

Consider, for example, that a remotely applied tensile force F is transmitted across a circular connection with no net moment about the center of the connection. The case (*iv*) formulae of the last section apply here and

$$H(n_1) = \frac{dK^0}{da_0} + \frac{n_1}{2a_0} K^0\ [a_0] = \frac{n_1 - 3}{2a_0} K^0\ [a_0] \tag{21}$$

By equations (20), we know that $n_1 = n + 2$ for $n > 1$. Therefore, $H > 0$ for $n > 1$ so that all perturbations of wavenumber greater than one are configurationally stable. For the translational mode, i.e., when $n = 1$ and $n_1 = -3$, it is obvious that $H < 0$ so that this mode is configurationally unstable. In fact, equation (19) becomes when $n = 1$,

$$K = K^0\ [1 + 3\ \mathrm{Re}(Ae^{i\theta})/a_0] \tag{22}$$

We get the same relation by applying equation (B-8) of Appendix B, for a connection under remotely applied force and moment, as in this case the center of the connection has simply been shifted by an amount $|A|$ so as to generate a net moment equal to $F|A|$ about the $\theta = 90° - \arg(A)$ axis (here $\arg(A)$ is the phase angle of A). Therefore, equation (22) is valid even for a shift of any finite amount. This suggests that translational shift is very likely to occur when the crack system is subjected only to a centered force. It should be noted that the shape the crack will take after finite amount of growth is hard to predict because once the translational shift occurs the net moment thus generated has to be considered. The stress intensity factor will become nonuniform along the shifted circle, and thus it will not remain circular.

A case is studied in Appendix C for which the crack system is subjected to a fixed vertical displacement of amount equal to c at infinity and the stress intensity factor and crack opening displacement thus induced are also derived there. Under this displacement boundary condition, the crack face weight function should be $k_d\ (\theta';r,\theta;a)$ of equation (A-6). Hence by equation (C-4) of Appendix C, we have

$$K^0\ [\theta;\ a] = \frac{Ec}{(1-\nu^2)\sqrt{\pi a}} \tag{23}$$

Therefore,

$$H_d = \frac{dK^0}{da_0} + \frac{n_1}{2a_0} K^0 = = \frac{(n-1)\ Ec}{2(1-\nu^2)\sqrt{\pi a_0^3}} \geq 0 \tag{24}$$

for $n \geq 1$. Equation (34) indicates that the translational mode is neutrally stable while perturbations of higher modes are stable. Growth in a circular shape should occur in this case.

The above results suggest that in a displacement controlled tensile test where we fix the amount of remote vertical displacement of a specimen which is constrained against rotation, growth in a circular shape should occur. In the load controlled tensile tests where a fixed, originally centered load is applied to the specimen (weight load, for example), a nonuniformity of growth which begins as an amplification of any initial nonuniformity in the translational shift mode is likely to take place, so that the crack could hardly grow in a uniform manner.

Acknowledgements

The work reported was supported by the ONR Mechanics Division, contract *N*00014 - 85 - *K* - 0405 with Harvard University.

References

Bueckner, H. F., 1970, "A Novel Principle for the Computation of Stress Intensity Factors," *Zeitschrift fuer Angewandte Mathematik und Mechanik*, Vol. 50, pp. 529-546.

Bueckner, H. F., 1973, "Field Singularities and Related Integral Representations," *Mechanics of Fracture 1: Methods of Analysis and Solution of Crack Problems*, Sih, G. C., ed., Noordhoff, Leyden, pp. 239-314.

Galin, L. A., 1953, *Contact Problems in the Theory of Elasticity*, in Russian, Translation by Moss, H., Sneddon, I. N., ed., School of Physical Sciences and Applied Mathematics, North Carolina State College Publications, 1961.

Gao, H., and Rice, J. R., 1986, "Shear Stress Intensity Factors for Planar Crack With Slightly Curved Front," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 774-778.

Gao, H., and Rice, J. R., 1987, "Somewhat Circular Tensile Cracks," *International Journal of Fracture*, pp. 155-174.

Green, A. E., and Zerna, W., 1954, *Theoretical Elasticity*, Oxford University Press.

Kassir, M. K., and Sih, G., 1975, "Three Dimensional Crack Problems," Nordhoff International Publishing, Leyden, The Netherlands.

Meade, K. P., and Keer, L. M., 1984, "Stress Intensity Factors for Semi-Infinite Plane Crack with a Wavy Front," *Journal of Elasticity*, Vol. 14, pp. 79-92.

Neuber, H., 1937, *Kerpspennungslehre*, Springer, Berlin, 1937 and 1958, English Translation available from Edward Bros., Ann Arbor, Michigan.

Panasyuk, V. V., 1962, "Spreading of a Crack Having a Shape Close to Circular in the Plane," *Dopovidi Academii Nauk Ukrainskoi RSR*, in Ukranian, No. 2, pp. 891-895.

Rice, J. R., 1972, "Some Remarks on Elastic Crack Tip Stress Fields," *International Journal of Solids and Structures*, Vol. 8, pp. 751-758.

Rice, J. R., 1985, "First Order Variations in Elastic Fields Due to Variation in Location of a Planar Crack Front," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 571-579.

Sankar, T. S., and Fabrikant, V. I., 1982, "Asymmetric Contact Problem Including Wear for Nonhomogeneous Half-Space." ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 43-46.

Sneddon, I. N., 1951, *Fourier Transforms*, McGraw Hill, New York, 1951.

Stallybrass, M. P., 1981, "On the Concentrated Loading of Certain Elastic Half-Space Problems and Related External Crack Problems. A New Approach," *International Journal of Engineering Science*, Vol. 21, pp. 781-791.

Tada, H., Paris, P. C., and Irwin, G. R., 1973, *The Stress Analysis of Cracks Handbook*, Del Research Corporation, Hellertown, Pa.

APENDIX A

Crack Face Weight Function for an External Circular Crack

From equation (6) to (8) of the text it is clear that the formulae for the crack face opening displacement all require the knowledge of crack face weight function $k(\theta';r,\theta;a)$, i.e., the stress intensity factor induced at θ' along the reference circular front by a pair of unit wedging forces at r, θ on the crack faces outside the circular connection. As pointed out by Stallybrass (1981), this weight function solution depends on the boundary condition at infinity. Examples of such boundary conditions at infinity could be vanishing displacements or traction free conditions. In the following we follow Stallybrass (1981) and categorize the forms of crack face weight function under different boundary conditions at infinity.

The normal stress distribution within the circular connection on the prolongation of the crack plane due to a unit point wedging force pair acting to open the crack faces at location r, θ is the same as the stress field induced within a circular punch area due to an externally applied unit concentrated force on the surface of the half space at the corresponding location. That stress distribution was derived by Galin (1953) using potential theory and later by Stallybrass (1981) using an integral equation approach and, for the case when the solid is restrained against displacement at infinity, the stress at location ρ, ϕ in the connection is

$$\sigma_{yy}^{(u)}(\rho,\phi;r,\theta)=\frac{1}{\pi^2\sqrt{a^2-\rho^2}}\frac{\sqrt{r^2-\rho^2}}{\rho^2-2r\rho\cos(\phi-\theta)+r^2} \tag{A-1}$$

Equation (A-1) enables us to calculate the net force and moments hence generated on each horizontal plane (planes parallel to the crack plane), i.e.,

$$P(r)=-\int_0^a\int_0^{2\pi}\sigma_{yy}^{(u)}(\rho,\phi;r,\theta)\,\rho d\rho d\phi+1$$

$$=\frac{2}{\pi}\cos^{-1}\left(\frac{a}{r}\right) \tag{A-2}$$

The net moment generated about the $\theta + 90°$ axis is

$$M(r)=-\int_0^a\int_0^{2\pi}\rho\cos(\phi-\theta)\sigma_{yy}^{(u)}(\rho,\phi;r,\theta)\rho d\rho d\phi+r$$

$$=\frac{2}{\pi}r\left[\cos^{-1}\left(\frac{a}{r}\right)+\frac{a}{r}\left(1-\frac{a^2}{r^2}\right)^{1/2}\right]\qquad(r>a) \tag{A-3}$$

where we also explicitly emphasized the dependence of P and M on position variable r, where the unit point force acts.

Since the stress distribution (A-1) represents the case when all displacements vanish at infinity, the above calculated net force P in (A-2) and net moment M in (A-3) are balanced by "reaction" force and moment from the restraint at infinity. In the situation when we have traction free boundary condition at infinity, i.e., when there is no restraint against displacement there, the "reaction" force and moment should be taken off by superposing equal, oppositely sensed force and moment at infinity to achieve such boundary conditions. Therefore, two auxiliary problems should be discussed prior to the full presentation of crack face weight functions, namely, the circular connection subjected to remote net centered force P and net moment M about the $\theta + 90°$ axis at infinity. Fortunately the stress distribution induced within the circular connection due to these loadings have been derived by Sneddon (1951) as

$$\sigma_{yy}^{(P)}(\rho,\phi;r,\theta)=\frac{P(r)}{2\pi a\sqrt{a^2-\rho^2}}\qquad(r>a,\rho<a) \tag{A-4}$$

and

$$\sigma_{yy}^{(M)}(\rho,\phi;r,\theta)=\frac{3M(r)\rho\cos(\phi-\theta)}{2\pi a^3\sqrt{a^2-\rho^2}}\qquad(r>a,\rho<a) \tag{A-5}$$

Equation (A-4) and (A-5) represents the stress distribution induced within the circular connection by net tensile force $P(r)$ and net moment $M(r)$ about $\theta + 90°$ axis at infinity. By the rule of superposition discussed before, the total stress distribution within the circular connection area would be (*i*) equation (A-1) if infinity is "clamped", i.e., with no displacements; (*ii*) equations (A-1) plus (A-4) if the solid is allowed only to move freely in the y direction (or vertically) but is fixed against rotation at infinity; (*iii*) equations (A-1) plus (A-5) if the solid could only rotate freely without displacement of points lying along the y axis at infinity; (*iv*) equations (A-1) plus (A-4) plus (A-5) if the solid is free to move without any restraint against displacement at infinity.

Assembling all the discussion made so far, we list the crack face weight function at the following typical boundary conditions at infinity:

***(i)* "Clamped" at Infinity, i.e., Fixed Against Any Displacements.** In this case, the crack face weight function is simply

$$k_d(\theta';r,\theta;a)=\lim_{\rho\to a^-}\sqrt{2\pi(a-\rho)}\sigma_{yy}^{(u)}(\rho,\theta';r,\theta)$$

$$=\frac{\sqrt{(r^2-a^2)/a\pi^3}}{a^2+r^2-2ar\cos(\theta'-\theta)} \tag{A-6}$$

Former discussion shows that the difference between solution under this condition and the solution under the condition of traction free, unrestrained displacement conditions at infinity lies only in terms representing the effect of a net force P as in (A-2) and a net moment M as in (A-3). Therefore, in solving the elasticity problems of a circular connection, in the first step we use above k_d as crack face weight function and in the second step we study separately the effect of the remote tensile forces and/or or moments and combine the results with those of the first step. An example of this way of thinking will be shown in Appendix B in deriving the crack face Green's function.

Solution (A-6) matches the point force solution proposed by Kassir and Sih (1975), although they failed to specify the limitation of the boundary condition at infinity on their solution.

***(ii)* Free Vertical Motion But Fixed Against Rotation at Infinity.** In this case, the remote tensile centered force P of equation (A-2) should be superposed. Therefore, the crack face weight function is

$$k_v(\theta';r,\theta;a)=\lim_{\rho\to a^-}\sqrt{2\pi(a-\rho)}\,[\sigma_{yy}^{(u)}(\rho,\theta';r,\theta)$$

$$+\sigma_{yy}^{(P)}(\rho,\theta';r,\theta)]=\frac{1}{(\pi a)^{3/2}}$$

$$\times\left[\cos^{-1}\left(\frac{a}{r}\right)+\frac{a\sqrt{r^2-a^2}}{a^2+r^2-2ar\cos(\theta'-\theta)}\right] \tag{A-7}$$

Note that this equation (A-7) coincides with the solution proposed by Tada et al. (1973), although they also did not specify the condition under which their solution would be valid.

***(iii)* Free Rotation But Fixed Against Vertical Displacement Along y Axis at Infinity.** In this case, there is an additional contribution from the superposed net moment only. Therefore,

$$k_r(\theta';r,\theta;a)=\lim_{\rho\to a^-}\sqrt{2\pi(a-\rho)}\,[\sigma_{yy}^{(u)}(\rho,\theta';r,\theta)$$

$$+\sigma_{yy}^{(M)}(\rho,\theta';r,\theta)]=\frac{1}{(\pi a)^{3/2}}\left\{\frac{a\sqrt{r^2-a^2}}{a^2+r^2-2ar\cos(\theta'-\theta)}\right.$$

$$\left.+3\left[\frac{r}{a}\cos^{-1}\left(\frac{a}{r}\right)+\left(1-\frac{a^2}{r^2}\right)^{1/2}\right]\cos(\theta'-\theta)\right\} \tag{A-8}$$

***(iv)* Traction Free at Infinity.** In this case, contributions from both net force and net moment should count, and we have the following solution by Stallybrass (1981)

$$k_v(\theta';r,\theta;a)=\lim_{\rho\to a^-}\sqrt{2\pi(a-\rho)}\,[\sigma_{yy}^{(u)}(\rho,\theta';r,\theta)$$

$$+\sigma_{yy}^{(P)}(\rho,\theta';r,\theta)+\sigma_{yy}^{(M)}(\rho,\theta';r,\theta)]=\frac{1}{(\pi a)^{3/2}}$$

$$\times\left\{\left[\cos^{-1}\left(\frac{a}{r}\right)+\frac{a\sqrt{r^2-a^2}}{a^2+r^2-2ar\cos(\theta'-\theta)}\right]+3\left[\frac{r}{a}\cos^{-1}\right.\right.$$

$$\left.\left.\left(\frac{a}{r}\right)+\left(1-\frac{a^2}{r^2}\right)^{1/2}\right]\cos(\theta'-\theta)\right\} \tag{A-9}$$

APPENDIX B

Crack Face Green's Function; Opening Displacement for Circular Connection under Remotely Applied Force and Moment

Equation (9) of the text gives the expression for the crack face Green's function as

$$D(r,\theta;\rho,\phi)=2\frac{1-\nu^2}{E}\int_a^{\min(r,\rho)}\int_0^{2\pi} \times k(\theta';\rho,\phi;a')k(\theta';r,\theta;a')a'\,da'\,d\theta' \quad (B\text{-}1)$$

Temporarily let us impose the condition that all the displacements vanish at infinity. In this case, we could replace crack face weight function k in equation (B-1) by k_d in equation (A-6). Therefore, replacing k by k_d and D by D_d in equation (B-1), we have

$$D_d(r,\theta;\rho,\phi)=2\frac{1-\nu^2}{E\pi^3}\int_a^{\min(r,\rho)}\int_0^{2\pi} \times\frac{\lambda(r/a',\theta-\theta')\lambda(\rho/a',\phi-\theta')}{\sqrt{r^2-a'^2}\sqrt{\rho^2-a'^2}}d\theta\,da' \quad (B\text{-}2)$$

where

$$\times\lambda(k,\phi)=(1-k^2)/(1-2k\cos\phi+k^2) \quad (B\text{-}3)$$

has been introduced for conciseness of the formulae, following Sankar and Fabrikant (1982).

Now consider the following transformation

$$r^*=1/r;\ \rho^*=1/\rho;\ a^*=1/a;\ x=1/a' \quad (B\text{-}4)$$

It may be shown that equation (B-2) becomes,

$$D_d(r,\theta;\rho,\phi)=2\frac{1-\nu^2}{E\pi^3 r\rho}\int_{\max(r^*,\rho^*)}^{a^*}\int_0^{2\pi} \times\frac{\lambda(x/r^*,\theta-\theta')\lambda(x/\rho^*,\phi-\theta')}{\sqrt{x^2-r^{*2}}\sqrt{x^2-\rho^{*2}}}d\theta'\,dx \quad (B\text{-}5)$$

The integral in equation (B-5) has been studied by Gao and Rice (1987). They pointed out that above integral can be reduced to a pseudo-elliptic integral, and by a standard transformation they proved

$$\int_{\max(r^*,\rho^*)}^{a^*}\int_0^{2\pi}\frac{\lambda(x/r^*,\theta-\theta')\lambda(x/\rho^*,\phi-\theta')}{\sqrt{x^2-r^{*2}}\sqrt{x^2-\rho^{*2}}}d\theta'\,dx =\frac{2\pi}{d^*}\arctan\left(\frac{\sqrt{(a^{*2}-r^{*2})(a^{*2}-\rho^{*2})}}{a^*d^*}\right) \quad (B\text{-}6)$$

where $d^*=\sqrt{r^{*2}-2r^*\rho^*\cos(\theta-\phi)+\rho^{*2}}$. Using equations ($B$-4) again to transform back to the original variables, we finally have,

$$D_d(r,\theta;\rho,\phi)=4\frac{1-\nu^2}{E\pi^2 d}\arctan\left(\frac{\sqrt{(r^2-a^2)(\rho^2-a^2)}}{ad}\right) \quad (B\text{-}7)$$

where $d=\sqrt{r^2-2r\rho\cos(\theta-\phi)+\rho^2}$ is the distance between r, θ and ρ, ϕ. Equation (B-7) matches the corresponding formulae given by Galin (1953) and Stallybrass (1981). It is discussed also in Appendix D. The above result looks very similar to the crack face Green's function for internal circular cracks given by Gao and Rice (1987), which is not unexpected because of the similarity of crack face weight functions in this case.

Recall that equation (B-7) represents the crack face Green's function when the condition of vanishing displacement field at infinity is imposed. This has been called case (*i*). We know from Appendix A that by superposing a net force P of equation (A-2) and net moment M of equation (A-3) about $\theta+90°$ axis, we can get rid of the restriction on the displacement field at infinity and achieve the traction free boundary condition there with no restraint against displacement, case (*iv*), and can similarly deal with cases (*ii*) and (*iii*). Therefore, to calculate the crack face Green's function for cases (*ii*), (*iii*), and (*iv*) we need to study two auxiliary problems, namely, crack opening displacement under remote applied centered for F and moment M, where the moment M is now assumed to be about the 90° axis for convenience.

Consider that the described external circular crack system is subjected to a remotely applied tensile force F with a net moment M about the $\theta=90°$ axis. In this case, $K^0\ [\theta';a]$ is given by Tada et a. (1973) (also, see Neuber, 1937, and Sneddon, 1951) as

$$K^0\ [\theta';a]=\frac{aF+3M\cos\theta'}{2a^2\sqrt{\pi a}} \quad (B\text{-}8)$$

Substituting equation (B-8) into equations (5), (6), and using (A-8), one may easily find

$$\frac{\partial\Delta^0 u[r,\theta;a]}{\partial a}=-\frac{(1-\nu^2)}{E(\pi a)^2}\int_0^{2\pi}(F+3M\cos\theta'/a) \times\left\{\left[\cos^{-1}\left(\frac{a}{r}\right)+\frac{a\sqrt{r^2-a^2}}{a^2+r^2-2ar\cos(\theta-\theta')}\right] +3\left[\frac{r}{a}\cos^{-1}\left(\frac{a}{r}\right)+\left(1-\frac{a^2}{r^2}\right)^{1/2}\right]\cos(\theta'-\theta)\right\}d\theta' =-2\frac{(1-\nu^2)}{\pi E}\left\{\frac{F}{a^2}\left[\cos^{-1}\left(\frac{a}{r}\right)+\frac{a}{\sqrt{r^2-a^2}}\right] +\frac{3M\cos\theta}{a^3}\left[\frac{3r}{2a}\cos^{-1}\left(\frac{a}{r}\right)+\frac{3r}{2\sqrt{r^2-a^2}}-\frac{a^2}{2r\sqrt{r^2-a^2}}\right]\right\} \quad (B\text{-}9)$$

Now we integrate over the radius of the connection between a and r, as in going from equation (5) to (6), and find that

$$\Delta u^0[r,\theta;a]=2\frac{(1-\nu^2)}{\pi E}\left\{\frac{F}{a}\cos^{-1}\left(\frac{a}{r}\right) +\frac{3M\cos\theta}{2a^2}\left[\frac{r}{a}\cos^{-1}\left(\frac{a}{r}\right)+\left(1-\frac{a^2}{r^2}\right)^{1/2}\right]\right\} \quad (B\text{-}10)$$

If we replace F by $P(\rho)$ of equation (A-2) and M by $M(\rho)$ (also the axis of the moment is changed to $\phi+90°$) of equation (A-3) of Appendix A, we therefore could rewrite equation (B-10) as

$$\Delta u^0[r,\theta;a]=4\frac{(1-\nu^2)}{\pi^2 Ea}\left\{P(r)P(\rho) +\frac{3M(r)M(\rho)}{2a^2}\cos(\phi-\theta)\right\} \quad (B\text{-}11)$$

If we combine equation (B-11) and equation (B-7), we have, for case (*iv*)

$$D(r,\theta;\rho,\phi)=4\frac{1-\nu^2}{E\pi^2 a}\left\{\frac{a}{d}\arctan\left(\frac{\sqrt{(r^2-a^2)(\rho^2-a^2)}}{ad}\right) +P(r)P(\rho)+\frac{3M(r)M(\rho)}{2a^2}\cos(\phi-\theta)\right\} \quad (B\text{-}12)$$

where $P(r)$, $M(r)$ are given by equations (A-2) and (A-3). We also observe that the symmetry is indeed preserved in equation (B-12). The M terms are deleted in (B-12) to give D for case (*ii*), and the P terms are deleted to give D for case (*iii*).

In going from equation (B-8) to equation (B-10), we have just shown an example of how to calculate the crack opening

displacement from the knowledge of the distribution of stress intensity factor along the crack front. It is, of course, also possible to go in the reverse direction, i.e., to extract the intensity factor distribution from the near front behavior of the opening displacement. When $r \to a^+$, $\cos^{-1}(a/r) \to \sqrt{2(r-a)/a}$ and $(1-a^2/r^2)^{1/2} \to \sqrt{2(r-a)/a}$, so that equation (*B*-10) shows

$$\Delta u^0[r,\theta;a] = 8\frac{(1-\nu^2)}{E}\left(\frac{F}{2a\sqrt{\pi a}} + \frac{3M\cos\theta}{2a^2\sqrt{\pi a}}\right)\sqrt{\frac{r-a}{2\pi}} + O[(r-a)^{3/2}]$$

$$= 8\frac{(1-\nu^2)}{E}\sqrt{\frac{r-a}{2\pi}}\,K^0[\theta;a] + O[(r-a)^{3/2}] \qquad (B\text{-}13)$$

The latter version of equation (*B*-13) represents the known asymptotic behavior of crack opening displacement near the crack front for any tensile crack.

When $r \to \infty$, $\cos^{-1}(a/r) \to \pi/2$,

$$\Delta u^0[r,\theta;a] \sim \frac{(1-\nu^2)}{E}\left\{\frac{F}{a} + \frac{3M}{2a^2}\frac{r\cos\theta}{a}\right\} \qquad (B\text{-}14)$$

Equation (*B*-14) shows that the crack faces far from the front would tend to be linear flat planes (free of stress) with slopes about $\theta = 90°$ of $\pm 3M(1-\nu^2)/(2a^3E)$. These planes would ultimately contact when $M \neq 0$, invalidating the present solution, if the jointed solids are truly unbounded half spaces, although this need not be a problem in practice for finite joined bodies, especially if the mathematically planar crack represents a shallow notch cut-out.

APPENDIX C

Stress Intensity Factor of an External Circular Crack with Fixed Displacement at Infinity

Assume that the elastic solid is subjected to a fixed amount of vertical displacement at infinity as following

$$u_{\pm\infty} = (c + \alpha x)\,\mathrm{sgn}(y) \qquad (C\text{-}1)$$

where $\mathrm{sgn}(y) = y/|y|$ for $y \neq 0$, and the same coordinate system as used in the text is adopted and $x = r\cos\theta$. We solve here for the stress intensity factor induced by this displacement in equation (*C*-1) and the crack opening displacement function.

Referring to equation (*B*-14) of Appendix B, we know that under remotely applied centered force and moment the crack faces far from the crack front would tend to become linear flat planes (free of stress). We also observe that under the imposed remote dispalcement field (*C*-1), crack faces far from the front should approach the same displacement field at infinity because the stresses approach zero there. Therefore imposing a fixed displacement field at infinity ($y \to \pm\infty$) is equivalent to imposing a net tensile force and a net moment at infinity for an external circular crack. Now consider a crack system subject only to a tensile force F and a net moment M at infinity but otherwise traction free. From Appendix B, we know crack opening displacement far from the crack front is

$$\Delta u^0[r,\theta;a] \sim \frac{(1-\nu^2)}{E}\left\{\frac{F}{a} + \frac{3M}{2a^2}\frac{r\cos\theta}{a}\right\}$$

Now let

$$\Delta u^0[r,\theta;a] = u_{+\infty} - u_{-\infty} \qquad (C\text{-}2)$$

Comparing both sides of equation (*C*-2), we find the following relations

$$F = \frac{2Eac}{1-\nu^2}, \qquad M = \frac{4Ea^3\alpha}{3(1-\nu^2)} \qquad (C\text{-}3)$$

Therefore by equation (*B*-8) and (*B*-10) the stress intensity factor induced is

$$K^0[\theta;a] = \frac{E}{1-\nu^2}\frac{c + 2a\alpha\cos\theta}{\sqrt{\pi a}} \qquad (C\text{-}4)$$

and the crack opening displacement is

$$\Delta u^0[r,\theta;a] = \frac{4}{\pi}\left\{c\cos^{-1}\left(\frac{a}{r}\right) + \alpha a\cos\theta\left[\frac{r}{a}\cos^{-1}\left(\frac{a}{r}\right) + \left(1-\frac{a^2}{r^2}\right)^{1/2}\right]\right\} \qquad (C\text{-}5)$$

APPENDIX D

General Displacement Green's Function and Stress Field for Internal and External Circular Cracks

When a three-dimensional crack system is subjected to tensile loading that is symmetric relative to the crack plane, it is known that the elasticity equations and boundary conditions can be satisfied if the displacement and stress field are written as (Galin, 1953; Green and Zerna, 1954; Meade and Keer, 1984)

$$u_y = -2[(1-\nu^2)/E]Y + [(1+\nu)/E]y\partial Y/\partial y$$
$$u_x = [(1+\nu)/E]\partial(F+yY)/\partial x \qquad (D\text{-}1)$$
$$u_z = [(1+\nu)/E]\partial(F+yY)/\partial z$$

where F and Y are harmonic functions related by $\partial F/\partial y = (1-2\nu)Y$. The coordinates are set up in the same manner as in the text with the crack on the $y = 0$ plane. The stress components that enter crack surface boundary conditions are calculated from stress-strain relations as

$$\sigma_{yy} = -\partial Y/\partial y + y\partial^2 Y/\partial y^2$$
$$\sigma_{yz} = y\partial^2 Y/\partial y\partial x, \quad \sigma_{yz} = y\partial^2 Y/\partial y\partial z \qquad (D\text{-}2)$$

It is seen from equations (*D*-2) that there is no shear traction on $y = 0$. Thus the problem of loading on the crack face is one of finding a function Y satisfying $\nabla^2 Y = 0$, vanishing at infinity (at least for case (*i*)), and generating stress $\sigma(x, z)$ and opening gap $\Delta u(x, z)$ on $y = 0$ given by

$$\sigma(x,z) = -\partial Y/\partial y|_{y=0} \text{ and } \Delta u = -[4(1-\nu^2)/E]Y|_{y=0^+} \qquad (D\text{-}3)$$

(a) Internal Circular Cracks. Now we consider the elasticity problem of a three-dimensional elastic solid with an internal circular crack of radius equal to a subjected to a point force pair in the $\pm y$ directions acting at ξ, η on the crack faces. According to equations (*D*-1), (*D*-2), and (*D*-3), we formulate following problem,

$$\nabla^2 Y = 0$$
$$Y = 0 \text{ when } x^2 + z^2 \geq a; \quad y = 0$$
$$\partial Y/\partial y = -\delta(x-\xi)\delta(z-\eta) \text{ when } x^2 + z^2 \leq a; \; y = 0 \qquad (D\text{-}4)$$
$$Y = 0 \text{ at } \infty$$

Let us denote the solution to equation (*D*-4) as $Y = H(x, y, z; \xi, \eta)$. It is known (Galin, 1953) that

$$H(x,y,z;\xi,\eta) = -\frac{1}{\pi^2 d}\arctan\left\{\frac{\sqrt{(a^2-\xi^2-\eta^2)(a^2-x^2-y^2-z^2+R)}}{\sqrt{2}ad}\right\} \qquad (D\text{-}5)$$

where

$$d^2=(x-\xi)^2+(z-\eta)^2+y^2$$

$$R^2=(a^2-x^2-y^2-z^2)^2+4a^2y^2 \qquad (D\text{-}6)$$

Replacing Y by H in equations (D-1) and (D-2), we then get the displacement Green's function and stress field at an arbitrary location in space. Specifically, the crack face Green's function is seen to be

$$D(x,z;\xi,\eta)=\frac{4(1-\nu^2)}{E}H(x,0,z;\xi,\eta)$$

$$=\frac{4(1-\nu^2)}{\pi^2 Ed}\arctan\left\{\frac{\sqrt{(a^2-\xi^2-\eta^2)(a^2-x^2-y^2)}}{ad}\right\} \qquad (D\text{-}7)$$

If expressed in polar coordinates, equation (D-7) becomes,

$$D(r,\theta;\rho,\phi)=\frac{4(1-\nu^2)}{\pi^2 Ed}\arctan\left\{\frac{\sqrt{(a^2-\rho^2)(a^2-r^2)}}{ad}\right\} \qquad (D\text{-}8)$$

where d reduced to $\sqrt{r^2-2r\rho\cos(\theta-\phi)+\rho^2}$. Equation ($D$-8) coincides with the solution derived through the perturbation analysis by Gao and Rice (1987, Appendix A).

(b) External Circular Cracks. Now we consider a similar crack system but with an external circular crack, or a circular connection of radius equal to a subjected to a point force pair in the $\pm y$ directions acting at ξ, η on the crack faces with zero displacement at infinity. In an analogous way we formulate the problem as solving

$$\nabla^2 Y=0$$

$$Y=0 \text{ when } x^2+z^2\le a;\ y=0$$

$$\partial Y/\partial y=-\delta(x-\xi)\delta(z-\eta) \text{ when } x^2+z^2\ge a;\ y=0 \qquad (D\text{-}9)$$

$$Y=0 \quad \text{at } \infty$$

Note that in this formulation we imply that there is no displacements at infinity and hence the solution thus generated can only be applied to case (i) of the text. The solution to equation (D-9), dentoed here as $Y=L(x, y, z; \xi, \eta)$, was given also by Galin (1953) as

$$L(x,y,z;\xi,\eta)=-\frac{1}{\pi^2 d}\arctan$$

$$\left\{\frac{\sqrt{(\xi^2+\eta^2-a^2)(x^2+y^2+z^2-a^2+R)}}{\sqrt{2}ad}\right\} \qquad (D\text{-}10)$$

where d and R are given by equations (D-6). Similarly if we replace Y by L in equations (D-1) and (D-2) the displacement Green's function and stress field at an arbitrary location in an elastic solid with an external crack are generated. Specifically the crack face Green's function can be extracted and expressed in polar coordinates as

$$D_d(r,\theta;\rho,\phi)=\frac{4(1-\nu^2)}{\pi^2 Ed}\arctan\left\{\frac{\sqrt{(\rho^2-a^2)(r^2-a^2)}}{ad}\right\} \qquad (D\text{-}11)$$

Equation (D-11) coincides with (B-7) of Appendix B, where the crack face Green's function is derived by the perturbation formalism.

Let us note that by using the solutions of equations (D-5) and (D-10) for Y in equations (D-1), we can compute the displacements u_z, u_y, and u_z at (x, y, z) due to unit opening point forces acting on the crack faces at (ξ, η, 0), for the respective internal and external circular crack cases. By the elastic reciprocal theorem, those very same results for u_z, u_y, and u_z also represent the opening gaps Δu on the crack faces induced at (ξ, η, 0) by unit point forces at (x, y, z) in the respective x, y, and z directions. But from the knowledge of that opening gap Δu in the vicinity of the crack front, one may also calculate (e.g., equation (B-13)) the tensile mode stress intensity factors induced by the unit point forces at (x, y, z) in the respective x, y, and z directions. These stress intensity factor defined the x, y, and z components of the tensile mode vector weight function $\mathbf{h}$ as introduced by Rice (1972, 1985). Hence, although we do not further pursue the details here, the results of this Appendix allow calculation of the vector tensile mode weight function at general field points for internal and external circular cracks.

The Dynamic Energy Release Rate for a Steadily Propagating Antiplane Shear Crack in a Linearly Viscoelastic Body

J. R. Walton
Department of Mathematics,
Texas A & M University,
College Station, Texas 77840
Mem. ASME

The steady-state propagation of a semi-infinite, antiplane shear crack is reconsidered for a general, infinite, homogeneous and isotropic linearly viscoelastic body. As with an earlier study, the inertial term in the equation of motion is retained and the shear modulus is only assumed to be positive, continuous, decreasing, and convex. A Barenblatt type failure zone is introduced in order to cancel the singular stress, and a numerically convenient expression for the dynamic Energy Release Rate (ERR) is derived for a rather general class of crack face loadings. The ERR is shown to have a complicated dependence on crack speed and material properties with significant qualitative differences between viscoelastic and elastic material. The results are illustrated with numerical calculations for both power-law material and a standard linear solid.

1 Introduction

A central issue in fracture mechanics is the development of fracture criteria. A great many experimental and analytical studies have addressed this topic in the nearly sixty years since Griffith's pioneering work. One fact that has emerged from this effort is that the choice of a fracture criterion is very much dependent upon the particular scenario considered. For example, the notion of a critical Stress Intensity Factor (SIF) has provided a highly successful criterion for quasi-static crack propagation in linearly elastic material. Important factors for the success of the SIF in this setting are that it is often easily computed and that the Energy Release Rate (ERR) can be determined in a simple manner from it. However, such is not necessarily the case for dynamically propagating cracks in viscoelastic material. Indeed, it is shown in this paper that for such models, the ERR may have a much more complicated dependence upon crack speed and the viscoelastic moduli.

Several studies of dynamic viscoelastic crack propagation have appeared in the literature, beginning with Willis' treatment of a steadily propagating, semi-infinite, Mode III crack in an infinite viscoelastic body (Willis, 1967). Using the Wiener-Hopf technique, Willis was able to construct the SIF for material that can be modelled as a standard linear solid. In 1977 Atkinson and Coleman used a perturbation technique to construct approximations to the SIF for a semi-infinite, Mode I crack propagating in steady-state through a viscoelastic layer, also modelled as a standard linear solid. In 1979, Atkinson and Popelar employed the Wiener-Hopf technique to analyze a semi-infinite, Mode III crack that suddenly begins to propagate at a constant speed through a viscoelastic layer subjected to constant antiplane loading. A fairly general constitutive relation in terms of differential operators is assumed. In 1980, Popelar and Atkinson consider the corresponding Mode I problem. In each of these last two papers, Atkinson and Popelar produced formal expressions for the SIF which were evaluated only for a standard linear solid. Using a local work argument, they also exhibited expressions for the ERR that are based upon the singular stress solution.

In 1982, using somewhat different analytical methods, Walton reconsidered the problem in Willis (1967). Under quite general constitutive assumptions, considerably more so than those adopted by Atkinson et al. which exclude, for example, the important class of power-law material, Walton constructed a simple expression for the SIF which exhibits quite clearly and precisely its dependence upon crack speed and material properties. More recently, this analysis was generalized to the case of a Mode III crack in a layer (Walton, 1985).

In this paper the methods in Walton (1982) are extended to carry out the calculation of the ERR. The notion of ERR considered here provides a phenomenologically meaningful and mathematically convenient fracture criterion for dynamic viscoelastic crack propagation. An essential feature of the model is that a Barenblatt type failure zone is assumed to exist at the crack tip. Use of the Barenblatt model not only simplifies the calculation of the ERR but, much more importantly, it introduces an additional length scale and produces bounded stresses and strains which result in the ERR having a fundamentally different dependence upon crack speed and

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 48th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October 28, 1985; final revision, February 17, 1987.

material properties from that derived by Atkinson and Popelar using the singular fields. For example, it is shown below that in the limit of vanishing failure zone length, the ERR equals the product of the square of the SIF and a simple function of crack speed and the glassy values of the viscoelastic shear modulus. This is in agreement with the result of Atkinson and Popelar (1979) and also Kostrov and Nikitin (1970). In contrast, when the failure zone has a nonzero length, the ERR is the product of the square of the SIF and a function of crack speed and the full form of the viscoelastic shear modulus, not just the glassy properties. Moreover, it is shown here that while the SIF is a monotonically decreasing function of crack speed, the ERR can exhibit much more complicated, nonmonotonic behavior depending upon combined viscoelastic and inertial effects. As discussed later, this has interesting implications with regard to stable versus unstable steady-state crack speeds.

The Barenblatt model has been utilized extensively in elastic fracture mechanics. This author would be remiss in not calling attention to Barenblatt's seminal 1962 paper. More recently, the Barenblatt model has been applied to studies of quasi-static viscoelastic fracture, most notably in Knauss (1973) and Schapery (1975), in which it was observed that whether a failure zone is incorporated or not significantly affects the behavior of the ERR.

Utilizing the techniques of this paper, this author has recently completed the analysis of the considerably more complicated corresponding Mode I problem. Moreover, L. Schovanec and this author have also recently calculated the ERR for two parallel, interacting Mode III cracks. Both of these investigations are the subject of forthcoming papers.

The specific boundary value problem considered here is that corresponding to the steady propagation (to the right) with speed V of a semi-infinite, antiplane shear crack in a general, homogeneous and isotropic, linearly viscoelastic body. The shear modulus, $\mu(t)$, is assumed only to be a positive, nonincreasing, and convex function of time, t. The governing equation of motion for the out-of-plane displacement, u_3, is

$$\mu^* d\Delta u_3 = \rho\, u_{3,tt}$$

where Δ is the two-dimensional Laplacian, $\Delta = (\partial^2/\partial x_1^2) + (\partial^2/\partial x_2^2)$, and $\mu^*\, d\epsilon$ denotes the Riemann-Stieltjes convolution

$$\mu^* d\epsilon = \int_{-\infty}^{t} \mu(t-\tau) d\epsilon(\tau).$$

Upon adoption of the Galilean variables $x = x_1 - Vt$, $y = x_2$, the boundary conditions may be written

$$\begin{aligned} \sigma_{23}(x,0) &= \frac{\partial}{\partial y}(\mu^* du_3) = f(x), && x<0 \\ u_3(x,0) &= 0, && x>0 \\ \sigma_{ij}(x,y) &\to 0, && x^2+y^2 \to \infty \end{aligned}$$

where σ_{ij} are the stress components and $f(x)$ is a system of tractions moving with the crack.

The starting point of the present investigation is the solution derived in Walton (1982) for the above boundary value problem. It was shown in Walton (1982) that two cases arise naturally in constructing the solution: $0 \le V < C^*$ and $C^* < V < C$ where $C^* = (\mu(\infty)/\rho)^{1/2}$ and $C = (\mu(0)/\rho)^{1/2}$ are the elastic shear wave speeds corresponding to the equilibrium and glassy values of the shear modules $\mu(t)$. For $0 \le V < C^*$, the stress field is that for static elastic fracture and is therefore independent of crack speed and material properties. Whereas, for $C^* < V < C$, the stress field is both speed and material dependent. Specifically, the SIF, K, is given by

$$K = \begin{cases} \dfrac{-1}{\pi}\displaystyle\int_{-\infty}^{0} \sigma_{23}^-(x,0)\,|x|^{\frac{-1}{2}}\,dx, & 0 \le V < C^* \\ -\dfrac{1}{\pi}\displaystyle\int_{-\infty}^{0} \sigma_{23}^-(x,0)\,|x|^{\frac{-1}{2}}\, e^{q_o x} dx, & C^* < V < C \end{cases} \tag{1}$$

where $g^-(x)$ ($g^+(x)$) denotes the restriction of $g(x)$ to $x < 0$ ($x > 0$) and q_o is the unique positive constant such that

$$V q_o \int_0^{\infty} \mu(t) e^{-q_o V t} dt = (V/C)^2 \mu(0). \tag{2}$$

In order to calculate the ERR, the object of principal interest in the present study, it is necessary, as discussed earlier, to modify the above boundary value problem by the introduction of a Barenblatt type failure zone. Specifically, it is now assumed that two loads are acting on the crack faces: the applied (external) tractions $\sigma_{23}^-(x, 0)$ discussed above, but now denoted $\sigma_e^-(x)$, and cohesive (failure) stresses $\sigma_f^-(x)$ acting in a failure zone of length a_f immediately behind the crack tip. The only assumptions about $\sigma_f^-(x)$ are that a_f is small relative to some length scale a_e associated with $\sigma_e^-(x)$ and that $K_e + K_f = 0$ where K_e and K_f are the SIF's corresponding to σ_e^- and σ_f^-, respectively. Hence the effect of the failure zone is to cancel the singular stresses ahead of the crack tip and thereby produce a cusp-shaped crack profile behind the tip.

The ERR, G, for steady-state crack propagation is now easily shown to be given by

$$G = \int_{-a_f}^{0} \sigma_f^-(x)\, u_{3,1}^-(x,0) dx \tag{3}$$

where $u_3(x, 0)$ is the crack face displacement corresponding to the combined loading $\sigma_e^- + \sigma_f^-$. Thus G has the interpretation of the work input to the crack tip, i.e., the energy available to the crack tip for propagating the crack. Unlike for elastic material, G for viscoelastic material is not merely a simple function of K. Rather, as is evident from Walton (1982), $u_{3,1}^-(x, 0)$ has a complicated dependence upon the loading $\sigma_e^-(x, 0)$, making impracticable the direct numerical evaluation of equation (3). In the next section, a computationally convenient expression for equation (3) is derived for a special, but still fairly general, class of loadings σ_e^- and σ_f^-.

2 The Calculation of G

For simplicity of argument and clarity of result, the ERR, G, given in equation (3) will be calculated first for a simple special case. Specifically, the external load, $\sigma_e^-(x)$ and failure zone stresses, $\sigma_f^-(x)$ will be assumed to have the forms

$$\begin{aligned} \sigma_e^-(x) &= L_e e^{x/a_e} \\ \sigma_f^-(x) &= -L_f e^{x/a_f} \end{aligned} \qquad -\infty < x < 0 \tag{4}$$

where $a_f/a_e \ll 1$. (cf Sills and Benveniste, 1981). For a_f/a_e small enough, the fact that $\sigma_f^-(x)$ does not have support in some small, compact, interval behind the crack tip will have a negligible effect on the results. The assumptions (4) then clearly incorporate the salient features of the Barenblatt model, namely, a set of cohesive stresses and associated length scale a_f and a length scale a_e associated with the applied load σ_e^- such that σ_f^- cancels the singular stresses produced by σ_e^- and $a_f/a_e \ll 1$. It should be noted that in this case, equation (3) is replaced by

$$G = \int_{-\infty}^{0} \sigma_f^-(x) u_{3,1}^-(x,0) dx. \tag{5}$$

In order to present the results in a suitable nondimensional form, it is useful to introduce certain parameters. First a characteristic timescale, τ, is defined and the shear modulus given the form

$$\mu(t)=\mu_\infty m(t/\tau)$$

where $m(s)$ is a nondimensional function of s with $\lim_{s\to\infty} m(s) = 1$ and $\mu_\infty = \lim_{t\to\infty} \mu(t)$. Also useful are the nondimensional parameters α, β, γ, ϵ and λ defined by

$$\alpha\equiv C^*\tau/a_e\ \beta\equiv q_o a_e\ \gamma\equiv V/C^*\ \epsilon\equiv a_f/a_e\ \lambda\equiv L_f/L_e. \tag{6}$$

It should be noted that α and β can assume any positive value while γ must be such that $0 < \gamma < C/C^*$. Also, $\beta = 0$ whenever $0 < V \le C^*$.

The desired expression for G is most easily expressed in terms of the Carson transform, $\bar{m}(s)$, of $m(t)$ defined by

$$\bar{m}(s)=m(0)+\int_0^\infty e^{-ts}dm(t).$$

It is useful to record for future reference the easily derived formula

$$\bar{\mu}(r)=\mu_\infty\bar{m}(r\tau).$$

It will now be shown that for loads of the form (4), G is given by

$$G=\frac{K^2}{2\mu_\infty}\left(\frac{1-\epsilon}{1+\epsilon}\right)\frac{1}{\bar{m}(\alpha\gamma/\epsilon)}\left[\frac{1-(\beta\epsilon)^2}{1-\gamma^2/\bar{m}(\alpha\gamma/\epsilon)}\right]^{\frac{1}{2}} \tag{7}$$

where $K = K_e = -K_f$ is the dynamic SIF. From equations (1) and (4) it follows easily that

$$K=\frac{L_f\sqrt{a_f}}{\sqrt{1+\beta\epsilon}}=\frac{L_e\sqrt{a_e}}{\sqrt{1+\beta}}. \tag{8}$$

The derivation of equation (7) from (5) utilizes the Fourier transform $\hat{f}$ of a function f. Specifically,

$$\hat{f}(p)=\int_{-\infty}^{\infty} f(x)e^{ixp}dx$$

with inverse, $\check{f}(x)$, given by

$$\check{f}(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} f(p)e^{-ixp}dp.$$

Applying the Parseval formula for the Fourier transform, it follows from equation (5) that

$$G=\int_{-\infty}^{\infty}\check{\sigma}_f^-(p)\hat{u}_{3,1}(p)dp. \tag{9}$$

In consideration of equation (4), a straight forward calculation shows that

$$\hat{\sigma}_f^-(p)=\frac{-a_fL_f}{(1+ia_fp)}$$

$$\check{\sigma}_f^-(p)=\frac{-a_fL_f}{2\pi(1-ia_fp)}. \tag{10}$$

The integral in equation (9) may be readily evaluated using residues since, from equations (10), it is clear that $\check{\sigma}_f^-(p)$ has a meromorphic extension to the lower complex half-plane with a simple pole at $-i/a_f$. Moreover, $u_{3,1}(x, 0)$ vanishes for $x > 0$ from which it follows that $\hat{u}_{3,1}(p) = \hat{u}_{3,1}^-(p)$ has an analytic extension, $F^-(z)$, to the lower half-plane with

$$\lim_{q\to\infty} F^-(p-iq)=0.$$

Consequently, for G one has

$$G=L_fF^-(-i/a_f). \tag{11}$$

It remains to evaluate $F^-(-i/a_f)$. To this end use will be made of the following formulas derived in Walton (1982):

$$\hat{u}_{3,1}(p)=\hat{\sigma}(p)/G(p), \tag{12}$$

$$G(p)=-i\,\mathrm{sgn}(p)\bar{\mu}(iVp)\gamma_1(iVp), \tag{13}$$

and

$$\gamma_1(ipV)=(1-\gamma^2/\bar{m}(iVp\tau))^{1/2}. \tag{14}$$

Here one has

$$\hat{\sigma}(p)=\hat{\sigma}^-(p)+\hat{\sigma}^+(p)$$

$$\sigma^-(x)=\sigma_e^-(x)+\sigma_f^-(x)$$

where $\sigma^+(x)$ ($\sigma^-(x)$) denotes the restriction of $\sigma(x)$ to the half-line $x > 0$ ($x < 0$). It should be noted that $\sigma^+(x)$ is the nonsingular stress field ahead of the crack that results from superposing σ_e^- and σ_f^-. Moreover, it is shown in Walton (1982) that

$$\hat{\sigma}^+(p)=\lim_{Im(z)\to 0+} F^+(z)$$

where

$$F^+(z)=-(q_o-iz)^{1/2}\frac{1}{2\pi i}\int_{-\infty}^{\infty}\hat{\sigma}^-(\tau)(q_o-i\tau)^{-1/2}\frac{d\tau}{(\tau-z)}, \tag{15}$$

with q_o given in equation (2). To evaluate equation (15), consider first

$$\hat{\sigma}_e^+(p)=\lim_{Im(z)\to 0+} F_e^+(z),$$

in which

$$F_e^+(z)=-(q_o-iz)^{1/2}\frac{1}{2\pi i}\int_{-\infty}^{\infty}\hat{\sigma}_e^-(\tau)(q_o-i\tau)^{-1/2}\frac{d\tau}{(\tau-z)}. \tag{16}$$

From the analog of equations (10) for $\hat{\sigma}_e^-(p)$, one sees that $\hat{\sigma}_e^-(p)$ has a simple pole at i/a_e. Since the branch of $(q_o-i\tau)^{-1/2}$ must be chosen to be analytic in the upper half-plane (see Walton, 1982), the integral in equation (16) can be calculated using residues. Thus for $\mathrm{Im}(z) > 0$,

$$F_e^+(z)=-(q_o-iz)^{1/2}[-iL_e(q_o+1/a_e)^{-1/2}/(i/a_e-z)+\hat{\sigma}_e^-(z)(q_o-iz)^{-1/2}].$$

Now letting $Im(z) \to 0$ and adding $\hat{\sigma}_e^-(p)$ to $\hat{\sigma}_e^+(p)$ there results

$$\begin{aligned}\hat{\sigma}(p)&=F_e^+(p)+\hat{\sigma}_e(p)\\&=iL_e(q_o-ip)^{1/2}(q_o+1/a_e)^{-1/2}/(i/a_e-p)\\&=a_eL_e(\beta/(1+\beta))^{1/2}(1-ip/q_o)^{1/2}/(1+ipa_e).\end{aligned} \tag{17}$$

Similarly, for $\hat{\sigma}_f(p)$ one can show that

$$\hat{\sigma}_f(p)=-a_fL_f(\beta\epsilon/(1+\beta\epsilon)^{1/2}(1-ip/q_o)^{1/2}/(1+ipa_f). \tag{18}$$

Combining equations (4), (8), (17), and (18) one concludes that

$$\hat{\sigma}(p)=-a_eL_e(\beta/(1+\beta))^{1/2}\frac{(1-\epsilon)(1-ip/q_o)^{1/2}}{(1+ia_ep)(1+ia_fp)}. \tag{19}$$

Substition of equation (19) into (12) gives

$$F^-(p)=\hat{u}_{3,1}(p)=a_eL_e(\beta/(1+\beta))^{1/2}(1-\epsilon)\Phi_1(p)\Phi_2(p) \tag{20}$$

with

$$\Phi_1(p)=(1+ia_ep)^{-1}(1+ia_fp)^{-1}$$

and

$$\Phi_2(p)=(1-ip/q_o)^{1/2}/G(p).$$

$\Phi_1(p)$ is obviously an analytic function in the lower half-plane and in Walton (1982) it was shown that the branches of $(1-ip/q_o)^{1/2}$ and $\gamma_1(iVp)$ must be chosen so that $\Phi_2(p)$ is also analytic there. Thus one may substitute equation (20) into (11) making use of equation (13) and (14) to conclude that

$$G=a_eL_eL_f\frac{(\beta/(1+\beta))^{1/2}(1-\epsilon)}{2(1+a_e/a_f)\bar{\mu}(V/a_f)}\left|\frac{1-1/(q_oa_f)}{1-\gamma^2/\bar{m}(V/a_f)}\right|^{1/2}. \tag{21}$$

Combining equations (6), (8), and (21), there finally results the desired formula (7).

The derivation given above is easiy modified to produce an analog to formula (7) for more general loads of the form

$$\sigma_e^-(x) = L_e \int_0^\infty e^{tx/a_e} d\,h_e(t)$$

$$\sigma_f^-(x) = -L_f \int_0^\infty e^{tx/a_f} d\,h_f(t) \tag{22}$$

where $h_e(t)$ and $h_f(t)$ are arbitrary signed (not necessarily positive) measures restricted only to the extent that the required integrals converge. For example, the special cases (4) correspond to $d\,h_e(t) = d\,h_f(t) = \delta(t-1)$, the Dirac measure concentrated at $t = 1$. As further examples, $d\,h_e(t) = \sin(t)dt$ and $d\,h_e(t) = \cos(t)dt$ produce

$$\sigma_e^-(x) = \frac{L_e}{1+(x/a_e)^2} \text{ and } \sigma_e^-(x) = \frac{(x/a_e)L_e}{1+(x/a_e)^2},$$

respectively.

Substitution of equations (22) into (1) followed by an interchange of the order of integration results in

$$|K_e| = \frac{L_e\sqrt{a_e}}{\sqrt{\pi}} \int_0^\infty (t+\beta)^{-1/2} d\,h_e(t)$$

$$= \frac{L_f\sqrt{a_f}}{\sqrt{\pi}} \int_0^\infty (t+\epsilon\beta)^{-1/2} d\,h_f(t)$$

$$= |K_f|$$

where, as before, $\beta \equiv a_e q_o$. Corresponding to equations (10) there is

$$\hat{\sigma}_e^-(p) = a_e L_e \int_0^\infty (t+ipa_e)^{-1} d\,h_e(t)$$

$$\check{\sigma}_e^-(p) = \frac{a_e L_e}{2\pi} \int_0^\infty (t-ia_e p)^{-1} d\,h_e(t). \tag{23}$$

Lines (15) and (16) are still valid but with $\hat{\sigma}_e^-$ given now by equations (23). In particular, one has for $I_m(z) > 0$ and after an interchange of integration that

$$F_e^+(z) = -(q_o - iz)^{1/2} \frac{1}{2\pi i} a_e L_e \int_0^\infty dh_e(\tau) \int_{-\infty}^\infty (q_o - it)^{-1/2} (\tau + ita_e)^{-1} \frac{dt}{(t-z)}. \tag{24}$$

The integrand in the inner integral in equation (24) is analytic for $I_m(t) > 0$ except for simple poles at $t = i\tau/a_e$ and $t = z$. Calculating the inner integral by residues, there then results

$$\frac{1}{2\pi i} \int_{-\infty}^\infty (q_o - it)^{-1/2} (\tau + ita_e)^{-1} \frac{dt}{(t-z)} = [(q_o - iz)^{-1/2} - (q_o + \tau/a_e)^{-1/2}]/(\tau + iza_e).$$

If one now lets $I_m(z) \to 0$ in this last result and makes use of equation (23*a*) and equation (24) it follows that

$$\hat{\sigma}_e(p) = \hat{\sigma}_e^+ + \hat{\sigma}_e^- = a_e L_e (1 - ip/q_o)^{1/2} \int_0^\infty (\beta/(\tau+\beta))^{1/2} \frac{dh_e(\tau)}{(\tau + ipa_e)}. \tag{25}$$

For $\hat{\sigma}_f(p)$ one has

$$\hat{\sigma}_f(p) = -a_f L_f (1 - ip/q_o)^{1/2} \int_0^\infty (\beta\epsilon/(\tau+\beta\epsilon))^{1/2} \frac{dh_f(\tau)}{(\tau + ipa_f)},$$

the analog of equation (18), which when combined with equation (25), gives

$$\hat{\sigma}(p) = \hat{\sigma}_e(p) + \hat{\sigma}_f(p) = (1 - i\,p/q_o)^{1/2} H(p) \tag{26}$$

with $H(p) \equiv \int_0^\infty \frac{a_e L_e (\beta/(\tau+\beta))^{1/2}}{(\tau + ia_e p)} dh_e(\tau)$

$$- \int_0^\infty \frac{a_f L_f (\epsilon\beta/(\tau+\epsilon\beta))^{1/2}}{(\tau + ia_f p)} dh_f(\tau). \tag{27}$$

It is easily seen that $H(p)$ is analytic for $I_m(p) < 0$ and $H(p) \to 0$ as $I_m(p) \to -\infty$. After substitution of equations (11), (12), (23*b*), (26), and (27) into equation (9) and an interchange of integration one obtains

$$G = a_f L_f \int_0^\infty dh_f(t) \frac{1}{2\pi} \int_{-\infty}^\infty \Psi(t,p)\,dp \tag{28}$$

where

$$\Psi(t,p) = \frac{H(p)(1 - ip/q_o)^{1/2}}{(t - ipa_f)G(p)}.$$

As before, the branches of $\gamma_1(ipV)$ and $\sqrt{1 - ip/q_o}$ can be chosen so that $i\,\mathrm{sgn}(p)\sqrt{1 - i\,p/q_o}/\gamma_1(ipV)$ is analytic for $I_m(p) < 0$. Thus, $\Psi(t, p)$ is analytic for $I_m(p) < 0$, except for a simple pole at $p = -it/a_f$. Evaluating the inner integral in equation (28) by residues, yields finally

$$G = \frac{L_f}{\mu_\infty} \int_0^\infty dh_f(t) \frac{H(-it/a_f)}{\bar{m}\left(\frac{\alpha\gamma t}{\epsilon}\right)} \left[\frac{1 - t/(\beta\epsilon)}{1 - \gamma^2/\bar{m}(\alpha\gamma t/\epsilon)}\right]^{1/2} \tag{29}$$

with $H(-it/a_f)$ defined in equation (27).

The two integrations required to evaluate equation (29) make it much more cumbersome to calculate numerically than equation (7), though still much easier than calculating G directly from either equations (3) or (5). However, since $\epsilon \ll 1$, it is not unreasonable to take $dh_f(t) = \delta(t-1)$, i.e., to take the simple form (4*b*) for $\sigma_f^-(x)$, since the details of the failure zone stress are not significant. Formula (29) then simplifies greatly to

$$G = \frac{L_f}{\mu_\infty} \frac{H(-i/a_f)}{\bar{m}(\alpha\gamma/\epsilon)} \left[\frac{1 - 1/(\beta\epsilon)}{1 - \gamma^2/\bar{m}(\alpha\gamma/\epsilon)}\right]^{1/2}$$

$$H(-i/a_f) = \frac{-a_f L_f}{2} (\epsilon\beta/(1+\epsilon\beta))^{1/2} + a_e L_e \beta^{1/2} \int_0^\infty \frac{dh_e(\tau)}{(\tau + 1/\epsilon)(\tau+\beta)^{1/2}}$$

subject to the auxiliary constraint

$$L_f \left(\frac{a_f}{1+\beta\epsilon}\right)^{1/2} = L_e (a_e)^{1/2} \int_0^\infty \frac{dh_e(t)}{(t+\beta)^{1/2}}.$$

In the next section, the qualitative behavior of G is investigated by considering the special cases of a power-law material and a standard linear solid.

3 Numerical Examples

The formula (7) will now be applied to the special cases of a standard linear solid and power-law material. First considered is the standard linear solid, which is modelled by a constant Poisson's ratio and a shear modulus, $\mu(t)$, of the form

$$\mu(t) = \mu_\infty(1 + \eta e^{-t/\tau}) = \mu_\infty m(t/\tau).$$

It follows that $1 + \eta = (C/C^*)^2$ and $\bar{m}(s)$ is given by

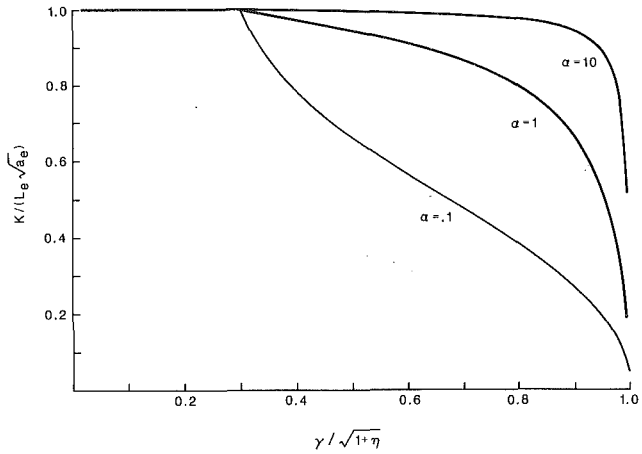


Fig. 1 k versus γ' for a standard linear solid with $\eta = 10$ and $\alpha = 0.1, 1.0, 10.0$

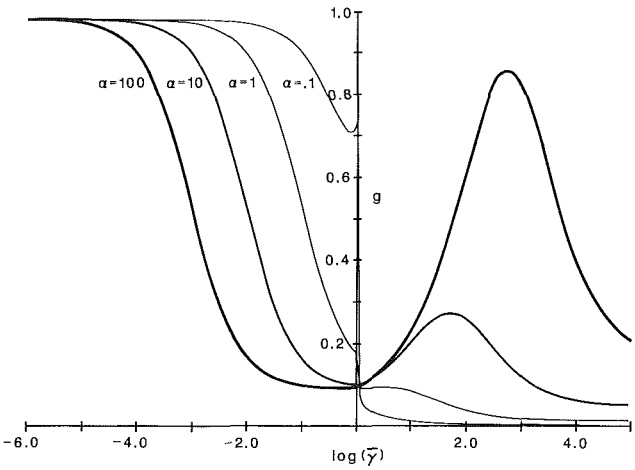


Fig. 2 g versus $\log(\bar{\gamma})$ for a standard linear solid with $\eta = 10$, $\epsilon = 0.01$, $\alpha = 0.1, 1.0, 10.0, 100.0$

$$\tilde{m}(s) = \frac{1 + s(1+\eta)}{(1+s)}. \tag{30}$$

From equations (2) and (30) one easily shows that

$$q_o = \frac{(\gamma^2 - 1)}{\tau V((1+\eta) - \gamma^2)}. \tag{31}$$

It should be noted that the restriction $0 < V < C$ corresponds to $0 \le \gamma^2 < 1 + \eta$, and moreover, that $q_o = 0$ for $0 \le \gamma \le 1$. From equations (7), (8), (13), (30), and (31), one readily obtains the formula

$$G = \frac{a_e L_e^2}{2\mu_\infty}\, g(\alpha,\gamma,\eta,\epsilon) \tag{32}$$

where

$$g(\alpha,\gamma,\eta,\epsilon) = \left(\frac{1-\epsilon}{1+\epsilon}\right) \frac{(\epsilon + \alpha\gamma)}{(\epsilon + \alpha\gamma(1+\eta))^{1/2}((1-\gamma^2)\epsilon + \alpha\gamma(1+\eta-\gamma^2))^{1/2}} \quad 0 \le \gamma \le 1 \tag{33}$$

and

$$g(\alpha,\gamma,\eta,\epsilon) = \left(\frac{1-\epsilon}{1+\epsilon}\right) \frac{(\epsilon + \alpha\gamma)(1+\beta\epsilon)^{1/2}}{(1+\beta)(\alpha\gamma(\eta+1-\gamma^2))^{1/2}(\epsilon + \alpha\gamma(1+\eta))^{1/2}} \quad 1 < \gamma < (1+\eta)^{1/2}.$$

Figure 1 displays a normalized SIF, $k \equiv K/(L_e\sqrt{a_e})$, which from equation (8) is seen to be just $k = (1+\beta)^{-1/2}$. From equation (31), β is seen to be given by

$$\beta = \frac{(\gamma^2 - 1)}{\alpha\gamma((1+\eta) - \gamma^2)}.$$

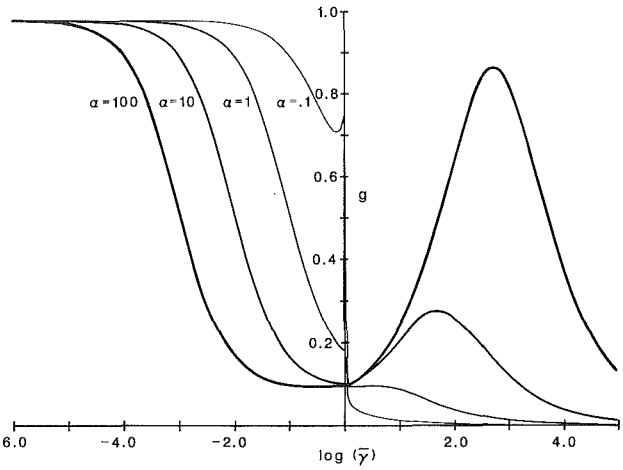


Fig. 3 g versus $\log(\bar{\gamma})$ for a standard linear solid with $\eta = 10$, $\lambda = 10.0$, and $\alpha = 0.1, 1.0, 10.0, 100.0$

In Fig. 1, k is plotted against $\gamma' \equiv \gamma/\sqrt{1+\eta}$ for $\alpha = 0.1$, 1, and 10. Clearly, k must vanish as γ' approaches 1.

In Fig. 2, $g(\alpha, \gamma, \eta, \epsilon)$ is plotted against $\log(\bar{\gamma})$, where $\bar{\gamma} = \gamma$ for $0 < \gamma < 1$ and $\bar{\gamma} = (\sqrt{\eta+1} - 1)/(\sqrt{\eta+1} - \gamma)$ for $1 < \gamma < \sqrt{\eta+1}$, for $\eta = 10$, $\epsilon = 0.01$ and $\alpha = 0.1, 1, 10, 100$. Thus the failure zone length is assumed to be constant. For many materials (such as rubber) a more realistic approximation is furnished by assuming a constant failure zone stress level, L_f. This is tantamount to holding λ constant and allowing ϵ to vary. From equation (8) one easily calculates ϵ as a function of λ to be

$$\epsilon = [\lambda^2(\beta+1) - \beta]^{-1}.$$

Consequently, one may regard g as a function of α, γ, η, and λ. Figure 3 is the analog of Fig. 2 for $\lambda = 10$, $\eta = 10$, and $\alpha = 0.1, 1, 10$. It should be noted that g vanishes as γ approaches $\sqrt{\eta+1}$ when λ is held constant. However, with ϵ constant, g tends to a nonzero finite limit as γ tends to $\sqrt{\eta+1}$, i.e., as V approaches the glassy shear wave speed. Indeed, from equation (33b) it is easily seen that

$$\lim_{\gamma \to \sqrt{\eta+1}} g(\alpha,\gamma,\eta,\epsilon) = \left(\frac{1-\epsilon}{1+\epsilon}\right) \frac{(\epsilon + \alpha\sqrt{\eta+1})(\epsilon/\eta)^{1/2}}{(\epsilon + \alpha(\eta+1)^{3/2})^{1/2}}.$$

The second example considered is a power-law material for which the shear modulus is assumed to have the form

$$\mu(t) = \mu_\infty(1 + (t/\tau)^{-n}), \quad 0 < n < 1$$
$$= \mu_\infty m(t/\tau).$$

For such material, the glassy wave speed, C, is infinite and $\tilde{m}(s)$ is given by

$$\tilde{m}(s) = 1 + \Gamma(1-n)s^n. \tag{34}$$

From equations (2) and (34) one sees that

$$q_o = \left(\frac{1}{V\tau}\right)\left[\frac{\gamma^2 - 1}{\Gamma(1-n)}\right]^{1/n}$$

and hence that

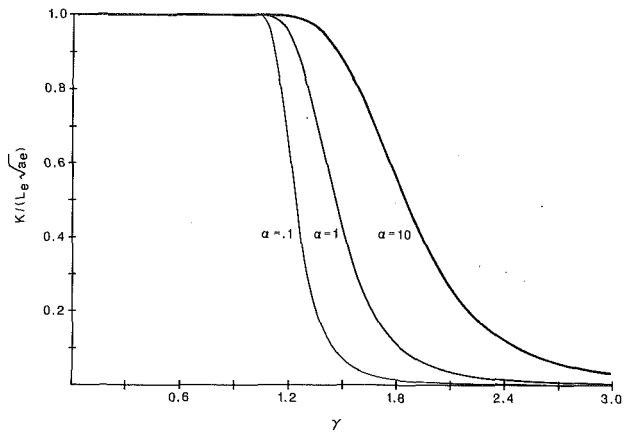


Fig. 4 k versus γ for a power-law material with $\eta = 0.3$ and $\alpha = 0.1, 1.0, 10.0$

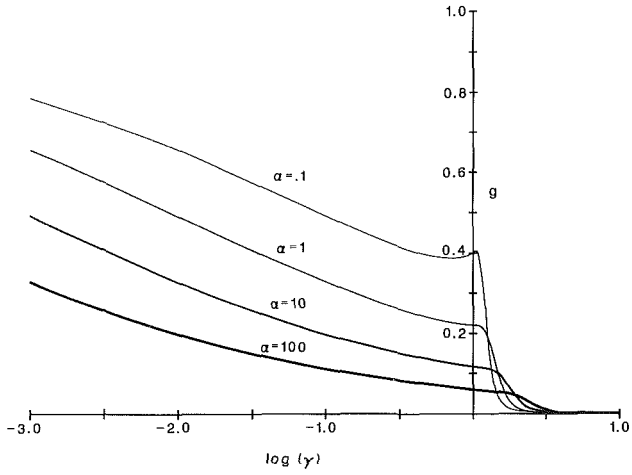


Fig. 5 g versus $\log(\gamma)$ for a power-law material with $\epsilon = 0.1, n = 0.3$ and $\alpha = 0.1, 1.0, 10.0, 100.0$

$$\beta = (\gamma^2 - 1)^{1/n}/(\alpha\gamma) \tag{35}$$

where, for convenience, α has been redefined by

$$\alpha = \Gamma(1-n)^{1/n} C^* \tau / a_e.$$

Combining equations (7), (8), (13), (34), and (35), it may be shown that

$$G = \frac{a_e L_e^2}{2\mu_\infty}\, g(\alpha,\gamma,n,\epsilon)$$

where

$$g(\alpha,\gamma,n,\epsilon) = \begin{cases} \left(\frac{1-\epsilon}{1+\epsilon}\right)\left[1-\gamma^2+\left(\frac{\alpha\gamma}{\epsilon}\right)^n\right]^{-1/2}\left[1+\left(\frac{\alpha\gamma}{\epsilon}\right)^n\right]^{-1/2} & 0<\gamma<1 \\ \left(\frac{1-\epsilon}{1+\epsilon}\right)\left(\frac{1-(\epsilon\beta)^2}{1-(\epsilon\beta)^n}\right)^{1/2}\left(\left(\frac{\alpha\gamma}{\epsilon}\right)^{2n}+\left(\frac{\alpha\gamma}{\epsilon}\right)^n\right)^{-1/2}/(1+\beta) & 1<\gamma \end{cases}$$

Figure 4 is a plot of the nondimensional SIF $k = (1+\beta)^{-1/2}$ against γ for $\alpha = 0.1$, 1, 10, and $n = 0.3$. Figure 5 shows g plotted against $\log(\gamma)$ for $\alpha = 0.1$, 1, 10, 100, $n = 0.3$, $\epsilon = 0.01$. Figure 6 has $\alpha = 1$, $\epsilon = 0.01$, $n = 0.1, 0.3, 0.5, 0.7$. The case λ constant is not exhibited here since it results in little change from the constant ϵ calculations.

Several comments on the numerical results should be made. It can be observed for a standard linear solid in Figs. 2 and 3 and may be shown analytically for general material that the slope of the curve g versus γ is discontinuous for $\gamma = 1$, i.e., as

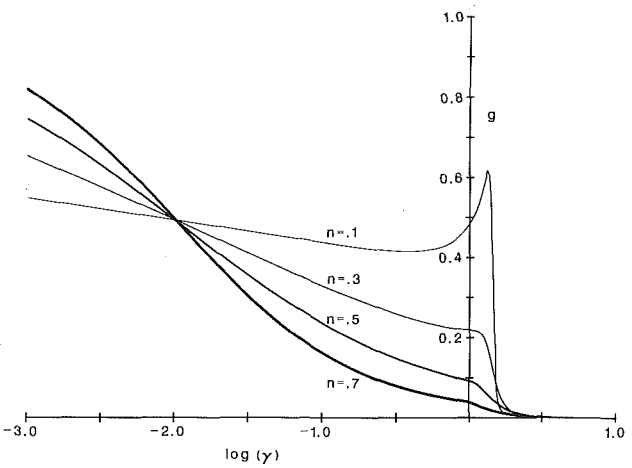


Fig. 6 g versus $\log(\gamma)$ for a power-law material with $\epsilon = 0.01$, $\alpha = 1.0$ and $n = 0.1, 0.3, 0.5, 0.7$

V passes through C^*. A more striking observation is the loss of monotonicity of g as a function of γ for certain ranges of the parameters. For example, in Figs. 2 and 3, it is seen that for $\alpha = 0.1$, g is monotone decreasing in γ, whereas for $\alpha = 1$, 10 g has a relative maximum on $1 < \gamma < \sqrt{1+\eta}$. As seen from Fig. 6, for power-law material, varying the exponent n also causes a transition from monotonicity to having a single relative maximum. The lack of monotonicity suggests that certain crack speeds are unstable. In particular, since G has the interpretation of work input to the crack tip, the γ-intervals on which g is increasing are those on which an increase in crack speed produces an increase in the work available to propagate the crack and hence should be considered unstable in steady-state. Evidently this lack of monotonicity is due to the combined inertial and viscoelastic effects considered here.

It is worth noting that these results illustrate that in contrast to elastic material, for viscoelastic material, there is no simple relationship between G and K. Indeed, though K is always a monotone decreasing function of V which vanishes at the glassy shear wave speed, G need not be monotone and need not vanish at $V = c$. Additional insight into the effect of material viscoelasticity can be gained by comparison with the ERR, G, for dynamic, elastic, steady-state crack propagation that arises as a simple limiting case of the analysis presented above. In the elastic limit, G is seen to be a monotone increasing function of crack speed that becomes infinite at the shear wave speed. This suggests that for elastic material, all crack speeds below the shear wave speed are unstable in steady-state which agrees with results contained in Freund (1986). Thus, the presence of material viscoelasticity tends to stabilize crack growth.

This paper closes with a brief illustration of the use of G as a fracture criterion. One must postulate the existence of a critical value of G, G_{cr}, which may be velocity dependent and which when exceeded results in crack acceleration. What is needed then is to invert equation (7) to find the γ versus L_e relationship for $G = G_{cr}$. Turning to the standard linear solid example, it is evident from equation (32) that the desired relation requires the inversion of $(L_e/\mu_\infty) = (G/a_e\mu_\infty)^{1/2}/(g(\alpha,$

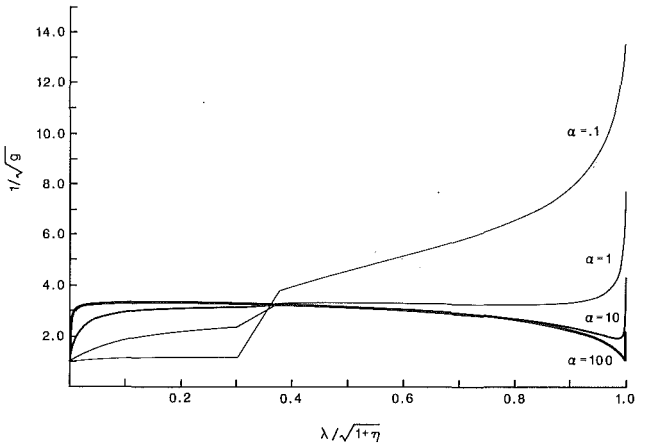


Fig. 7 $1/(g(\alpha, \gamma, \eta, \epsilon))^{1/2}$ **versus** $\log(\tilde{\gamma})$ **with** $\epsilon = 0.01$, $\eta = 10$, **and** $\alpha = 0.1, 1, 10, 100$

$\gamma, \eta, \epsilon))^{1/2}$ for γ as a function of the nondimensional load L_e/μ_∞. Figure 7 shows plots of $\log(\tilde{\gamma})$ versus $1/(g(\alpha, \gamma, \eta, \epsilon))^{1/2}$ for $\epsilon = 0.01$, $\eta = 10$, and $\alpha = 0.1, 1, 10, 100$, where $\tilde{\gamma} \equiv \gamma/\sqrt{\eta+1}$. The points on the ordinate axis represent, modulo multiplication by a constant, values of the nondimensional load L_e/γ_∞; the corresponding crack speed may then be found on the abscissa. It should be noticed that the speed intervals on which the curves are increasing (decreasing) correspond to stable (unstable) crack speeds.

Acknowledgment

This author wishes to acknowledge with much appreciation the encouragement and the many useful comments and suggestions offered by Prof. R. A. Schapery during the course of this research.

This research was supported by the Office of Naval Research under contract No. N00014-75-C-0325.

References

Atkinson, C., 1979, "A Note on Some Dynamic Crack Problems in Linear Viscoelasticity," *Arch. Mech. Stos.*, Vol. 31, pp. 829–849.

Atkinson, C., and Coleman, C. J., 1977, "On Some Steady-State Moving Boundary Problems in the Linear Theory of Viscoelasticity," *J. Inst. Maths Applics*, Vol. 20, pp. 85–106.

Atkinson, C., and Popelar, C., 1979, "Antiplane Dynamic Crack Propagation in a Viscoelastic Strip," *J. Mech. Phys. Solids*, Vol. 27, pp. 431–439.

Barenblatt, G. I., 1962, "The Mathematical Theory of Equilibrium Cracks in Brittle Fracture," in *Advances in Applied Mechanics*, Vol. VII, Academic Press, pp. 55–129.

Freund, L. B., 1986, "The Mechanics of Brittle Fracture," to appear in the *Proceedings of the Tenth U.S. National Congress of Applied Mechanics*.

Knauss, W. G., 1973, "On the Steady Propagation of a Crack in a Viscoelastic Sheet: Experiments and Analysis," in *Deformation and Fracture of High Polymers*, Kausch, H. H., and Jaffee, R., eds., Plenum Press, pp. 501–541.

Kostrov, B. V., and Nikitin, L. V., 1970, "Some General Problems of Mechanics of Brittle Fracture," *Arch. Mech. Stos.*, Vol. 22, pp. 749–776.

Popelar, C. H., and Atkinson, C., 1980, "Dynamic Crack Propagation in a Viscoelastic Strip," *J. Mech. Phys. Solids*, Vol. 28, pp. 79–93.

Schapery, R. A., 1975, "A Theory of Crack Initiation and Growth in Viscoelastic Media I. Theoretical Development," *Int. J. Frac.*, Vol. 11, pp. 141–159.

Sills, L. B., and Benveniste, Y., 1981, "Steady-State Propagation of a Mode III Interface Crack Between Dissimilar Viscoelastic Media," *Int. J. Eng. Sci.*, Vol. 19, pp. 1255–1268.

Walton, J. R., 1982, "On the Steady-State Propagation of an Anti-Plane Shear Crack in an Infinite Linearly Viscoelastic Body," *Quart. Appl. Math.*, pp. 37–52.

Walton, J. R., 1985, "The Dynamic, Steady-State Propagation of an Anti-Plane Shear Crack in a General Linearly Viscoelastic Layer," JOURNAL OF APPLIED MECHANICS, Vol. 54, No. 4, pp. 853–856.

Willis, J. R., 1967, "Crack Propagation in Viscoelastic Media," *J. Mech. Phys. Solids*, Vol. 15, pp. 229–240.

Periodic Array of Cracks in a Half-Plane Subjected to Arbitrary Loading

H. F. Nied
General Electric Company,
Corporate Research and Development,
Schenectady, NY 12301
Mem. ASME

The plane elastic problem for a periodic array of cracks in a half-plane subjected to equal, but otherwise arbitrary normal crack surface tractions is examined. The mixed boundary value problem, which is formulated directly in terms of the crack surface displacements, results in a hypersingular integral equation in which the unknown function is the crack opening displacement. Based on the theory of finite part integrals, a least squares numerical algorithm is employed to efficiently solve the singular integral equation. Numerical results include crack opening displacements, stress intensity factors, and Green's functions for the entire range of possible periodic crack spacing.

1 Introduction

In this paper the elasticity problem for an infinite array of periodic cracks in a half-plane is examined to determine the stress intensity factors and the crack opening displacements as a function of the crack spacing. Of particular interest for fracture mechanics applications are the calculated values of the stress intensity factors for a row of edge cracks subjected to arbitrary surface tractions normal to the crack surface. It is well-known that with decreasing edge crack spacing, crack interactions reduce the magnitude of the stress intensity factors at the crack tips to a level well below the stress intensity factors associated with a single crack of the same length. Therefore, generalized solutions for such crack configurations, i.e., solutions which are applicable for arbitrary loading normal to the crack surface, are more useful for determining the conditions for sudden fracture in brittle materials that are known to have a large number of small, equal length, surfaces cracks, than similar solutions developed for a single crack. Typical examples of nonuniform stress fields which would be of particular interest include transient thermal stresses and wedge loading of the crack surfaces.

The solution for the problem of a half-plane with an infinite row of periodic edge cracks subjected to uniform axial stress has been given by Benthem and Koiter (1973) using an asymptotic approximation and Bowie (1973) who used conformal mapping. More recently, solutions for interacting arrays of parallel edge cracks subjected to specific thermal stress conditions have been developed by Nemat-Nasser et al. (1978) by formulating the problem in terms of a singular integral equation with the usual Cauchy type singularity.

In this paper we will consider the elasticity problem depicted in Fig. 1; an infinite array of stacked internal cracks, uniformly spaced apart at a distance h. The special case of an array of edge cracks (i.e., $a = 0$ in Fig. 1) will also be considered, with special attention devoted towards developing a Green's function which can be used to easily calculate the stress intensity factors which arise when the system of edge cracks is subjected to arbitrary crack surface tractions normal to the crack surface.

Instead of taking the usual approach of formulating the problem in terms of a singular integral equation of the Cauchy type (a consequence of the somewhat artificial step of specifying the unknown function in terms of the derivative of the crack surface displacements), an alternative approach will be followed in which the unknown function is simply taken to be the crack surface displacement. The hypersingular integral equation which results from such a formulation contains a singularity of order $1/x^2$. Integral equations with such strong singularities have only recently become amenable to direct numerical solution (as opposed to reducing the equation to an equivalent Cauchy singular integral equation). Kaya and Erdogan (1984, 1987), have developed a numerical approach which utilizes the concept of singular integrals interpreted in the finite-part sense, a concept which was introduced by Hadamard (1923). In the present study, it was found that the finite-part interpretation of the integrals with strong singularities, which arise in this particular crack problem, leads to a very accurate and efficient method for numerical solution.

2 Formulation of the Problem

In addition to satisfying the two-dimensional equations of equilibrium and the elastic constitutive relationships, the solution to the problem depicted in Fig. 1 must also satisfy the following boundary and symmetry conditions

$$\sigma_{xx}(0,y) = 0 \tag{1}$$

$$\tau_{xy}(0,y) = 0 \tag{2}$$

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March 19, 1986; final revision, February 6, 1987.

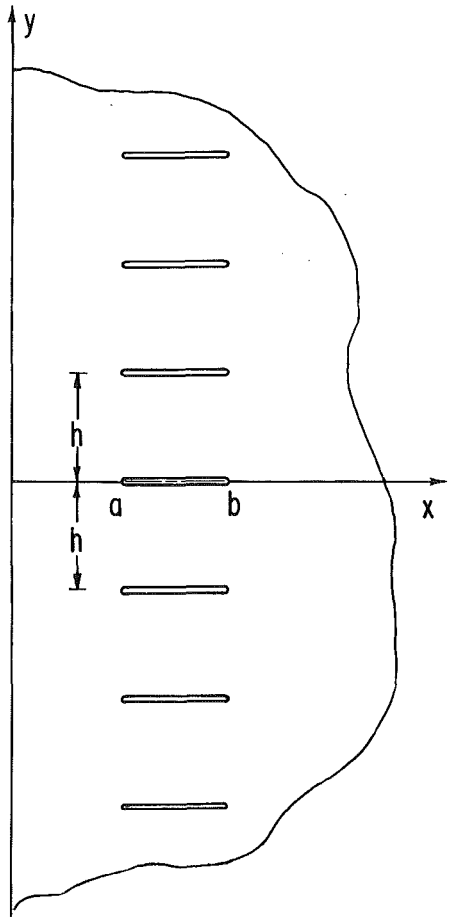


Fig. 1 Array of periodically spaced cracks in a half-plane

$$v(x,nh)=0,\ (x<a,x>b),\ n=-\infty,\ \ldots,\infty \tag{3}$$

$$\sigma_{yy}(x,nh)=-p(x),\ (a<x<b),\ n=-\infty,\ \ldots,\infty \tag{4}$$

$$\tau_{xy}(x,nh)=0,\ n=-\infty,\ \ldots,\infty \tag{5}$$

where it's assumed that all the crack surfaces are subjected to equal and symmetric loading specified by $p(x)$.

The crack problem depicted by Fig. 1 can be conveniently formulated by applying integral transform techniques to the two-dimensional equations of elasticity. The resulting algebra is somewhat simplified, if first we develop the stress field for a single crack in the half-plane and then use superposition to obtain the expressions for the multiple crack problem. Furthermore, the derivation of the single crack expressions are facilitated by using superposition to combine expressions for the uncracked half-plane and a cracked infinite space. For example, the x and y components of displacement for the single crack problem can be written as

$$u(x,y)=u^1(x,y)+u^2(x,y) \tag{6}$$

$$v(x,y)=v^1(x,y)+v^2(x,y) \tag{7}$$

respectively, where the superscript 1 refers to the infinite space with a single crack and superscript 2 indicates the contribution due to the half-space. In the same manner the stresses for a half-space with a single crack can also be expressed by

$$\sigma_{ij}(x,y)=\sigma^1_{ij}(x,y)+\sigma^2_{ij}(x,y),\ (i,j{:}x,y) \tag{8}$$

Integral expressions are developed for the variables in equations (6)–(8) by direct substitution of the applicable Fourier transform into the two-dimensional equations of equilibrium and Hooke's law. After applying the symmetry condition τ_{xy} $(x, 0) = 0$ and the shear condition for a single crack τ_{xy} $(0, y)$ $= 0$, the superposed equations for displacement (6)–(7) can be shown to have the form

$$u(x,y)=\frac{2}{\pi}\int_0^\infty\left(\frac{1-\kappa}{2\alpha}+y\right)A(\alpha)e^{-\alpha y}\sin\alpha x\,d\alpha$$

$$+\frac{2}{\pi}\int_0^\infty\left(\frac{1+\kappa}{2\beta}+x\right)B(\beta)e^{-\beta x}\cos\beta y\,d\beta \tag{9}$$

$$v(x,y)=\frac{2}{\pi}\int_0^\infty\left(\frac{1+\kappa}{2\alpha}+y\right)A(\alpha)e^{-\alpha y}\cos\alpha x\,d\alpha$$

$$+\frac{2}{\pi}\int_0^\infty\left(\frac{1-\kappa}{2\beta}+x\right)B(\beta)e^{-\beta x}\sin\beta y\,d\beta \tag{10}$$

where

$$\kappa=3-4\nu,\ \text{for plane strain} \tag{11}$$

$$\kappa=\frac{3-\nu}{1+\nu},\ \text{for plane stress} \tag{12}$$

and $A(\alpha)$ and $B(\beta)$ are unknown functions which must be determined from the boundary conditions. With the displacements given by equations (9)–(10) it is not difficult to show that the components of stress are given by

$$\frac{1}{2\mu}\sigma_{xx}=\frac{2}{\pi}\int_0^\infty(-1+\alpha y)A(\alpha)e^{-\alpha y}\cos\alpha x\,d\alpha$$

$$-\frac{2}{\pi}\int_0^\infty(1+\beta x)B(\beta)e^{-\beta x}\cos\beta y\,d\beta \tag{13}$$

$$\frac{1}{2\mu}\sigma_{yy}=\frac{-2}{\pi}\int_0^\infty(1+\alpha y)A(\alpha)e^{-\alpha y}\cos\alpha x\,d\alpha$$

$$+\frac{2}{\pi}\int_0^\infty(-1+\beta x)B(\beta)e^{-\beta x}\cos\beta y\,d\beta \tag{14}$$

$$\frac{1}{2\mu}\tau_{xy}=\frac{-2}{\pi}\int_0^\infty\alpha y\,A(\alpha)e^{-\alpha y}\sin\alpha x\,d\alpha$$

$$-\frac{2}{\pi}\int_0^\infty\beta x\,B(\beta)e^{-\beta x}\sin\beta y\,d\beta \tag{15}$$

where μ is the elastic shear modulus.

If we define a displacement function in the following manner:

$$V(x)=v(x,0^+)-v(x,0^-)=2v(x,0),\ (a<x<b) \tag{16}$$

$$V(x)=0,\ (x<a,x>b) \tag{17}$$

and apply displacement boundary condition (3), stress boundary condition (1), and invert the resulting expressions in equations (10) and (13); for the case of a single crack (i.e., $n =$ 0) we can obtain the unknowns A and B. Thus, it is straightforward to show that the y component of stress, in terms of the unknown crack opening displacement $V(x)$, for a single crack is given by

$$\frac{\pi\sigma_{yy}}{4\mu}=-\int_a^b V(t)dt\int_0^\infty(1+\alpha y)\frac{\alpha}{1+\kappa}\cos\alpha t\cos\alpha x\,e^{-\alpha y}d\alpha$$

$$+\int_a^b V(t)dt\int_0^\infty(-1+\beta x)\frac{\beta}{1+\kappa}(\beta t-1)e^{-\beta(t+x)}\cos\beta y\,d\beta \tag{18}$$

After substitution of the trigonometric identity

$$\cos\alpha t\cos\alpha x=\frac{1}{2}\left[\cos\alpha(t-x)+\cos\alpha(t+x)\right] \tag{19}$$

and evaluation of the definite integrals in equation (18), the perturbation in the y component of the stress field due to a single crack located at $y = 0$ can be expressed in terms of the crack opening displacement $V(x)$ as

$$\frac{\kappa+1}{2\mu}\,\pi\,\sigma_{yy}(x,y)=\int_a^b V(t)G(t,x,y)\,dt \tag{20}$$

where $G(t, x, y)$ is expressed by

$$G(t,x,y)=\frac{(t-x)^2-y^2}{[(t-x)^2+y^2]^2}+\frac{(t+x)^2-y^2}{[(t+x)^2+y^2]^2}$$

$$+\frac{2y^2[3(t-x)^2-y^2]}{[(t-x)^2+y^2]^3}+\frac{2y^2[3(t+x)^2-y^2]}{[(t+x)^2+y^2]^3}$$

$$+\frac{2[(t+x)^2-y^2]}{[(t+x)^2+y^2]^2}+\frac{-4(t+x)^2[(t+x)^2-3y^2]}{[(t+x)^2+y^2]^3}$$

$$+\frac{12tx[(t+x)^4-6y^2(t+x)^2+y^4]}{[(t+x)^2+y^2]^4} \tag{21}$$

The complete expression for an infinite array of cracks can be obtained from equations (20) and (21) by invoking superposition. That is, the perturbation on the stress field along the line $y = 0$ due to an infinite row of cracks, is simply the infinite sum of contributions determined from equation (20) for values of y spaced apart at an interval h. Thus, by setting $y = nh$ in equation (21), summing on n from $-\infty$ to -1 and from 1 to ∞, and adding these terms to the limiting expression obtained from equation (20) as $y \to 0$, we obtain the correct integral equation for an array of cracks. The integral equation for an infinite array of cracks differs from the single crack equation in that it has an additional Fredholm kernel as a result of the summation contributions. Due to symmetry, the summation of G from $-\infty, \ldots, -1, 1, \ldots, \infty$ can be reexpressed as twice the sum of G for $n = 1, 2, \ldots, \infty$. For the case of periodic spacing these infinite sums can be obtained in closed form. Two identities which aid in the reduction of these sums are (Gradshteyn and Rhyzhik, 1965)

$$\sum_{n=1}^{\infty}\frac{1}{x^2+n^2}=\frac{1}{2x^2}\left[\pi x\coth\pi x-1\right] \tag{22}$$

$$\sum_{n=1}^{\infty}\frac{x^2-n^2}{(x^2+n^2)^2}=\frac{\pi^2}{2}\operatorname{csch}^2\pi x-\frac{1}{2x^2} \tag{23}$$

The closed form expressions which are needed to evaluate the sums arising out of equations (20) and (21) can be obtained by successively differentiating equations (22) and (23) and combining intermediate expressions.

After making the substitutions,

$$y=nh \tag{24}$$

$$\alpha=\frac{(t-x)}{h} \tag{25}$$

$$\beta=\frac{(t+x)}{h} \tag{26}$$

in equations (20) and (21), evaluating all sums from 1 to ∞ and regrouping terms, the net perturbation along the line $y = 0$ due to the influence of all cracks extending from $-\infty$ to $+\infty$ is

$$\frac{\kappa+1}{2\mu}\,\pi\,\sigma_{yy}=\int_a^b V(t)K_2(t,x,h)\,dt \tag{27}$$

where

$$K_2(t,x,h)=\frac{1}{h^2}\left\{-\frac{1}{\alpha^2}+\frac{1}{\beta^2}+3\pi^2\operatorname{csch}^2\pi\alpha+5\pi^2\operatorname{csch}^2\pi\beta\right.$$

$$-2\pi^3\alpha\operatorname{csch}^2\pi\alpha\coth\pi\alpha-6\pi^3\beta\operatorname{csch}^2\pi\beta\coth\pi\beta$$

$$\left.+(\beta^2-\alpha^2)\left[2\pi^4\operatorname{csch}^2\pi\beta\coth^2\pi\beta+\pi^4\operatorname{csch}^4\pi\beta-\frac{3}{\beta^4}\right]\right\} \tag{28}$$

Evaluation of expression (20) $\lim_{y\to 0}$ yields

$$\int_a^b\left[\frac{1}{(t-x)^2}+\frac{-1}{(t+x)^2}+\frac{12x}{(t+x)^3}+\frac{-12x^2}{(t+x)^4}\right]V(t)\,dt$$

$$=\frac{\kappa+1}{2\mu}\,\pi\,\sigma_{yy} \tag{29}$$

After combining equation (29) with (27), the integral equation for the displacement function $V(x)$ can be expressed by

$$\fint_a^b\frac{V(t)\,dt}{(t-x)^2}+\int_a^b V(t)\left[K_1(t,x)\right.$$

$$\left.+K_2(t,x,h)\right]dt=\frac{-\pi(\kappa+1)}{2\mu}\,p(x) \tag{30}$$

where $p(x)$ is the specified crack surface tractions, K_1 is

$$K_1(t,x)=\frac{-1}{(t+x)^2}+\frac{12x}{(t+x)^3}+\frac{-12x^2}{(t+x)^4} \tag{31}$$

and K_2 is given by expression (28). The symbol ($\fint$) denotes that the integral is to be interpreted as a finite-part integral in the sense of Hadamard (Kaya and Erdogan, 1987).

3 Solution of the Integral Equation

For the case of internal cracks, the integral equation in (30) can be normalized between -1 and 1 with the substitutions

$$t=\frac{(b-a)}{2}\,r+\frac{(b+a)}{2} \tag{32}$$

$$x=\frac{(b-a)}{2}\,s+\frac{(b+a)}{2} \tag{33}$$

$$V(t)=\frac{b-a}{2}\,\tilde{V}(r) \tag{34}$$

The normalized integral equation in this case is

$$\fint_{-1}^{1}\frac{\tilde{V}(r)}{(r-s)^2}\,dr+\int_{-1}^{1}\tilde{V}(r)L(r,s)\,dr=\tilde{p}(s) \tag{35}$$

with

$$L(r,s)=\left(\frac{b-a}{2}\right)^2(K_1(t,x)+K_2(t,x,h)) \tag{36}$$

$$\tilde{p}(s)=-\pi\,\frac{\kappa+1}{2\mu}\,p(x) \tag{37}$$

The theory of finite-part integrals will not be explored in detail in this study. It is sufficient to simply state that a Cauchy principal value integral can be written as the sum of two finite-part integrals and that direct differentiation of a Cauchy integral gives a finite-part integral (Kaya, 1984). For example,

$$\frac{d}{dx}\fint_a^b\frac{f(t)}{t-x}\,dt=\fint_a^b\frac{f(t)}{(t-x)^2}\,dt \tag{38}$$

Making use of relation (38) it is possible to determine the finite-part integral for a wide range of kernels. A finite-part integral of particular interest in this study is given by Kaya and Erdogan (1987)

$$\fint_{-1}^{1}\frac{\sqrt{1-t^2}\,U_n(t)}{(t-x)^2}\,dt=-\pi(n+1)U_n(t) \tag{39}$$

where $U_n(t)$ is the Chebyshev polynomial of the second kind. If we let

$$\bar{V}(r) = F(r)(1-r^2)^{1/2} \tag{40}$$

and expand $F(r)$ in terms of a Chebyshev polynomial series of the second kind, it is possible to evaluate the resulting finite-part integrals with the aid of equation (39). Thus,

$$F(r) = \sum_{i=0}^{N} a_i U_i(r) \tag{41}$$

and

$$\bar{V}(r) = \left[\sum_{i=0}^{N} a_i U_i(r)\right](1-r^2)^{1/2} \tag{42}$$

Of course the method is not restricted to expansions only in Chebyshev polynomials and, in particular, power series expansions seem to be especially convenient for many applications (Kaya and Erdogan, 1987).

In equation (41) the a_i's are unknown coefficients, which once determined specify the crack opening displacement function $\bar{V}(r)$. Direct substitution of equation (42) into the integral equation (35) and evaluation of the finite-part integral given by equation (39) results in

$$\sum_{i=0}^{N} a_i\left[-\pi(i+1)U_i(s) + h_i(s)\right] = \bar{p}(s) \tag{43}$$

where

$$h_i(s) = \int_{-1}^{+1} U_i(r)L(r,s)(1-r^2)^{1/2}dr \tag{44}$$

Equation (43) can be solved by simple collocation. That is, we can construct N linearly independent equations for the N unknown a_i's by evaluating equation (43) at N station points s_j. The resulting system of equations for the solution of the a_i's is given by

$$\sum_{i=0}^{N} a_i\left[-\pi(i+1)U_i(s_j) + h_i(s_j)\right] = \bar{p}(s_j), \quad j=0,1,\ldots,N \tag{45}$$

where s_j's, for example, can be determined from the roots of the Chebyshev polynomial of the first kind

$$T_{n+1}(s_j) = 0,\ s_j = \cos\left(\frac{2j+1}{N+1}\frac{\pi}{2}\right),\ j=0,1,\ldots,N \tag{46}$$

The stress intensity factors are defined by

$$k_1(a) = \left(\frac{2\mu}{\kappa+1}\right)\lim_{t\to a}\frac{V(t)}{\sqrt{2(t-a)}} \tag{47}$$

$$k_1(b) = \left(\frac{2\mu}{\kappa+1}\right)\lim_{t\to b}\frac{V(t)}{\sqrt{2(b-t)}} \tag{48}$$

Thus, once $F(r)$ (equation (41)) is known, the stress intensity factors at either crack tip may be expressed as

$$k_1(a) = \left(\frac{2\mu}{\kappa+1}\right)\sqrt{\frac{b-a}{2}}\,F(-1) \tag{49}$$

$$k_1(b) = \left(\frac{2\mu}{\kappa+1}\right)\sqrt{\frac{b-a}{2}}\,F(1) \tag{50}$$

Since $F(-1)$ and $F(1)$ are given by

$$F(-1) = \lim_{x\to -1}\sum_{i=0}^{N} a_i U_i(x),\quad F(1) = \lim_{x\to 1}\sum_{i=0}^{N} a_i U_i(x) \tag{51}$$

respectively, the values of $F(r)$ at either endpoint in terms of the a_i's is

$$F(-1) = \sum_{i=0}^{N} a_i(i+1)(-1)^i \tag{52}$$

$$F(1) = \sum_{i=0}^{N} a_i(i+1) \tag{53}$$

For the case of edge cracks, i.e., when $a = 0$, the normalized integral equation has the form

$$\int_0^1 \frac{\bar{V}(r)}{(r-s)^2}\,dr + \int_0^1 \bar{V}(r)L(r,s)dr = \bar{p}(s) \tag{54}$$

with

$$r = \frac{t}{b},\ s = \frac{x}{b},\ \bar{V}(r) = \frac{V(t)}{b},$$

$$L(r,s) = b^2(K_1(t,x) + K_2(t,x,h)),$$

$$\bar{p}(s) = -\pi\frac{\kappa+1}{2\mu}p(x) \tag{55}$$

As in the case of internal cracks, we can express $\bar{V}(r)$ by

$$\bar{V}(r) = \left[\sum_{i=0}^{N} a_i U_i(r)\right](1-r^2)^{1/2} \tag{56}$$

Substitution of equations (56) into (54) leads to an expression similar to equation (45) for the evaluation of the unknown coefficients a_i.

$$\sum_{i=0}^{N} a_i\left\{\int_0^1 \frac{U_i(r)\sqrt{1-r^2}}{(r-s_j)^2}\,dr + \int_0^1 U_i(r)L(r,s_j)\sqrt{1-r^2}\,dr\right\} = \bar{p}(s_j) \tag{57}$$

The difference between this expression and equation (45) is that a closed form expression for the finite-part integral in equation (57) is not known. Kaya (1984) has shown that we can take advantage of the identity (39) by rewriting equation (57) as

$$\sum_{i=0}^{N} a_i\left\{\int_{-1}^1 \frac{U_i(r)\sqrt{1-r^2}}{(r-s_j)^2}\,dr - \int_0^1 \frac{U_i(-r)\sqrt{1-r^2}}{(r+s_j)^2}\,dr + \int_0^1 U_i(r)L(r,s_j)\sqrt{1-r^2}\,dr\right\} = \bar{p}(s_j) \tag{58}$$

and thus the algebraic system of equations for solution of the a_i's if collocation is used, becomes

$$\sum_{i=0}^{N} a_i\left[-\pi(i+1)U_i(s_j) + h_i(s_j)\right] = \bar{p}(s_j), \quad 0<s_j<1,\ j=0,1,2,\ldots,N \tag{59}$$

where

$$h_i(s_j) = \int_0^1\left[-\frac{U_i(-r)}{(r+s_j)^2} + U_i(r)L(r,s_j)\right]\sqrt{1-r^2}\,dr \tag{60}$$

Once the unknown a_i's are known for the edge crack problem, the stress intensity factor is determined from equation (48) after substitution from equations (55) and (56). Thus, for edge cracks the stress intensity factor is given by

$$k_1(b) = \left(\frac{2\mu}{\kappa+1}\right)\sqrt{b}\sum_{i=0}^{N} a_i(i+1) \tag{61}$$

4 Numerical Solution and Results

When collocation is used for the solution of equations (45) or (59), the number of terms N, in the Chebyshev polynomial expansion, is equal to the number of collocation points s_j. It is not difficult to show that for large values of N, the system of equations defined by either (45) or (59) becomes badly ill-conditioned. For example, the determinant of a typical system of equations from equation (59), for a 10th order Chebyshev polynomial expansion ($N = 11$), is of $0(10^6)$. For a Chebyshev polynomial expansion of order 24 ($N = 25$), the determinant is of $0(10^{-22})$. Since, in this study it was desired that stress intensity factors be calculated for arbitrary loading, it was necessary to use a numerical algorithm which incorporated information at a sufficiently large number of load points s_j on the crack surface. Thus, to keep the order of the Chebyshev polynomial expansion relatively low, but still incorporate information from a greater number of load points than unknowns, the method of least squares was employed. Following the usual procedure for deriving the set of Normal Equations used in the least squares algorithm, the system of $N \times N$ algebraic equations which minimizes the square of the error for M sampling points is given by

$$\sum_{i=0}^{N} a_i \left(\sum_{j=0}^{M} A_{ij} A_{kj} \right) = \sum_{j=0}^{M} \tilde{p}_j A_{kj}, \qquad k = 0,1,2, \ldots, N \tag{62}$$

where

$$A_{ij} = -\pi (i+1) U_i(s_j) + h_i(s_j) \tag{63}$$

For the case of completely embedded cracks, the integration of the bounded kernel implied by equation (44) was handled numerically using Gauss-Chebyshev quadrature. The numerical solution of the $N \times N$ system of algebraic equations given by equations (62) results in a very accurate determination of the stress intensity factors as well as crack opening displacement (COD). However, for the case of edge cracks, the integral given by equation (60) requires greater effort to evaluate accurately. Rewriting equation (60) and temporarily excluding the portion of the kernel which remains bounded for all values of r and s, we obtain

$$I = \int_0^1 \left[\frac{(-1)^{i+1}}{(r+s)^2} + \frac{-1}{(r+s)^2} + \frac{12s}{(r+s)^3} + \frac{-12s^2}{(r+s)^4} \right] U_i(r)(1-r^2)^{1/2} dr \tag{64}$$

which exhibits a singular behavior as both s and r simultaneously go to 0. For $r = 0$, evaluation of the integral in equation (64) yields a singular behavior of the form

$$I_s = \left[\frac{[(-1)^{i+1} - 1]}{s(s+1)} + \frac{6(1+2s)}{s(1+s)^2} + \frac{-4(1+3s+3s^2)}{s(1+s)^3} \right] \sin\left[(i+1)\frac{\pi}{2} \right] \tag{65}$$

and thus we can improve convergence in the numerical evaluation of equation (64) by adding and subtracting the singular behavior exhibited in equation (65). That is,

$$I = I_s + \int_0^1 \left[\frac{(-1)^{i+1}}{(r+s)^2} + \frac{-1}{(r+s)^2} + \frac{12s}{(r+s)^3} + \frac{-12s^2}{(r+s)^4} \right] * \left[U_i(r)\sqrt{1-r^2} - \sin\left[(i+1)\frac{\pi}{2} \right] \right] dr \tag{66}$$

This rearrangement of terms satisfactorily improves the rate of convergence during numerical evaluation of equation (64),

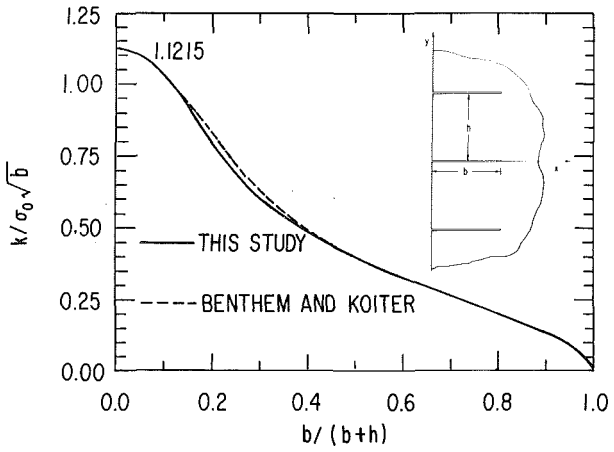


Fig. 2 Stress intensity factors for a row of edge cracks subjected to uniform applied stress σ_0. Comparison with Bentham and Koiter.

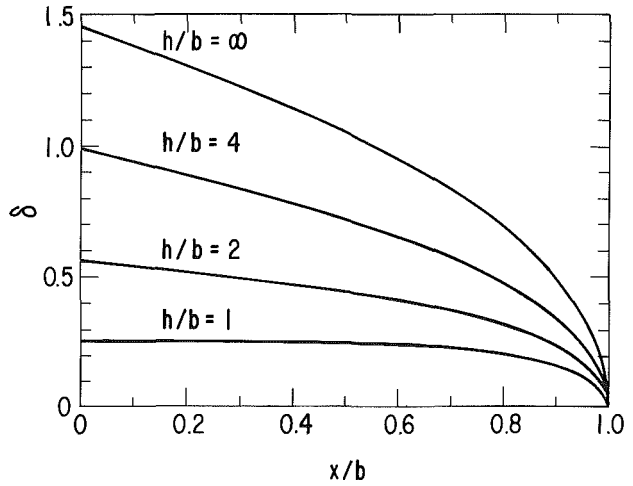


Fig. 3 Crack opening displacement for a row of edge cracks subjected to uniform applied stress σ_0.

$\delta = \left(\frac{2\mu}{1+\kappa} \right) \frac{V(s)}{\sigma_0 b}$.

with the integral in equation (66) calculated using Gauss-Chebyshev quadrature.

Figure 2 gives the stress intensity factors for an array of edge cracks subjected to uniform loading σ_0. The normalization parameter S is given by

$$S = \frac{b}{b+h} \tag{67}$$

In this figure the results for an array of edge cracks is compared with Benthem and Koiter's published results (Benthem and Koiter, 1973). It can be seen that over the entire range of possible crack spacing, the two results are almost identical. In generating the stress intensity factors it was found that an 11-term Chebyshev series expansion was sufficient for highly accurate results. For example, the stress intensity factor for an edge crack with $S = 0$, i.e., $h = \infty$, was determined to be $k/\sigma_0\sqrt{b} = 1.121522$. For comparison, the exact solution can be determined from the numerical evaluation of a closed form expression given by Koiter (1965) and is known to be $k/\sigma_0\sqrt{b} = 1.12152226$. It is interesting to note that when $S = 0.5$, i.e., $h/b = 1$, the stress intensity factor $k/\sigma_0\sqrt{b} = 0.398662$; a value which is only 35.5 percent of the maximum stress intensity factor for a single isolated edge crack. In Fig. 3 the crack opening displacement (δ), calculated from the mouth of the crack to the crack tip, is displayed for interacting edge cracks

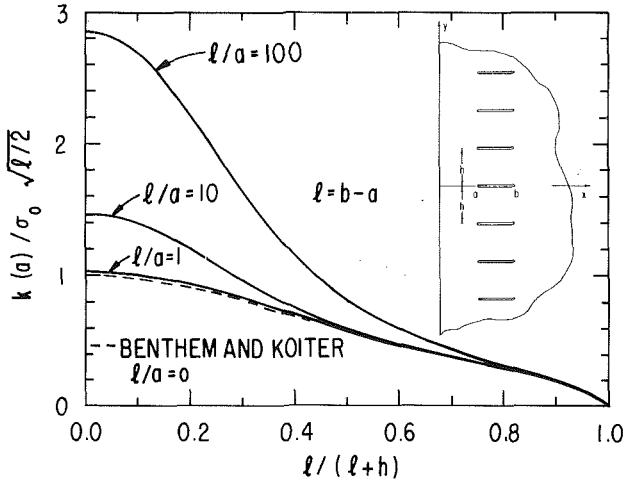


Fig. 4 Stress intensity factors at crack tip a, for an array of embedded cracks subjected to uniform applied stress σ_0 ($\ell = b - a$.)

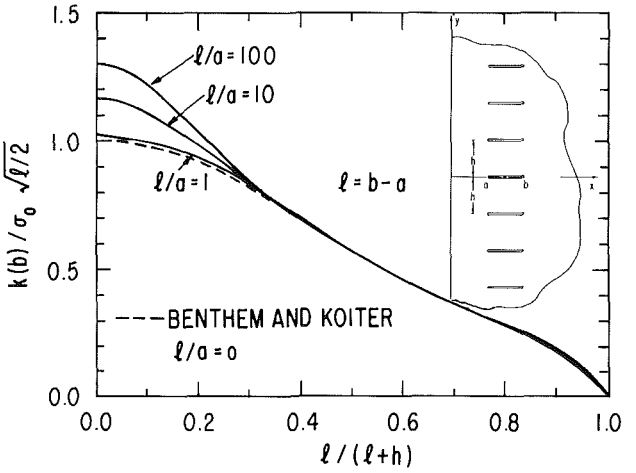


Fig 5 Stress intensity factors at crack tip b, for an array of embedded cracks subjected to uniform applied stress σ_0. ($\ell = b - a$.)

with different h/b ratios. Again, it can be seen that the overall effect of nearby surface cracks interactions is to greatly reduce the COD along the entire crack length.

Figures 4 and 5 contain plots of stress intensity factors calculated for cracks subjected to uniform remote stress σ_0 and located below the surface of the half-space. Note that for the embedded crack the normalized crack spacing parameter S is given by

$$S = \frac{(b-a)}{(b-a)+h} \tag{68}$$

In Fig. 4, the stress intensity factors are shown for the crack tip located closest to the free surface of the plane as a function of S with $(b - a)/a = 1$, 10, 100. Figure 5 contains a similar plot for the crack tip at the point of deepest penetration into the half-space. In both these figures the dashed line represents an approximate solution given by Benthem and Koiter (1973) for the case of a "stack" of cracks in an infinite space, i.e., the limiting case when $a = \infty$ for finite crack lengths. It is remarkable that this limiting case is so closely approximated when $(b - a)/a = 1$. This means that the half-space free surface has little effect on the stress intensity factors when the unbroken ligament length is as small as one crack length. In comparing Fig. 5 to Fig. 4 it can be seen that the crack tip located at the point of deepest penetration in the half-space is relatively insensitive to the crack length/ligament length ratio for $S >$

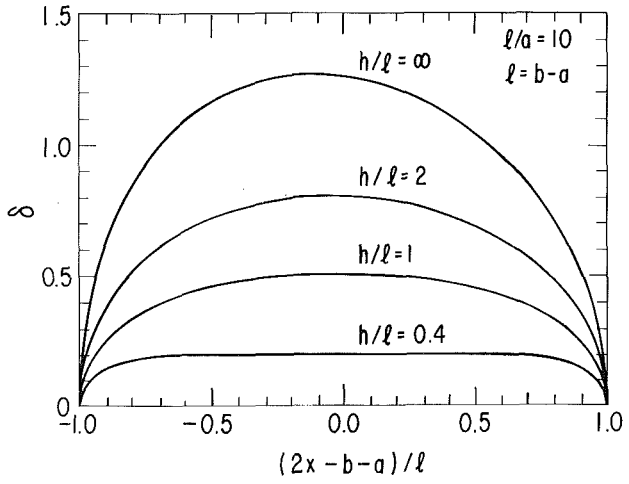


Fig. 6 Crack opening displacement for a row of embedded cracks subjected to uniform applied stress σ_0.

$\delta = \left(\frac{2\mu}{1+\mu}\right)\frac{V(s)}{\sigma_0(\ell/2)}$.

0.25. However, at the crack tip nearest to the free surface (Fig. 4), the effect of the crack length to the unbroken ligament length ratio $(b - a)/a$ is more pronounced, with larger differences between the stress intensity factors for different $(b - a)/a$ ratios seen over a much wider range of crack array spacing. Figure 6 contains plots of the COD δ for embedded cracks with $(b - a)/a = 10$ and various values of crack spacing. For the case of the isolated embedded crack ($h/(b - a) = \infty$), the COD is noticeably nonsymmetric for this value of $(b - a)/a$. With decreasing distance between interacting cracks ($h/(b - a)$), the COD becomes smaller in magnitude, has a "flatter" slope along the crack axis, and appears increasingly symmetric.

It is possible to generate the results for concentrated wedge force loading applied on the crack surface at $s = s_j$ by setting the right-hand side of the integral equation (54) to

$$\tilde{p}(s) = \begin{cases} \dfrac{2P}{s_{j+1} - s_{j-1}} & s = s_j \\ 0 & s = s_i,\ i \neq j \end{cases} \tag{69}$$

where P represents the concentrated wedge force. By sequentially applying unit wedge force loading to the entire crack surface, it is possible to generate a numerical Green's function for the stress intensity factors. In a strict sense, the quantities calculated by using equation (69) represent the response to a crack surface traction distributed on a small area around $s = s_j$ rather than a concentrated wedge force acting at s_j. In any case the numerical approximation of the concentrated force loading rapidly converges to a unique solution for the Green's function with an increasing number of abscissa load points s_j. Figure 7 is a plot of the numerically generated Green's functions for the stress intensity factors at the tip of an edge crack. The function $G(s)$ is given by

$$G(s) = \frac{k}{P}\sqrt{b}\,\sqrt{1-s} \tag{70}$$

and for different h/b ratios represents the stress intensity factor resulting from concentrated force loading at a given value of s ($s = x/b$). The numerical evaluation of equation (64) becomes increasingly difficult for very small values of s, i.e., when the unit loading is applied close to the crack mouth, due to the singularity which arises when both r and s simultaneously go to 0. Thus, the limiting values given in Fig. 7 for $G(0)$ were obtained by extrapolating a least-squares curve fit

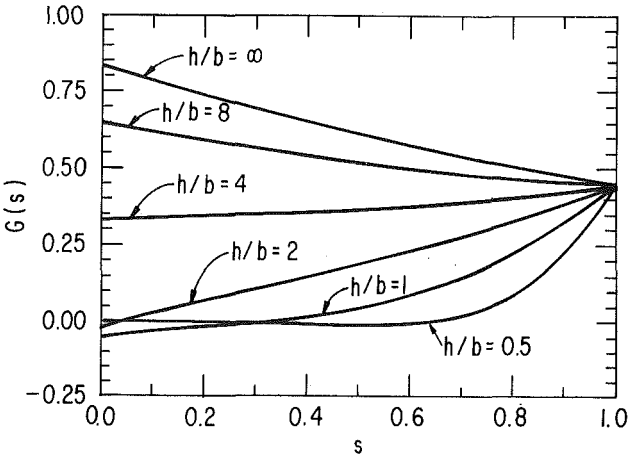


Fig. 7 Edge crack Green's functions for stress intensity factors.

$G(s) = \frac{k}{P}\sqrt{b}\sqrt{1-s}$, $s = x/b$.

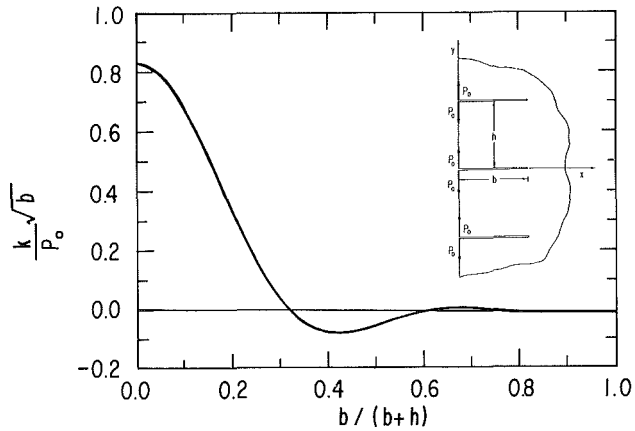


Fig. 8 Stress intensity factors due to concentrated force loading located at crack mouth

polynomial derived from the numerical values of G over the entire range x/b for a given h/b ratio. It is interesting to note in Fig. 7 that for certain small values of h/b the stress intensity factors due to concentrated force loading close to the crack mouth have become negative. This indicates that for these particular values of h/b, wedge force loading at the crack mouth results in compressive stresses close to the crack tip. Of course the negative stress intensity factors are only useful wedge force solutions if after superposition with another stress field the resulting stress intensity factor is positive. For nonconstant stress fields given by $\sigma(s)$ the stress intensity factor can be determined by superposition from

$$k = \frac{1}{\sqrt{b}} \int_0^1 \frac{G(s)}{\sqrt{1-s}}\, \sigma(s)\, ds \tag{71}$$

For values of $h/b \sim 0.65$ and smaller, the contribution to the stress intensity factor due to concentrated force loading at the crack mouth is negligible. This can be seen clearly in Fig. 8, which plots the stress intensity factor due to concentrated force loading located at the crack mouth as a function of the periodic crack spacing (S is given by equation (67)). For wedge loading of the crack mouth the magnitude of the stress intensity factor decreases rapidly as the spacing between the surface cracks decreases. When the value of h/b is between approximately 2 and 0.65, the stress intensity factor due to wedge loading is negative. Smaller values of h/b ($S > 0.6$) yield very small contributions to the stress intensity factor.

Acknowledgment

The support provided by General Electric's Lighting Research and Technical Services Operation, which made this investigation possible, is gratefully acknowledged.

References

Bentham, J. P., and Koiter, W. T., 1973, "Asymptotic Approximations to Crack Problems," *Methods of Analysis and Solutions of Crack Problems*, Sih, G. C., ed., Noordhoff, Netherlands, pp. 131–178.

Bowie, O. L., 1973, "Solutions of Plane Crack Problems by Mapping Technique," *Methods of Analysis and Solutions of Crack Problems*, Sih, G. C., ed., Noordhoff, Netherlands, pp. 1–55.

Gradshteyn, I. S., and Rhyzhik, I. M., 1965, *Tables of Integrals Series and Products*, Academic Press, New York.

Hadamard, J., 1923, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Yale University Press.

Kaya, A. C., 1984, "Applications of Integral Equations with Strong Singularities in Fracture Mechanics," Ph.D. dissertation, Lehigh University.

Kaya, A. C., and Erdogan, F., 1984, "On the Solution of Integral Equations with Strong Singularities," *Numerical Solutions of Singular Integral Equations, Proceedings of IMACS Symposium*, Gerasoulis, A., and Vichnevetsky, R., eds., Rutgers University, pp. 54–57.

Kaya, A. C., and Erdogan, F., 1987, "On the Solution of Integral Equations with Strongly Singular Kernels," *Quarterly of Applied Mathematics*, Vol. 45, No. 1, pp. 105–122.

Koiter, W. T., 1965, Discussion of "Rectangular Tensile Sheet with Symmetrical Edge Cracks," by Bowie, O. L., ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, p. 237.

Nemat-Nasser, S., Keer, L. M., and Parihar, K. S., 1978, "Unstable Growth of Thermally Induced Interacting Cracks in Brittle Solids," *Int. J. Solids Structures*, Vol. 14, pp. 409–430.

Cumulant-Neglect Closure Method for Nonlinear Systems Under Random Excitations

Jian-Qiao Sun
Graduate Student.

C. S. Hsu
Professor,
Fellow ASME

Department of Mechanical Engineering,
University of California,
Berkeley, CA 94720

The validity of the cumulant-neglect closure method is examined by applying it to a system for which an exact solution is available. A comparison of the results indicates that the Gaussian closure technique usually leads to a mean-square versus excitation strength curve which follows the same general shape as that of the exact solution but has substantial errors in some cases. The 4th order cumulant-neglect method is found to be inapplicable and to predict erroneous behavior for systems in certain parameter ranges, including a faulty prediction of a jump in response as the excitation varies through a certain critical value. On the other hand, for systems in other ranges the 4th order cumulant-neglect closure method predicts the mean square response quite well. These two parameter ranges are delineated in the paper. The 6th order cumulant-neglect closure method is also examined, leading to similar conclusions.

1 Introduction

There are two popular closure methods for studying the statistical properties of the responses of stochastic nonlinear systems. The first one is the method of non-Gaussian closure (see Crandall, 1985, for example, and references therein). The idea of the method is to assume a non-Gaussian probability density function with adjustable parameters for the response and to use the moment relations derived from the system equations to obtain equations for the unknown parameters. The resulting probability density function can then provide approximate response statistics. The choice of trial density functions is of course open and important. As shown by Crandall (1985), when an inappropriate choice is made the approximate solution becomes worse as the order of the method goes higher. The second is the method of cumulant-neglect closure (Wu and Lin, 1984; Ibrahim, 1985). The method simply neglects the cumulants of system variables above certain order in order to close the infinite hierarchy of the equations governing the statistic moments of system variables. It has been shown by Wu and Lin (1984) that for certain problems this method is more versatile than the method of non-Gaussian closure and that, when the external excitation is dominant it gives better approximations as the order of the method is higher. However, only a limited range of parameter values is considered in Wu and Lin (1984). The general question of the validity of the cumulant-neglect closure method remains open.

It is understood that both methods essentially put constraints on the response statistics. As such, it is important to know whether the methods yield reasonable approximations to the true solutions and whether more constraints bring approximations closer to, or to cause them to deviate further from the true solutions. A mathematical proof of convergence of the method for general nonlinear systems is difficult and not available. What we propose to do is a case study on a class of systems.

In the paper, we study a stochastic nonlinear system which was originally studied by Dimentberg (1982) and which has an exact analytic solution for the stationary probability density function. We first present the exact solution by Dimentberg. Subsequently, we study the system by the method of Gaussian closure and the cumulant-neglect closure method with cumulants retained up to the fourth order. The purpose is to find in what way these methods of closure are adequate or inadequate. For the cases where the fourth order cumulant-neglect method is inadequate we study the sources of difficulties which lead to the inadequacies and give a qualitative assessment of the ranges of parameters in which the method is useful. Finally, we briefly study the sixth order cumulant-neglect method.

2 Exact Solution

Consider the following nonlinear system (Dimentberg, 1982)

$$\ddot{x}+2\alpha\dot{x}[1+\eta(t)]+\beta_1\dot{x}\left[x^2+\frac{\dot{x}^2}{\Omega^2}\right]+\Omega^2 x[1+\xi(t)] = \zeta(t),\ \beta_1\geq 0,\ \alpha>0 \tag{1}$$

where $\eta(t)$, $\xi(t)$ and $\zeta(t)$ are independent zero-mean Gaussian physical white noises in the sense of Gray and Caughey (1965) and

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

Discussion on this paper should be adressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 17, 1986; final revision, December 31, 1986.

$$\left.\begin{aligned} E[\eta(t)\eta(t+\tau)] &= D_\eta\delta(\tau) \\ E[\xi(t)\xi(t+\tau)] &= D_\xi\delta(\tau) \\ E[\zeta(t)\zeta(t+\tau)] &= D_\zeta\delta(\tau) \end{aligned}\right\}. \tag{2}$$

Let $v=\dot{x}$, the stationary joint probability density function $p(x,v)$ is governed by *FPK* equation

$$v\frac{\partial p}{\partial x}=\Omega^2 x\frac{\partial p}{\partial v}+\frac{\partial}{\partial v}\left\{\left[2\alpha(1-\alpha D_\eta)v+\beta_1 v\left(x^2+\frac{v^2}{\Omega^2}\right)\right]p\right\}$$

$$+\tfrac{1}{2}\frac{\partial^2}{\partial v^2}[(4\alpha^2 D_\eta v^2+\Omega^4 D_\xi x^2+D_\zeta)p]. \tag{3}$$

If

$$\Omega^2 D_\xi = 4\alpha^2 D_\eta, \tag{4}$$

then the equation (3) has an exact solution:

$$p(x,v)=C\frac{\exp\left[-\beta\left(x^2+\frac{v^2}{\Omega^2}\right)\right]}{\left(\kappa+x^2+\frac{v^2}{\Omega^2}\right)^{\delta-\kappa\beta}} \tag{5}$$

where

$$\kappa=\frac{D_\zeta}{\Omega^4 D_\xi},\quad \delta=\frac{2\alpha}{\Omega^2 D_\xi}+\tfrac{1}{2},\quad \beta=\frac{\beta_1}{\Omega^2 D_\xi} \tag{6}$$

$$C^{-1}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\left[-\beta\left(x^2+\frac{v^2}{\Omega^2}\right)\right]}{\left(\kappa+x^2+\frac{v^2}{\Omega^2}\right)^{\delta-\kappa\beta}}dxdv. \tag{7}$$

It can be shown from equation (5) that

$$E(x^2)=\frac{1}{\Omega^2}E(v^2). \tag{8}$$

Introducing a transformation

$$x=r\cos\varphi\quad v=\Omega r\sin\varphi, \tag{9}$$

we can obtain following results

$$p(r)=2\pi C\Omega\frac{r\exp(-\beta r^2)}{(\kappa+r^2)^{\delta-\kappa\beta}} \tag{10}$$

$$E(x^2)=\tfrac{1}{2}E(r^2)=\tfrac{1}{2}\pi C\Omega\int_0^\infty\frac{r\exp(-\beta r)}{(\kappa+r)^{\delta-\kappa\beta}}dr \tag{11}$$

$$C^{-1}=\pi\Omega\int_0^\infty\frac{\exp(-\beta r)}{(\kappa+r)^{\delta-\kappa\beta}}dr. \tag{12}$$

When $\beta>0$ and $\kappa\neq 0$, the stationary statistic moments of any order exist for an arbitrary δ and there is no bifurcation possible in this case.

An interesting case is when $\beta>0$ and $\kappa=0$. $\kappa=0$ implying $\zeta(t)=0$, the system is only under parametric excitations. Equations (11) and (12) become

$$E(x^2)=\tfrac{1}{2}\pi C\Omega\int_0^\infty r^{1-\delta}\exp(-\beta r)dr \tag{13}$$

$$C^{-1}=\pi\Omega\int_0^\infty r^{-\delta}\exp(-\beta r)dr. \tag{14}$$

Replacing C in equations (13) by (14), we have a rather simple expression:

$$E(x^2)=\frac{1}{2\beta}\frac{\Gamma(2-\delta)}{\Gamma(1-\delta)}=\frac{1}{2\beta}(1-\delta). \tag{15}$$

The conditions for the existence of integrals in equations (13) and (14) are, respectively,

$$\delta\leq 2,\quad \delta\leq 1. \tag{16}$$

Together, the condition for the existence of normalization constant C and $E(x^2)$ is

$$\delta\leq 1. \tag{17}$$

It is easy to check that (17) is also the condition for the existence of statistical moments of any order of the system. Moreover, $\delta=1$ is the condition of bifurcation. When $\delta>1$, equation (15) is not valid. This suggests that the trivial solution exists and is stable. The stationary statistical moments of all orders are zero. In particular, $E(x^2)=0$. So we have

$$E(x^2)=\begin{cases}0 & \delta>1\\ 0 & \delta=1\\ \dfrac{1}{2\beta}(1-\delta) & \delta<1\end{cases}. \tag{18}$$

Equation (18) shows that, as δ varies across the bifurcation point, there is no jump in $E(x^2)$. It can also be shown that there is no jump in any other moments.

3 Cumulant-Neglect Closure Methods

The Ito stochastic differential equations for the Markov vector (x,v) can be written down from equation (3). Let $x_1=x$, $x_2=v$, then

$$\left.\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= \left(-2\alpha x_2-\beta_1 x_1^2 x_2-\frac{\beta_1}{\Omega^2}x_2^3-\Omega^2 x_1\right)dt+2\alpha^2 D_\eta x_2 dt \\ &\quad +(4\alpha^2 D_\eta x_2^2+\Omega^4 D_\xi x_1^2+D_\zeta)^{1/2}dW(t) \end{aligned}\right\} \tag{19}$$

where $W(t)$ is a unit Wiener process, $2\alpha^2 D_\eta x_2$ is the drift correction term of Wong and Zakai (1965)

Let

$$m_{ij}=E(x_1^i x_2^j) \tag{20}$$

where $E(x_1^i x_2^j)$ is stationary mixed moment of $(i+j)$th order. We are only interested in stationary solutions. From equations (19), we derive the equations of stationary moments up to the 4th order.

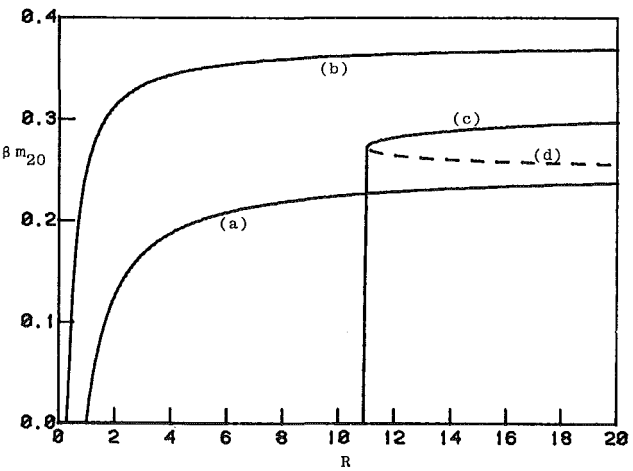


Fig. 1 Stationary mean square value of the response versus R for $\kappa=0$: (*a*) equation (18), the exact solution; (*b*) equation (31), Gaussian closure; (*c*) equation (41), the fourth order method; (*d*) neglected branch of equation (37)

$$\left.\begin{aligned} &m_{01}=0\\ &(2\alpha^2D_\eta-2\alpha)m_{01}-\Omega^2m_{10}-\beta_1m_{21}-\frac{\beta_1}{\Omega^2}\,m_{03}=0 \end{aligned}\right\} \tag{21}$$

$$\left.\begin{aligned} &m_{11}=0\\ &m_{02}-\Omega^2m_{20}+(2\alpha^2D_\eta-2\alpha)m_{11}-\beta_1m_{31}-\frac{\beta_1}{\Omega^2}\,m_{13}=0\\ &(8\alpha^2D_\eta-4\alpha)m_{02}-2\beta_1m_{22}-2\frac{\beta_1}{\Omega^2}\,m_{04}-2\Omega^2m_{11}\\ &+\Omega^4D_\xi m_{20}+D_\zeta=0 \end{aligned}\right\} \tag{22}$$

$$\left.\begin{aligned} &m_{21}=0\\ &2m_{12}+(2\alpha^2D_\eta-2\alpha)m_{21}-\beta_1m_{41}-\frac{\beta_1}{\Omega^2}\,m_{23}-\Omega^2m_{30}=0\\ &m_{03}+(8\alpha^2D_\eta-4\alpha)m_{12}-2\beta_1m_{32}-2\,\frac{\beta_1}{\Omega^2}\,m_{14}-2\Omega^2m_{21}\\ &+\Omega^4D_\xi m_{30}+D_\zeta m_{10}=0\\ &(18\alpha^2D_\eta-6\alpha)m_{03}-3\beta_1m_{23}-3\,\frac{\beta_1}{\Omega^2}\,m_{05}-3\Omega^2m_{12}\\ &+3\Omega^4D_\xi m_{21}+3D_\zeta m_{01}=0 \end{aligned}\right\} \tag{23}$$

$$\left.\begin{aligned} &m_{31}=0\\ &3m_{22}+(2\alpha^2D_\eta-2\alpha)m_{31}-\beta_1m_{51}-\frac{\beta_1}{\Omega^2}\,m_{33}-\Omega^2m_{40}=0\\ &2m_{13}+(8\alpha^2D_\eta-4\alpha)m_{22}-2\beta_1m_{42}-2\,\frac{\beta_1}{\Omega^2}\,m_{24}\\ &+\Omega^4D_\xi m_{40}+D_\zeta m_{20}=0\\ &m_{04}+(18\alpha^2D_\eta-6\alpha)m_{13}-3\beta_1m_{33}-3\,\frac{\beta_1}{\Omega^2}\,m_{15}\\ &-3\Omega^2m_{22}+3\Omega^4D_\xi m_{31}+3D_\zeta m_{11}=0\\ &(32\alpha^2D_\eta-8\alpha)m_{04}-4\beta_1m_{24}-4\frac{\beta_1}{\Omega^2}\,m_{06}-4\Omega^2m_{13}\\ &+6\Omega^4D_\xi m_{22}+6D_\zeta m_{02}=0 \end{aligned}\right\} \tag{24}$$

3.1 Gaussian Closure. Let cumulant functions of x_1 and x_2 be denoted by $\lambda_n(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$. In the method of Gaussian closure we set the cumulants, corresponding to the moments in equations (21) and (22) of order greater than and equal to three, equal to zero in order to close the hierarchy of moment equations at the second order. Then these moments are expressed in terms of the first and second order moments. Use the formulas of cumulant functions in (Ibrahim, 1985), for example, we have the following relations after using equations (21) and (22)

$$\left.\begin{aligned} m_{03} &= 0\\ m_{21} &= 0\\ m_{12} &= m_{02}m_{10}\\ m_{30} &= 3m_{10}m_{20}-2m_{10}^3 \end{aligned}\right\} \tag{25}$$

$$\left.\begin{aligned} m_{04} &= 3m_{02}^2\\ m_{13} &= m_{10}m_{30}\\ m_{22} &= m_{20}m_{02}+2m_{10}m_{12}-2m_{02}m_{10}^2\\ m_{31} &= 0 \end{aligned}\right\} . \tag{26}$$

Put equations (25) and (26) in equations (21) and (22), we have

$$\left.\begin{aligned} m_{10} &= m_{30}=m_{13}=m_{12}=0\\ m_{02} &= \Omega^2m_{20} \end{aligned}\right\} \tag{27}$$

$$8\beta_1m_{20}^2-(8\alpha^2D_\eta+\Omega^2D_\xi-4\alpha)m_{20}-\frac{D_\zeta}{\Omega^2}=0. \tag{28}$$

When $4\alpha^2D_\eta=\Omega^2D_\xi$, we have

$$8\beta m_{20}^2-2(2-\delta)m_{20}-\kappa=0. \tag{29}$$

Special cases.

1. When $D_\eta=D_\zeta=0$ and $\beta_1>0$, from equation (28) we have

$$m_{20}=0 \text{ or } m_{20}=\frac{\Omega^2D_\xi-4\alpha}{8\beta_1}$$

The first one is the trivial solution, while the second one gives the well known bifurcation condition by Gaussian closure (Wu and Lin, 1984; Ibrahim, 1985; Ariaratnam, 1980):

$$\Omega^2D_\xi=4\alpha.$$

2. When $\beta_1=0$, $D_\eta=D_\xi=0$ and $D_\zeta=2\pi S_0$, then equation (28) gives the exact linear solution

$$m_{20}=\frac{\pi S_0}{2\alpha\Omega^2}.$$

3. When $4\alpha^2D_\eta=\Omega^2D_\xi$, $\kappa=0$ and $\beta>0$, we have from equation (29)

$$m_{20}=0 \text{ or } m_{20}=\frac{1}{4\beta}(2-\delta). \tag{30}$$

Notice that when $\kappa=0$ ($D_\zeta=0$), the trivial (identically zero) solution always exists and its stability can be ascertained by examining its bifurcation into the nontrivial solution. Equation (30) shows that the bifurcation condition is $\delta=2$ and

$$m_{20}=\begin{cases} 0 & \delta>2\\ 0 & \delta=2\\ \dfrac{1}{4\beta}(2-\delta) & \delta<2\,. \end{cases} \tag{31}$$

Once again, there is no jump in m_{20} at bifurcation point.

3.2 The 4th Order Method of Cumulant-Neglect Closure. For the remainder of the paper we take $\beta>0$. Consider now the 4th order cumulant-neglect closure method. In this method, in the same spirit as in the method of Gaussian closure, we set the cumulants, corresponding to the moments of order 5 and 6 in equations (23) and (24), to zero to close the hierarchy of the moment equations at the order 4. This leads to

$$\lambda_n(x_{i_1}x_{i_2}, \ldots, x_{i_n})=0,\qquad n=5,6;\ (i_1,i_2,\ldots,i_n)\neq(1,1,\ldots,1). \tag{32}$$

From equations (21) to (24) and (32), one can show that

$$m_{10}=m_{01}=m_{11}=m_{12}=m_{21}=m_{31}=m_{30}=m_{03}=0$$
$$m_{05}=m_{14}=m_{23}=m_{32}=m_{41}=m_{51}=0.$$

In general $m_{13} \neq 0$. However, when $4\alpha^2 D_\eta = \Omega^2 D_\xi$ it can be shown that $m_{13} = 0$ and

$$\left.\begin{aligned} m_{02} &= \Omega^2 m_{20} \\ m_{22} &= \frac{m_{04}}{3\Omega^2} \\ m_{04} &= \frac{3\Omega^4(24\beta m_{20}^2 + \kappa)m_{20}}{36\beta m_{20} + 2\delta - 6} \end{aligned}\right\}. \tag{33}$$

Substituting (33) into $(22)_3$, we obtain the equation for m_{20}

$$96(\beta m_{20})^3 - 36(2-\delta)(\beta m_{20})^2 - [14\beta\kappa + 2(2-\delta)(\delta-3)](\beta m_{20}) - (\delta-3)\beta\kappa = 0. \tag{34}$$

Once m_{20} is determined by equation (34), m_{02} can be computed from equation $(33)_1$, m_{22} from $(33)_2$, and m_{04} from $(33)_3$. However, since m_{04} should be nonnegative, it is seen from $(33)_3$ that for nontrivial solutions there is a constraint on m_{20}

$$\beta m_{20} \geq (3-\delta)/18. \tag{35}$$

3.2.1. Purely Parametric Excitation $\kappa = 0$. To proceed further, let us first examine the case $\kappa = 0$ for which the analysis is simple and transparent. When $\kappa = 0$, the external forcing is absent and the excitation is entirely parametric. In that case we have

$$m_{20} = 0 \text{ or } m_{20} = \frac{18(2-\delta) \pm \sqrt{324(2-\delta)^2 + 192(2-\delta)(\delta-3)}}{96\beta}. \tag{36}$$

The solution $m_{20} = 0$ corresponds to the trivial solution of identically zero response which is always possible in this case. The nontrivial solutions for m_{20} are obtained from equation (36). It is readily seen that there are real solutions for m_{20} only if

$$324(2-\delta)^2 + 192(2-\delta)(\delta-3) \geq 0. \tag{37}$$

This condition in turn requires

$$\delta \leq \frac{6}{11} \text{ or } \delta \geq 2. \tag{38}$$

However, for $\delta \geq 3$, the two real solutions are both negative and therefore are not valid ones for m_{20}. For $2 \leq \delta \leq 3$, there is one positive real solution for m_{20}, but its value violates the constraint equation (35) and hence is to be ruled out. Consequently, there are real and valid nontrivial solutions for m_{20} only in the range $\delta \leq (6/11)$. In this range there are positive real solutions for m_{20}. But, it can be shown that the solution associated with the negative square root decreases as the strength of the excitations D_ξ increases. It is thus physically not acceptable. This branch of the solution is shown in Fig. 1 by the dashed line. Summarizing the results, we have for the case $\kappa = 0$

$$m_{20} = \begin{cases} 0 & \delta > \dfrac{6}{11} \\ \dfrac{0.2727}{\beta} & \delta = \dfrac{6}{11} \\ \dfrac{18(2-\delta) + \sqrt{324(2-\delta)^2 + 192(2-\delta)(\delta-3)}}{96\beta} & \delta < \dfrac{6}{11}. \end{cases} \tag{39}$$

As readily seen from equation (39), there is a jump in m_{20} as δ varies across the point $\delta = 6/11$.

A Discussion. Shown in Fig. 1 are the curves of $(\beta \times m_{20})$ versus R of the exact solution and the solutions by the method of Gaussian closure and by the 4th order method, where R is defined as:

$$R = \frac{1}{2\delta - 1} = \frac{\Omega^2 D_\xi}{4\alpha}.$$

The bifurcation conditions in terms of R are the following

$$\left.\begin{aligned} &R = 1 \quad \text{(exact)} \\ &R = 1/3 \quad \text{(Gaussian closure)} \\ &R = 11 \quad \text{(the 4th order method,} \\ &\qquad \text{not a bifurcation but a jump)} \end{aligned}\right\}.$$

As can be seen from the Fig. 1, the solution by Gaussian closure has large error consistently for all $R > 1/3$. The 4th order method gives a better approximation for $R \geq 11$ and the wrong solution for $1 < R < 11$ and a faulty jump in m_{20}. It is interesting to indicate why this happens.

For the 4th order method, the condition (32) gives following moment relations besides some trivial ones:

$$\left.\begin{aligned} m_{06} &= 15 m_{02} m_{04} - 30 m_{02}^3 \\ m_{42} &= m_{02} m_{40} + 6 m_{20} m_{22} - 6 m_{02} m_{20}^2 \\ m_{24} &= m_{20} m_{04} + 6 m_{02} m_{22} - 6 m_{20} m_{02}^2 \end{aligned}\right\}. \tag{40}$$

The exact relations corresponding to equations (40) can be obtained explicitly from the exact solution (5) when $\kappa = 0$

$$\left.\begin{aligned} m_{06} &= \frac{5\Omega^2}{2\beta}\left(1 - \frac{\delta}{3}\right) m_{04} \\ m_{42} &= \frac{1}{2\beta\Omega^2}\left(1 - \frac{\delta}{3}\right) m_{04} \\ m_{24} &= \frac{1}{2\beta}\left(1 - \frac{\delta}{3}\right) m_{04} \end{aligned}\right\} \tag{41}$$

and

$$m_{04} = \frac{3\Omega^4}{2\beta}\left(1 - \frac{\delta}{2}\right) m_{20}. \tag{42}$$

Comparing equations (40) and (41), one can see that the condition (32) imposes nonlinear relations on moments. It is understood that these nonlinear constraints are the causes of the inequalities (35) and (37) which in turn result in the wrong bifurcation condition and the faulty jump in moments at $\delta = 6/11$ $(R = 11)$.

Applying equations (41) to (21) to (24), we find, as might have been expected, that some of the equations become identities while others become the exact relations of moments such as

$$\left.\begin{aligned} &m_{02} = \Omega^2 m_{20} \\ &\Omega^2 m_{40} = 3 m_{22} \\ &m_{04} = 3\Omega^2 m_{22} \\ &m_{22} = \frac{\Omega^2}{2\beta}\left(1 - \frac{\delta}{2}\right) m_{20} \end{aligned}\right\}. \tag{43}$$

Equations (41), (42), and (43) show that the moments of any order of the system are linearly related and of a non-Gaussian nature. The method of cumulant-neglect closure, on the other hand, requires the moments above certain order to satisfy some nonlinear relations; this is where the solution by the method deviates from the exact solution. When the order of

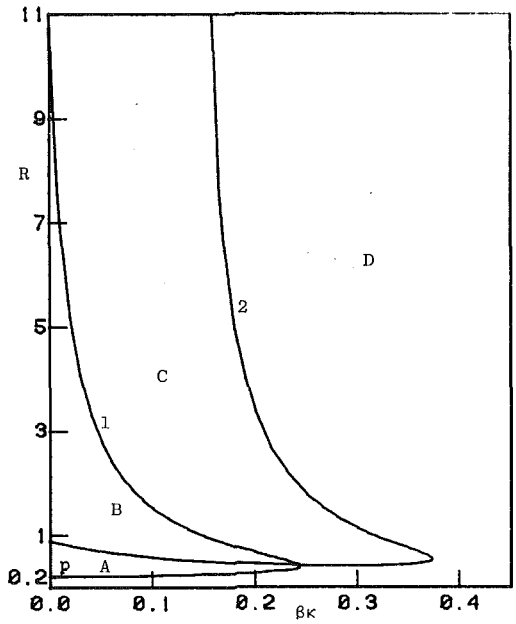


Fig. 2 Invalid parameter domain of the fourth and sixth order methods

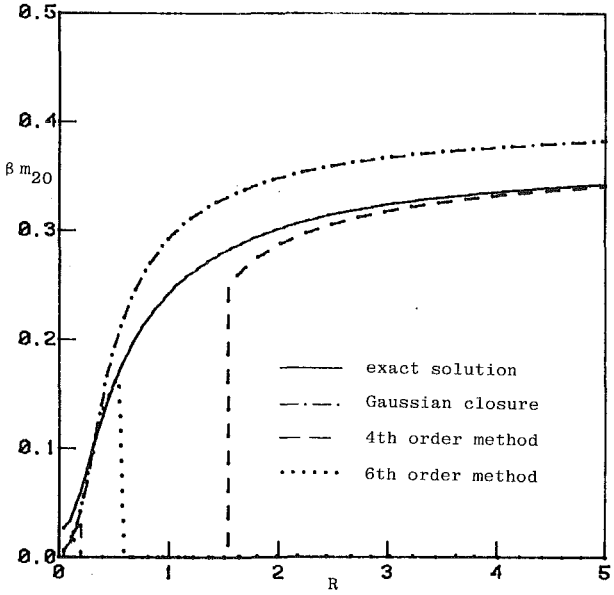


Fig. 3 Stationary mean square value of the response versus R for $\kappa = 0.1$, $\beta = 1$, $\beta\kappa = 0.1$

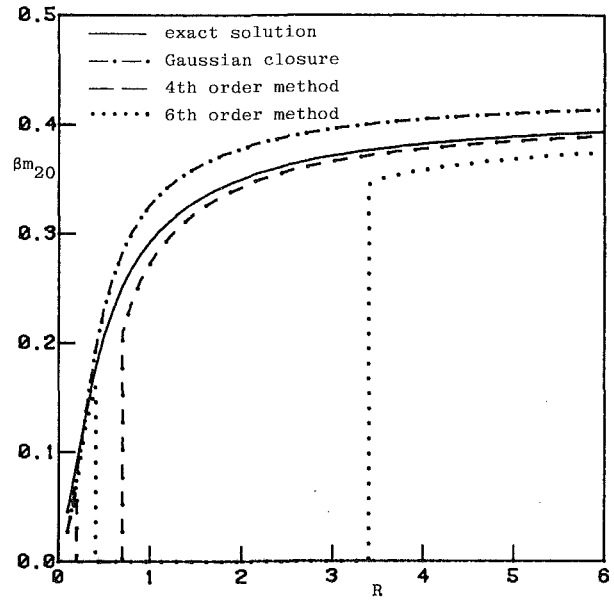


Fig. 4 Stationary mean square value of the response versus R for $\kappa = 0.2$, $\beta = 1$, $\beta\kappa = 0.2$

the method becomes higher, higher-order nonlinear constraint relations on moments are introduced. The higher-order nonlinear relations may introduce more constraints like equations (35) and (37). It is difficult to know whether such higher order constraints make the approximate solutions converge to the exact solution. This specific case $\kappa = 0$ considered here, however, clearly shows that the constraints introduced at the fourth order level actually make the method invalid and bring about a faulty solution for a range of the system parameters.

3.2.2. The General Case $\kappa \neq 0$. Next, let us consider the general case $\kappa \neq 0$. Here what we need is to find all the real and positive solutions of m_{20} from equation (34) which also satisfy the constraint equation (35). These solutions will be referred to as meaningful solutions. A detailed calculation for this task yields curve 1 in Fig. 2. In this figure, area $C+D$ is a region in the $\beta\kappa - R$ parameter plane where there exist meaningful solutions. Area $A+B$ of the figure is a region where there is no meaningful solution because the solutions are either complex, real but negative, or violating (35).

The neighborhood near point p on the R axis with $R = 0.2$ needs some special comments. The R coordinate axis, or $\kappa = 0$, is a special case. As discussed earlier, when $\kappa = 0$ there is the trivial solution for all positive values of R. The nontrivial solutions exist, however, only for $R > 11$, according to the fourth order cumulant-neglect method. As soon as κ deviates from zero, meaningful solutions with small values of m_{20} appear for R-values less than approximately 0.2.

Figure 2 implies that it is inappropriate to use the fourth order cumulant-neglect method for systems with parameters lying in region $A+B$. For systems in region $C+D$ the method will yield meaningful solutions. It is then important to know how good are the solutions of the method for these systems. Here we again compare the results against the exact solutions. Some of the calculated results are shown in Figs. 3–6. They are for $\beta\kappa = 0.1$, 0.2, 1.0, and 10, respectively. It is interesting to observe from these figures that if a system is located in area $C+D$ of Fig. 2 and not too near the boundary of the area, then the predicted mean square of the response is a quite good approximation to the true value. The cases considered by Wu and Lin (1984) fall in the area $C+D$ (with $R = 0.05$); hence, they have found good agreement between the results from the 4th order cumulant-neglect method and the true solutions. This is consistent with the findings of this study. For the fourth order method, besides the curve shown in each of Figs. 3–6, there exists also a second branch with much lower βm_{20} values. These are not considered in the above discussion. A further study is needed if we wish to have a more rigorous basis for excluding them.

3.3 The Sixth Order Method of Cumulant-Neglect Closure. Motivated by the results of the fourth order method, we now study the sixth order method. By keeping the moment equations up to the sixth order, we have to solve 27 equations together with 15 nonlinear relations for the moments of seven and eight orders. It has been observed that under the condition (4), both the Gaussian closure and the fourth order method give the following result.

$$m_{ij} = 0, \text{ if either } i \text{ or } j \text{ is odd.} \tag{44}$$

which is in agreement with the exact solution. It can also be shown that there is a branch of the stationary solutions by the sixth order method for which equation (44) is true. We intend

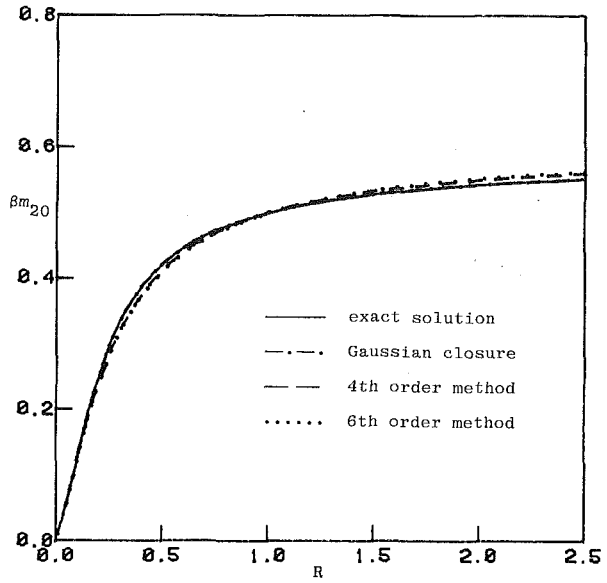


Fig. 5 Stationary mean square value of the response versus R for $\kappa = 1$, $\beta = 1$, $\beta\kappa = 1$

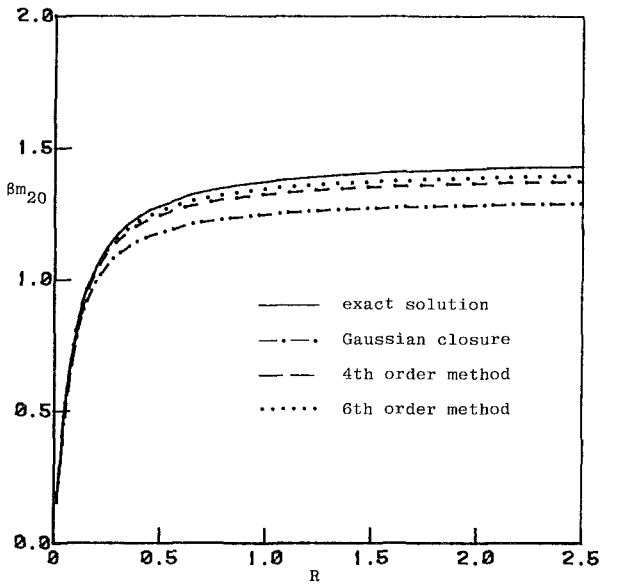


Fig. 6 Stationary mean square value of the response versus R for $\kappa = 1$, $\beta = 10$, $\beta\kappa = 10$

to find and examine such a branch of solution. With equation (44) and the condition (4), we obtain following three equations

$$\left.\begin{aligned} &2(2-\delta)m_{20} - \frac{8}{3}\beta m_{40} + \kappa = 0 \\ &3\kappa m_{20} + 2(3-\delta)m_{40} - \frac{12}{5}\beta m_{60} = 0 \\ &5\kappa m_{40} + 2(4-\delta)m_{60} - \frac{16}{7}\beta m_{80} = 0 \end{aligned}\right\} \tag{45}$$

and the moment relations

$$\left.\begin{aligned} m_{02} &= \Omega^2 m_{20} \\ m_{42} &= \frac{\Omega^2}{5} m_{60}, m_{24} = \frac{\Omega^4}{5} m_{60}, m_{06} = \Omega^6 m_{60} \\ m_{62} &= \frac{\Omega^2}{7} m_{80}, m_{44} = \frac{3\Omega^4}{35} m_{80}, m_{26} = \frac{\Omega^6}{7} m_{80}, m_{08} = \Omega^8 m_{80} \end{aligned}\right\} \tag{46}$$

$$m_{80} = 630m_{20}^4 - 420m_{20}^2 m_{40} + 35m_{40}^2 + 28m_{20}m_{60}. \tag{47}$$

From equations (45) and (47), we obtain the equation for m_{20}

$$a_0(\beta m_{20})^4 + a_1(\beta m_{20})^3 + a_2(\beta m_{20})^2 + a_3(\beta m_{20}) + a_4 = 0 \tag{48}$$

where

$$\begin{aligned} a_0 &= 288 \\ a_1 &= -144(2-\delta) \\ a_2 &= (2-\delta)(42-17\delta) - 56\beta\kappa \\ a_3 &= -(11.75\delta - 26.5)\beta\kappa - \frac{1}{4}(2-\delta)(3-\delta)(4-\delta) \\ a_4 &= \frac{1}{8}[15\beta\kappa - (3-\delta)(4-\delta)]\beta\kappa \,. \end{aligned}$$

Notice that all the relations in equations (46) are in agreement with exact ones. From equations (45), one may also derive two constraints on m_{20} to make m_{40} and m_{60} positive. As in the case of the fourth order method, a comprehensive numerical study for the meaningful solutions of equation (48) is carried out in the $\beta\kappa - R$ parameter plane. The result is curve 2 in Fig. 2. The meaningful solutions are chosen out of several branches of the solutions of equation (48) strictly based upon comparison with the exact solutions. Thus area $B + C$ in Fig. 2 is a region where no meaningful solution exists because the solutions are either complex, real but negative, or real and positive but of the value which remains small for all excitation strength, or decreases as the excitation strength increases. Area $A + D$ is a region where there exist meaningful solutions.

3.3.1. Comparison with the Fourth Order Method. It is interesting to discuss the features of the cumulant-neglect closure methods revealed in Fig. 2. Area A in Fig. 2 is the region where the 4th order method is invalid but the sixth order method is valid. In this region one can expect to improve the approximate solutions of lower order methods by using the sixth order method. Area C is the reverse of Area A. In area C, the sixth order method is invalid but the fourth order method is valid. This suggests that in area C, one should stop at the fourth order method because the higher order method may not give better approximations. Area B is the region where both the fourth and sixth methods are invalid. It seems to suggest that the non-Gaussian closure methods are inapplicable in area B. Here, it is interesting to note from Figs. 3-4 that in the area B, the Gaussian closure method, while it does not give accurate results, does yield a qualitatively correct response pattern. Area D is the region where both the fourth and sixth order methods are valid. In the region, the sixth order method gives better approximations than the fourth order method if the parameters are located far away from the boundary of the invalid domain. All the features discussed thus far are clearly displayed in Figs. 3–6.

4 Conclusion

A nonlinear system under random parametric and external excitations is studied with the method of Gaussian closure and the fourth and sixth order cumulant-neglect closure methods. The approximate solutions are compared with an exact analytic solution. It is shown that for the fourth and sixth order cumulant-neglect closure methods one can determine two regions in the parameter plane. In one region the methods are invalid and give faulty results. For the other region the methods give meaningful solutions and, in fact, give fairly good approximate solutions in most instances.

This is merely a case study on one class of systems. It seems to indicate that the cumulant-neglect closure method is a very useful tool, but it also suggests that the validity of the method needs verifying when it is applied to general nonlinear systems under random excitations. One way of verifying is to check the

results obtained by the cumulant-neglect closure methods with direct simulation. Another way is to use the statistic hypothesis testing technique to establish the guideline of validity of the method. This is, however, a future research topic and can not be covered here.

Acknowledgment

Authors wish to acknowledge the reviewers for their valuable comments. This material is based upon work supported by the National Science Foundation under grant No. MSM-8519950.

References

Ariaratnam, S. T., 1980, "Bifurcation in Nonlinear Stochastic Systems," *New Approaches to Nonlinear Problems in Dynamics*, Holmes, P. J., ed., SIAM, pp. 470–474.

Crandall, S. H., 1985, "Non-Gaussian Closure Techniques for Stationary Random Vibration," *Int. J. Non-Lin. Mech.*, Vol. 20, pp. 1–8.

Dimentberg, M. F., 1982, An Exact Solution to a Certain Non-Linear Random Vibration Problem," *Int. J. Non-Linear Mechanics*, Vol. 17, No. 4, pp. 231–236.

Gray, A. H. Jr., and Caughey, T. K., 1965, "A Controversy in Problems Involving Random Parametric Excitation," *J. of Math. and Phy.*, Vol. 44, pp. 288–296.

Ibrahim, R. A., et al., 1985, "Stochastic Response of Nonlinear Dynamic Systems Based on a Non-Gaussian Closure," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 965–970.

Wong, E., and Zakai, M., 1965, "On the Relation Between Ordinary and Stochastic Differential Equations," *Int. J. Eng. Sci.*, Vol. 3, pp. 213–229.

Wu, W. F., and Lin, Y. K., 1984, "Cumulant-Neglect Closure for Non-Linear Oscillators Under Random Parametric and External Excitations," *Int. J. Non-Lin. Mech.*, Vol. 19, No. 4, pp. 349–362.

Integrals of Linearized Differential Equations of Motion of Mechanical Systems; Part I: Linearized Differential Equations

T. R. Kane
Professor of Applied Mechanics,
Fellow ASME

S. Djerassi
Stanford University,
Stanford, CA 94305

Part I of this paper deals with relationships between integrals of a set of nonlinear differential equations and integrals of equations obtained from such a set by a process of linearization. The terms "total linearization," "reduction," and "partial linearization" are introduced, a theorem is stated and proved in connection with each of these, and the use of each theorem is illustrated by means of an example. In Part II, the theorems are applied to differential equations of motion of mechanical systems.

1 Introduction

The subject of this paper is introduced most easily by reference to a simple example. Figure 1 shows a mechanical system S consisting of two particles, P_1 and P_2, each of mass M, and two light, rigid rods, R_1 and R_2, each of length L. R_1 and R_2 are connected to each other by a revolute joint and a linear torsion spring of modulus k at P_2, and R_2 is connected by means of a revolute joint at O to a vertical shaft that is made to rotate at a constant angular speed Ω.

If u_1 and u_2 are defined as $\dot{q}_1$ and $\dot{q}_2$, respectively, where q_1 and q_2 are the angles between the vertical and R_1 and R_2, respectively, and dots denote time-differentiation, then all motions of S are governed by the four first-order differential equations:

$$\dot{q}_1 = u_1 \tag{1.1}$$

$$\dot{q}_2 = u_2 \tag{1.2}$$

$$\dot{u}_1 + \dot{u}_2\cos(q_1 - q_2) = \Omega^2 c_1(s_1 + s_2) - u_2^2\sin(q_1 - q_2) - [k/(ML^2)](q_1 - q_2) - (g/L)s_1 \tag{1.3}$$

$$\dot{u}_1\cos(q_1 - q_2) + 2\dot{u}_2 = \Omega^2 c_2(s_1 + 2s_2) + u_1^2\sin(q_1 - q_2) + [k/(ML^2)](q_1 - q_2) - 2(g/L)s_2 \tag{1.4}$$

where s_i and c_i denote $\sin q_i$ and $\cos q_i$ ($i = 1,2$), respectively. These equations cannot be solved in closed form. Hence, to obtain a description of the motion of S that takes place subsequent to an instant at which u_1, u_2, q_1, and q_2 have specified initial values, one resorts to numerical integration. In this way one can generate results such as, for example, those represented by the two curves labeled q_1 and q_2 in Fig. 2 or, equivalently, those reported in columns 2 and 3 of Table 1, which apply if $g/L = 1\text{s}^{-2}$, $ML^2 = 1$ kg m^2, $k/(ML^2) = 1.27289\ \text{s}^{-2}/\text{rad}$, $\Omega^2 = 0.69773\ \text{rad}^2\ \text{s}^{-2}$, and, at $t = 0$, $u_1 = u_2 = 0$, $q_1 = 23$ deg and $q_2 = 30$ deg.

To check on the validity of results obtained by the means just described, one can take advantage of the fact that equations (1.1)–(1.4) possess the integral

$$G(u_1, u_2, q_1, q_2) = C, \quad \text{a constant} \tag{1.5}$$

where the function G (the Hamiltonian of S) is given by

$$G = \frac{1}{2}\left[u_1^2 + 2u_2^2 + 2u_1u_2\cos(q_1 - q_2) - \Omega^2(s_1^2 + 2s_2^2 + 2s_1s_2)\right] + \frac{1}{2}\frac{k}{ML^2}(q_1 - q_2)^2 - \frac{g}{L}(c_1 + 2c_2) \tag{1.6}$$

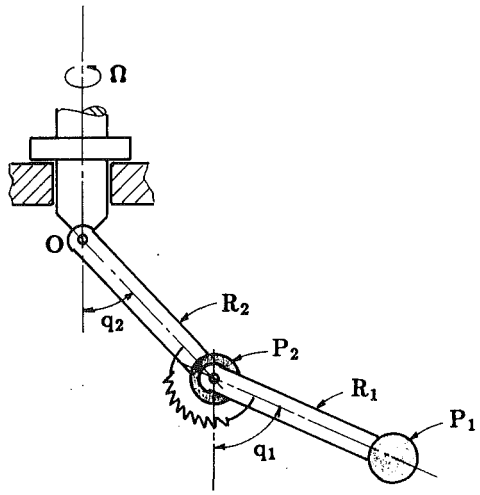


Fig. 1 Mechanical system

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August 22, 1986; final revision, December 1, 1986.

Table 1

1 t sec	2 q_1 deg	3 q_2 deg	4 G Nm	5 $\hat{q}_1$ deg	6 $\hat{q}_2$ deg	7 G Nm	8 $\tilde{G}$ Nm	9 $\hat{q}_1$ deg	10 q_2 deg	11 $\tilde{G}$ Nm
0.0	23.0	30.0	-3.0071	-10.0	0.0	-2.9759	-3.0062	- 7.0	30.0	-3.0068
1.0	34.9	22.5	-3.0071	2.0	-7.5	-2.9739	-3.0062	12.5	22.5	-3.0081
2.0	24.2	31.1	-3.0071	8.51	1.39	-2.9736	-3.0062	- 6.9	31.3	-3.0064
3.0	37.0	24.6	-3.0071	4.7	-4.9	-2.9717	-3.0062	12.7	25.0	-3.0075
4.0	27.2	33.7	-3.0071	- 5.1	4.6	-2.9716	-3.0062	- 6.6	34.6	-3.0054
5.0	40.1	27.6	-3.0071	8.2	-1.7	-2.9732	-3.0062	12.9	28.4	-3.0066
6.0	30.1	36.3	-3.0071	- 2.2	7.3	-2.9736	-3.0062	- 6.5	37.6	-3.0044
7.0	41.9	29.3	-3.0071	10.0	0.0	-2.9759	-3.0062	13.1	30.4	-3.0061
8.0	30.6	36.8	-3.0071	- 1.8	7.6	-2.9741	-3.0062	- 6.5	38.0	-3.0042
9.0	41.0	28.4	-3.0071	8.8	-1.1	-2.9740	-3.0062	13.0	29.1	-3.0064
10.0	28.4	34.7	-3.0071	- 4.3	5.3	-2.9718	-3.0062	- 6.6	35.4	-3.0051

To this end, one incorporates in the computer program used to carry out the numerical integration of equations (1.1)–(1.4) the evaluation of G as given by equation (1.6), and determines whether or not the value thus found stays constant to an extent compatible with the accuracy of the numerical integration scheme being employed. In column 4 of Table 1, the value of G generated in this way is seen to remain constant to five significant figures. The same result is reported in Fig. 2 as the straight line labeled G.

Particular solutions of equations (1.1)–(1.4) can be found by setting $u_1 = u_2 = 0$ and letting $\bar{q}_1$ and $\bar{q}_2$ be constants such that

$$\Omega^2 \bar{c}_1(\bar{s}_1 + \bar{s}_2) - [k/(ML^2)](\bar{q}_1 - \bar{q}_2) - (g/L)\bar{s}_1 = 0 \tag{1.7}$$

$$\Omega^2 \bar{c}_2(\bar{s}_1 + 2\bar{s}_2) + [k/(ML^2)](\bar{q}_1 - \bar{q}_2) - 2(g/L)\bar{s}_2 = 0 \tag{1.8}$$

where $\bar{s}_i$ and $\bar{c}_i$ denote sin $\bar{q}_i$ and cos $\bar{q}_i$ ($i = 1, 2$), respectively. For example, with g/L, ML^2, $k/(ML^2)$, and Ω^2 as before, these equations are satisfied with $\bar{q}_1 = 33$ deg, $\bar{q}_2 = 30$ deg. Suppose now that, to study motions differing only slightly from that corresponding to such a particular solution, one introduces perturbations $\hat{u}_1$, $\hat{u}_2$, $\hat{q}_1$, and $\hat{q}_2$ as $\hat{u}_i \triangleq u_i$, $\hat{q}_i \triangleq q_i - \bar{q}_i$ ($i = 1, 2$), substitutes into equations (1.1)–(1.4), and then linearizes in the perturbations and their time-derivatives. This yields the equations

$$\dot{\hat{q}}_1 = \hat{u}_1 \tag{1.9}$$

$$\dot{\hat{q}}_2 = \hat{u}_2 \tag{1.10}$$

$$\dot{\hat{u}}_1 + \dot{\hat{u}}_2\cos(\bar{q}_1 - \bar{q}_2) = \Omega^2\left[(\bar{c}_1^2 - \bar{s}_1^2 - \bar{s}_1\bar{s}_2)\hat{q}_1 + \bar{c}_1\bar{c}_2\hat{q}_2\right] - [k/(ML^2)](\hat{q}_1 - \hat{q}_2) - (g/L)\bar{c}_1\hat{q}_1 \tag{1.11}$$

$$\dot{\hat{u}}_1\cos(\bar{q}_1 - \bar{q}_2) + 2\dot{\hat{u}}_2 = \Omega^2\left[\bar{c}_1\bar{c}_2\hat{q}_1 - (\bar{s}_1\bar{s}_2 - 2\bar{c}_2^2 + 2\bar{s}_2^2)\hat{q}_2\right] + [k/(ML^2)](\hat{q}_1 - \hat{q}_2) - 2(g/L)\bar{c}_2\hat{q}_2 \tag{1.12}$$

which, being linear, can be solved either in closed form or by numerical integration. Using the same parameter values and initial conditions as before, one obtains the curves labeled $\hat{q}_1$ and $\hat{q}_2$ in Fig. 3, as well as the entries in columns 5 and 6 of Table 1; and, evaluating G after replacing u_i with $\hat{u}_i$ and q_i with $\bar{q}_i + \hat{q}_i$ ($i = 1, 2$) in equation (1.6), one is led to the curve labeled G in Fig. 3, and to the numerical values recorded in column 7 of Table 1. Both show that G fails to remain constant. However, this does not mean that $\hat{u}_i$ and $\hat{q}_i$ ($i = 1, 2$) have been evaluated incorrectly. It simply indicates that equation (1.6) is not an integral of equations (1.9)–(1.12). To find such an integral, one may proceed as follows. In equation (1.6), after replacing u_i with $\hat{u}_i$ and q_i with $\bar{q}_i + \hat{q}_i$ ($i = 1, 2$), expand each trigonometric function of $\hat{q}_i$ in a Maclaurin series, and then drop all terms of third or higher degree in $\hat{u}_i$ and/or $\hat{q}_i$ ($i = 1, 2$). Denote the resulting function of $\hat{u}_i$, $\hat{q}_i$, and $\bar{q}_i$ ($i = 1, 2$) by $\tilde{G}$. Then the equation

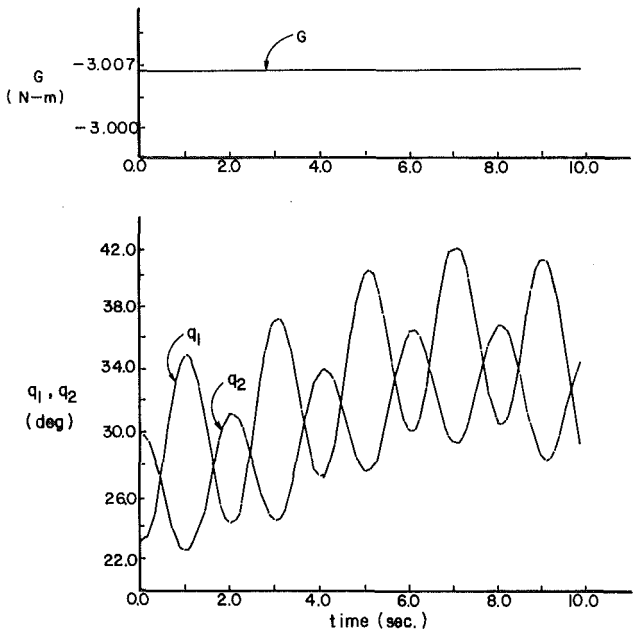


Fig. 2 A solution of the differential equations

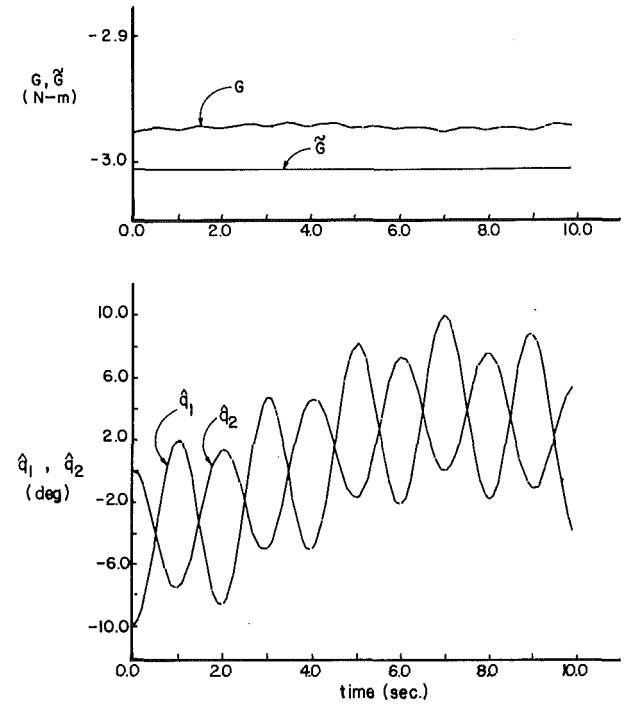


Fig. 3 A solution of the linearized differential equations

$$\tilde{G} = \tilde{C} \tag{1.13}$$

where $\tilde{C}$ is a constant, is an integral of equations (1.9)–(1.12). The validity of this contention is confirmed by the plot of $\tilde{G}$ versus t, shown in Fig. 3, and by the numerical value of $\tilde{G}$ at various instants of time, recorded in column 8 of Table 1.

The foregoing observations lead rather naturally to the following question: can one *always* construct an integral of a set of linearized equations of motion by proceeding as in the above example? It is one of the objectives of this paper to answer this question. Before doing so, however, let us briefly examine the following situation. Suppose that the spring connecting R_1 and R_2 is so stiff that one may expect q_1 to be nearly equal to q_2 throughout all motions of S. Under these cir-

cumstances, one may wish to replace u_1 with $u_2 + \hat{u}_1$, and q_1 with $q_2 + \hat{q}_1$, in equations (1.1)-(1.4), and then linearize in $\hat{u}_1$, $\hat{q}_1$, and their time-derivatives, which leads to a set of four partially linearized, but nevertheless nonlinear, differential equations; and, to test the validity of a numerical integration of these equations, one can attempt to make corresponding substitutions in equation (1.6), drop terms of degree three or higher in $\hat{u}_1$ and $\hat{q}_1$, and then see whether or not the function $\tilde{G}$ thus formed remains constant as time varies. Results of such a numerical integration are recorded in columns 9 and 10 of Table 1, and the associated values of $\tilde{G}$ appear in column 11, which reveals that $\tilde{G} = \tilde{C}$, a constant, is not an integral of the set of differential equations under consideration. Thus, one is led to wonder whether or not it is possible to derive an integral of a set of partially linearized equations of motion from an integral of the associated set of nonlinear equations of motion and, if so, how one might proceed in order to attain this goal. In Part I we answer both this question and the one raised earlier for a class of differential equations arising in dynamical analyses. We state and prove three theorems in the sections that follow, and illustrate the use of each theorem by means of an example. In Part II we apply the theorems to differential equations of motion of mechanical systems.

2 Total Linearization

Let $x(t)$ denote an $n \times 1$ matrix whose elements, $x_1, \ldots, x_n$, are functions of a scalar, independent variable t, and let $F(x, t)$ stand for an $n \times 1$ matrix whose elements $F_1, \ldots, F_n$ are functions of $x_1, \ldots, x_n$ and t such that $F(0, t)$, denoted by $\bar{F}$, vanishes for all t, that is,

$$\bar{F} \triangleq F(0,t) = 0 \tag{2.1}$$

Furthermore, require x to satisfy the differential equation

$$\dot{x} = F \tag{2.2}$$

where the dot denotes differentiation with respect to t; and let $G(x, t)$ denote a scalar function of $x_1, \ldots, x_n$ and t such that, whenever x satisfies equation (2.2), then

$$G = C, \quad \text{a constant} \tag{2.3}$$

Under these circumstances, equation (2.3) is called an *integral* of equation (2.2), and G is referred to as an *integral generating function* or, for short, an IGF, for equation (2.2).

To produce what we call the totally linearized form of equation (2.2), we introduce $\hat{x}(t)$, an $n \times 1$ matrix whose elements $\hat{x}_1, \ldots, \hat{x}_n$ are functions of t; define an $n \times n$ matrix $\partial F/\partial x$ as

$$\frac{\partial F}{\partial x} \triangleq \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial F_n}{\partial x_1} & \cdots & \dfrac{\partial F_n}{\partial x_n} \end{bmatrix} \tag{2.4}$$

and use an overbar to denote evaluation of $\partial F/\partial x$ at $x = 0$ [as in equation (2.1)]. The *totally linearized* form of equation (2.2) is, by definition, the equation

$$\dot{\hat{x}} = \frac{\overline{\partial F}}{\partial x}\,\hat{x} \tag{2.5}$$

Finally, we let $\hat{G}(\hat{x}, t)$ denote the scalar function defined as

$$\hat{G}(\hat{x},t) \triangleq \bar{G} + \frac{\overline{\partial G}}{\partial x}\,\hat{x} \tag{2.6}$$

where $\overline{\partial G}/\partial x$ is the value at $x = 0$ of the $n \times 1$ matrix $\partial G/\partial x$ defined as

$$\frac{\partial G}{\partial x} \triangleq \left[\frac{\partial G}{\partial x_1} \cdots \frac{\partial G}{\partial x_n}\right] \tag{2.7}$$

$\hat{G}$ is called the *totally linearized form* of G.

The *total linearization theorem* asserts that $\hat{G}$ is an IGF for equation (2.5) if G is an IFG for equation (2.2). In other words, $\hat{G} = \hat{C}$, a constant, is an integral of equation (2.5) if equation (2.3) is an integral of equation (2.2).

To prove this theorem, we begin by differentiating G with respect to t, which yields[1]

$$\frac{\partial G}{\partial x}\underset{(2.2)}{F} + \frac{\partial G}{\partial t}\underset{(2.3)}{=0} \tag{2.8}$$

Next, differentiating equation (2.6) with respect to t, we obtain

$$\begin{aligned} \dot{\hat{G}} &\underset{(2.6)}{=} \frac{\overline{\partial G}}{\partial t} + \frac{\overline{\partial^2 G}}{\partial t \partial x}\,\hat{x} + \frac{\overline{\partial G}}{\partial x}\,\dot{\hat{x}} \\ &= -\frac{\overline{\partial G}}{\partial x}\underset{(2.8)}{\bar{F}} + \frac{\overline{\partial^2 G}}{\partial t \partial x}\,\hat{x} + \frac{\overline{\partial G}}{\partial x}\,\frac{\overline{\partial F}}{\partial x}\,\hat{x}\underset{(2.5)}{} \\ &= 0\underset{(2.1)}{} + \left(\frac{\overline{\partial^2 G}}{\partial t \partial x} + \frac{\overline{\partial G}}{\partial x}\,\frac{\overline{\partial F}}{\partial x}\right)\hat{x} \end{aligned} \tag{2.9}$$

and, after defining $\partial^2 G/\partial x^2$ as

$$\frac{\partial^2 G}{\partial x^2} \triangleq \begin{bmatrix} \dfrac{\partial^2 G}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 G}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial^2 G}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 G}{\partial x_n \partial x_n} \end{bmatrix} \tag{2.10}$$

note that differentiation of equation (2.8) with respect to x and then setting x equal to zero produces

$$\frac{\overline{\partial^2 G}}{\partial x^2}\,\bar{F} + \frac{\overline{\partial G}}{\partial x}\,\frac{\overline{\partial F}}{\partial x} + \frac{\overline{\partial^2 G}}{\partial x \partial t}\underset{(2.8)}{=0} \tag{2.11}$$

The first term of this equation vanishes by virtue of equation (2.1), and the remaining two terms are the coefficient of $\hat{x}$ in equation (2.9). Consequently,

$$\dot{\hat{G}} \underset{(2.9)\ (2.11)}{=} 0 \tag{2.12}$$

which means that $\hat{G}$ is an IGF for equation (2.5).

3 Reduction

The *reduced form* of G is a function $\tilde{G}(\hat{x}, t)$ defined as

$$\tilde{G}(\hat{x},t) \triangleq \hat{G} + \frac{1}{2}\,\hat{x}^T\,\frac{\overline{\partial^2 G}}{\partial x^2}\,\hat{x} \tag{3.1}$$

where $\hat{G}$ and $\overline{\partial^2 G}/\partial x^2$ are given by equations (2.6) and (2.10), respectively. The $n \times n$ symmetric matrix whose ith row is $(\overline{\partial G}/\partial x)$ $(\overline{\partial^2 F}/\partial x_i \partial x)$ is denoted by $(\overline{\partial G}/\partial x)$ $(\overline{\partial^2 F}/\partial x^2)$; that is,

[1] Numbers beneath signs of equality or terms of an equation refer to corresponding equations.

$$\overline{\frac{\partial G}{\partial x}}\ \overline{\frac{\partial^2 F}{\partial x^2}} \triangleq \begin{bmatrix} \overline{\frac{\partial G}{\partial x}}\ \overline{\frac{\partial^2 F}{\partial x_1 \partial x}} \\ \hline \vdots \\ \hline \overline{\frac{\partial G}{\partial x}}\ \overline{\frac{\partial^2 F}{\partial x_n \partial x}} \end{bmatrix} \tag{3.2}$$

The *reduction theorem* asserts that, when G is an IGF for equation (2.2), then $\tilde{G}$ is an IGF for equation (2.5) if and only if

$$\overline{\frac{\partial G}{\partial x}}\ \overline{\frac{\partial^2 F}{\partial x^2}} = 0 \tag{3.3}$$

The proof of this theorem involves the first time-derivative of $\tilde{G}$ and the second derivative of equation (2.8) with respect to x. Now,

$$\dot{\tilde{G}} \underset{(3.1)}{=} \dot{\tilde{G}} + \frac{1}{2}\left(\dot{\hat{x}}^T \overline{\frac{\partial^2 G}{\partial x^2}}\,\hat{x} + \hat{x}^T \overline{\frac{\partial^3 G}{\partial t \partial x^2}}\,\hat{x} + \hat{x}^T \overline{\frac{\partial^2 G}{\partial x^2}}\,\dot{\hat{x}}\right)$$

$$\underset{(2.12)}{=} 0 + \frac{1}{2}\left(\hat{x}^T \underset{(2.5)}{\overline{\frac{\partial F^T}{\partial x}}}\ \overline{\frac{\partial^2 G}{\partial x^2}}\,\hat{x} + \hat{x}^T \overline{\frac{\partial^3 G}{\partial t \partial x^2}}\,\hat{x} + \hat{x}^T \overline{\frac{\partial^2 G}{\partial x^2}}\ \underset{(2.5)}{\overline{\frac{\partial F}{\partial x}}}\,\hat{x}\right) \tag{3.4}$$

or, equivalently,

$$\dot{\tilde{G}} = \frac{1}{2}\,\hat{x}^T\left(\overline{\frac{\partial F}{\partial x}}^T \overline{\frac{\partial^2 G}{\partial x^2}} + \overline{\frac{\partial^3 G}{\partial t \partial x^2}} + \overline{\frac{\partial^2 G}{\partial x^2}}\ \overline{\frac{\partial F}{\partial x}}\right)\hat{x} \tag{3.5}$$

Differentiating equation (2.8) with respect to x gives

$$F^T \frac{\partial^2 G}{\partial x^2} + \frac{\partial G}{\partial x}\,\frac{\partial F}{\partial x} + \underset{(2.8)}{\frac{\partial^2 G}{\partial x \partial t}} = 0 \tag{3.6}$$

and, setting $x = 0$ after differentiating this equation with respect to x, one obtains with the aid of equation (2.1)

$$\overline{\frac{\partial F}{\partial x}}^T \overline{\frac{\partial^2 G}{\partial x^2}} + \overline{\frac{\partial^2 G}{\partial x^2}}\ \overline{\frac{\partial F}{\partial x}} + \overline{\frac{\partial G}{\partial x}}\ \overline{\frac{\partial^2 F}{\partial x^2}} + \overline{\frac{\partial^3 G}{\partial x^2 \partial t}} = 0 \tag{3.7}$$

Comparing the left-hand member of equation (3.7) with the expression within the parentheses in equation (3.5), one sees that $\dot{\tilde{G}} = 0$, which means that $\tilde{G}$ is an IGF for equation (2.5), if and only if equation (3.3) is satisfied.

4 Partial Linearization

Let i be an integer smaller than n; define j as

$$j \triangleq n - i \tag{4.1}$$

and let y and z be an $i \times 1$ and a $j \times 1$ matrix, respectively, such that

$$x = \begin{bmatrix} y \\ -- \\ z \end{bmatrix} \tag{4.2}$$

Also, let $I(y, z, t)$ and $J(y, z, t)$ be an $i \times 1$ and a $j \times 1$ matrix, respectively, such that

$$F = \begin{bmatrix} I \\ -- \\ J \end{bmatrix} \tag{4.3}$$

and define $H(y, z, t)$ as

$$H(y,z,t) \triangleq G\left(\begin{bmatrix} y \\ -- \\ z \end{bmatrix}, t\right) \tag{4.4}$$

Finally, use an overbar to denote evaluation at $y = 0$, so that, for example, $\bar{H} = H(0, z, t)$; and let $\hat{y}(t)$ be an $i \times 1$ matrix. Then we refer to the equation

$$\begin{bmatrix} \dot{\hat{y}} \\ -- \\ \dot{z} \end{bmatrix} = \begin{bmatrix} \bar{I} + \overline{\frac{\partial I}{\partial y}}\,\hat{y} \\ \hline \bar{J} + \overline{\frac{\partial J}{\partial y}}\,\hat{y} \end{bmatrix} \tag{4.5}$$

and to the function $\hat{H}(\hat{y}, z, t)$ defined as

$$\hat{H}(\hat{y},z,t) \triangleq \bar{H} + \overline{\frac{\partial H}{\partial y}}\,\hat{y} \tag{4.6}$$

as a *partially linearized form* of equation (2.2) and a *partially linearized form* of H, respectively.

The *partial linearization theorem* asserts that, when G is an IGF for equation (2.2), then $\hat{H}$ is an IGF for equation (4.5) if and only if

$$\bar{I}^T \overline{\frac{\partial^2 H}{\partial y^2}} = 0 \tag{4.7}$$

and

$$\overline{\frac{\partial J}{\partial y}}^T \overline{\frac{\partial^2 H}{\partial z \partial y}} + \left(\overline{\frac{\partial J}{\partial y}}^T \overline{\frac{\partial^2 H}{\partial z \partial y}}\right)^T = 0 \tag{4.8}$$

To prove this, we note first that

$$\dot{\hat{H}} \underset{(4.6)}{=} \overline{\frac{\partial H}{\partial z}}\,\dot{z} + \overline{\frac{\partial H}{\partial t}} + \left(\dot{z}^T \overline{\frac{\partial^2 H}{\partial z \partial y}} + \overline{\frac{\partial^2 H}{\partial t \partial y}}\right)\hat{y} + \overline{\frac{\partial H}{\partial y}}\,\dot{\hat{y}} \tag{4.9}$$

$$= \overline{\frac{\partial H}{\partial z}}\left(\bar{J} + \underset{(4.5)}{\overline{\frac{\partial J}{\partial y}}}\,\hat{y}\right) + \overline{\frac{\partial H}{\partial t}} + \left[\left(\bar{J}^T + \hat{y}^T \underset{(4.5)}{\overline{\frac{\partial J}{\partial y}}}^T\right)\overline{\frac{\partial^2 H}{\partial z \partial y}} + \overline{\frac{\partial^2 H}{\partial t \partial y}}\right]\hat{y}$$

$$+ \overline{\frac{\partial H}{\partial y}}\left(\bar{I} + \underset{(4.5)}{\overline{\frac{\partial I}{\partial y}}}\,\hat{y}\right) \tag{4.10}$$

so that $\dot{\hat{H}} = 0$ for all $\hat{y}$ if and only if

$$\overline{\frac{\partial H}{\partial z}}\,\bar{J} + \overline{\frac{\partial H}{\partial t}} + \overline{\frac{\partial H}{\partial y}}\,\bar{I} = 0 \tag{4.11}$$

and

$$\overline{\frac{\partial H}{\partial z}}\ \overline{\frac{\partial J}{\partial y}} + \bar{J}^T \overline{\frac{\partial^2 H}{\partial z \partial y}} + \overline{\frac{\partial^2 H}{\partial t \partial y}} + \overline{\frac{\partial H}{\partial y}}\ \overline{\frac{\partial I}{\partial y}} = 0 \tag{4.12}$$

and $\overline{\frac{\partial J}{\partial y}}^T \overline{\frac{\partial^2 H}{\partial z \partial y}}$ is skew-symmetric, that is,

$$\overline{\frac{\partial J}{\partial y}}^T \overline{\frac{\partial^2 H}{\partial z \partial y}} + \left(\overline{\frac{\partial J}{\partial y}}^T \overline{\frac{\partial^2 H}{\partial z \partial y}}\right)^T = 0 \tag{4.13}$$

We shall now show that equation (4.11) is satisifed whenever G satisfies equation (2.8), and that equation (4.12) may be replaced with equation (4.7). To these ends, we observe that

$$\frac{\partial H}{\partial y} \underset{(4.4)}{=} \left[\underset{(4.2)}{\frac{\partial G}{\partial x_1}} \quad \cdots \quad \frac{\partial G}{\partial x_i}\right] \tag{4.14}$$

and

$$\frac{\partial H}{\partial z} \underset{(4.4)}{=} \left[\frac{\partial G}{\partial x_{i+1}} \underset{(4.2)}{\cdots} \frac{\partial G}{\partial x_n} \right] \tag{4.15}$$

whence

$$\frac{\partial G}{\partial x} \underset{(2.7)}{=} \left[\frac{\partial H}{\partial y} \underset{(4.14)}{\vdots} \frac{\partial H}{\partial z} \underset{(4.15)}{} \right] \tag{4.16}$$

Hence, equation (2.8) can be rewritten

$$\left[\frac{\partial H}{\partial y} \underset{(4.16)}{\vdots} \frac{\partial H}{\partial z} \right] \left[\frac{I}{J} \underset{(4.3)}{} \right] + \frac{\partial H}{\partial t} \underset{(4.4)}{} = 0 \tag{4.17}$$

Now, when $y = 0$, this becomes equation (4.11), which means that the latter is satisfied whenever G satisfies equation (2.8). Additionally, differentiation of equation (4.17) with respect to y yields

$$\left\{ \left[\frac{\overline{\partial^2 H}}{\partial y^2} \vdots \frac{\overline{\partial^2 H}}{\partial y \partial z} \right] \left[\frac{\bar{I}}{\bar{I}} \right] \right\}^T + \left[\frac{\overline{\partial H}}{\partial y} \vdots \frac{\overline{\partial H}}{\partial z} \right] \left[\frac{\overline{\partial I}}{\partial y} \atop \overline{\partial J} \right] + \frac{\overline{\partial^2 H}}{\partial y \partial t} = 0 \tag{4.18}$$

or, equivalently,

$$\bar{I}^T \frac{\overline{\partial^2 H}}{\partial y^2} + \bar{J}^T \frac{\overline{\partial^2 H}}{\partial z \partial y} + \frac{\overline{\partial H}}{\partial y} \frac{\overline{\partial I}}{\partial y} + \frac{\overline{\partial H}}{\partial z} \frac{\overline{\partial J}}{\partial y} + \frac{\overline{\partial^2 H}}{\partial y \partial t} = 0 \tag{4.19}$$

which makes it possible to replace equation (4.12) with equation (4.7).

5 Examples

The function $G(x_1, \ldots, x_6)$ defined as

$$G \triangleq (x_1+1)^2 + x_2^2 + x_3^2 + 2(x_4+4)^2 + 3x_5^2 + 4x_6^2 \tag{5.1}$$

is an IGF for the set of differential equations

$$\dot{x}_1 = x_2 x_6 - x_3 x_5, \quad \dot{x}_2 = (x_4+4)x_3 - (x_1+1)x_6 \tag{5.2}$$

$$\dot{x}_3 = (x_1+1)x_5 - (x_4+4)x_2 \quad , \quad \dot{x}_4 = -x_5 x_6/2 \tag{5.3}$$

$$\dot{x}_5 = 2(x_4+4)x_6/3 \quad , \quad \dot{x}_6 = -(x_4+4)x_5/4 \tag{5.4}$$

and the totally linearized form of these equations is the set of equations

$$\dot{\hat{x}}_1 = 0 \quad , \quad \dot{\hat{x}}_2 = 4\hat{x}_3 - \hat{x}_6 \quad , \quad \dot{\hat{x}}_3 = \hat{x}_5 - 4\hat{x}_2 \tag{5.5}$$

$$\dot{\hat{x}}_4 = 0 \quad , \quad \dot{\hat{x}}_5 = 8\hat{x}_6/3 \quad , \quad \dot{\hat{x}}_6 = -\hat{x}_5 \tag{5.6}$$

while the totally linearized form of G is

$$\hat{G} = 2\hat{x}_1 + 16\hat{x}_4 + 33 \tag{5.7}$$

In conformity with the total linearization theorem, $\hat{G}$ is an IGF for equations (5.5) and (5.6), a fact that is easily established by noting that

$$\dot{\hat{G}} \underset{(5.7)}{=} 2\dot{\hat{x}}_1 + 16\dot{\hat{x}}_4 = \underset{(5.5)}{0} + \underset{(5.6)}{0} \tag{5.8}$$

Since G as given in equation (5.1) contains no terms of degree higher than the second, the reduced form of G is given by

$$\tilde{G} = (\hat{x}_1+1)^2 + \hat{x}_2^2 + \hat{x}_3^2 + 2(\hat{x}_4+4)^2 + 3\hat{x}_5^2 + 4\hat{x}_6^2 \tag{5.9}$$

Referring to equation (3.2), one obtains

$$\frac{\overline{\partial G}}{\partial x} \frac{\overline{\partial^2 F}}{\partial x^2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & -8 \\ 0 & 2 & 0 & 0 & -8 & 0 \end{bmatrix} \tag{5.10}$$

Hence, the requirement imposed by equation (3.3) is violated, and the reduction theorem leads one to conclude that $\tilde{G}$ is not an IGF for equations (5.5) and (5.6).

Finally, consider the set of four equations

$$\dot{x}_1 = x_2 \quad , \quad \dot{x}_2 = -\sin x_1 (x_2^2 \cos x_1 + 2)/(2 - \cos^2 x_1) \tag{5.11}$$

$$\dot{x}_3 = x_3 x_4 \quad , \quad \dot{x}_4 = \sin x_1 (x_2^2 + \cos x_1)/(2 - \cos^2 x_1) \tag{5.12}$$

and suppose that these are to be linearized in x_1 and x_2, but not in x_3 and x_4. Referring to equations (4.2) and (4.3), and noting that $i = j = 2$, one finds that $y_1 = x_1$, $y_2 = x_2$, $z_1 = x_3$, $z_2 = x_4$, and

$$I = \begin{bmatrix} y_2 \\ -\sin y_1 (y_2^2 \cos y_1 + 2)/(2 - \cos^2 y_1) \end{bmatrix} \tag{5.13}$$

$$J = \begin{bmatrix} z_1 z_2 \\ \sin y_1 (y_2^2 + \cos y_1)/(2 - \cos^2 y_1) \end{bmatrix} \tag{5.14}$$

Furthermore, G, defined as $G \triangleq x_2 + 2x_4$ is an IGF for equations (5.11) and (5.12); H, formed in accordance with equation (4.4), is given by $H = y_2 \cos y_1 + 2z_2$; and H, I, and J satisfy equations (4.7) and (4.8). In accordance with the partial linearization theorem, $\hat{H}$, here given by $\hat{H} = \hat{y}_2 + 2z_2$, must thus be an IGF for the partially linearized form of equations (5.11) and (5.12), that is, for the equations

$$\dot{\hat{y}}_1 = \hat{y}_2, \quad \dot{\hat{y}}_2 = -2\hat{y}_1, \quad \dot{z}_1 = z_1 z_2, \quad \dot{z}_2 = \hat{y}_1 \tag{5.15}$$

And, indeed,

$$\dot{\hat{H}} = \dot{\hat{y}}_2 + 2\dot{z}_2 \underset{(5.15)}{=} -2\hat{y}_1 + 2\hat{y}_1 = 0 \tag{5.16}$$

T. R. Kane
Professor of Applied Mechanics,
Fellow ASME

S. Djerassi

Stanford University
Stanford, CA 94305

Integrals of Linearized Differential Equations of Motion of Mechanical Systems; Part II: Linearized Equations of Motion

Theorems derived in Part I are here applied to differential equations of motion of mechanical systems. The theorems are reformulated in terms of variables appearing in dynamical equations of motion, and their use is illustrated by means of an example.

1 Introduction

Systems of linear differential equations deduced by a process of linearization from the nonlinear differential equations governing the behavior of a mechanical system play an important role in many engineering analyses (see, e.g., Kailath, 1980). This paper shows how one can construct integrals for such systems of linear differential equations of motion by making use of available integrals of the associated nonlinear equations of motion.

The present work draws heavily upon Part I. Indeed, it consists of the application of material from Part I to the field of dynamics. This leads to theorems which furnish analytical tools especially well suited for dealing with motions of mechanical systems.

In Part I it is shown how one can construct integrals for linearized forms of the differential equation (I2.2)[1] when this equation possesses an integral. In what follows, we show that the differential equations of motion of any simple, nonholonomic system can be cast in the form of equation (I2.2), and we correlate the variables and functions appearing in the equations of motion to x and F in equation (I2.2). Next, we appeal to the three theorems of Part I to formulate theorems involving the variables that appear in the differential equations of motion. Finally, we illustrate the use of this material with an example.

2 Equations of Motion

Consider a simple, nonholonomic system S (Kane, 1985) possessing P degrees of freedom, N ($N \geq P$) generalized coordinates $q_1, \ldots, q_N$, and P independent generalized speeds $u_1, \ldots, u_P$. A set of equations of motion of S consists of N *kinematical* equations and P *dynamical* equations. To show that these can be cast in the form of equation (I2.2), we refer to Kane (1985), Sec. 2.12, where N (not necessarily independent) generalized speeds $u_1, \ldots, u_N$ are defined as linear combinations of the time derivatives of N generalized coordinates. Solving for the latter, we have [Kane, 1985, equation (2.14.5)]

$$\dot{q}_s = \sum_{r=1}^{N} W_{sr} u_r + X_s \quad (s = 1, \ldots, N) \tag{2.1}$$

where W_{sr} and X_s are functions of $q_1, \ldots, q_N$ and t. Simple, nonholonomic constraints are introduced as m linear relations of the form [Kane, 1985, Eq. (2.13.1)]

$$u_k = \sum_{r=1}^{P} A_{kr} u_r + B_k \quad (k = P+1, \ldots, N) \tag{2.2}$$

where $P \triangleq N - m$. Here A_{kr} and B_k are functions of $q_1, \ldots, q_N$ and t; and $u_1, \ldots, u_P$ are *independent* generalized speeds. Substitution from equations (2.2) into equations (2.1) leads to expressions relating $\dot{q}_1, \ldots, \dot{q}_N$ to $u_1, \ldots, u_P$. Specifically, after defining D_{sr} and E_s as

$$\left.\begin{aligned} D_{sr} &\triangleq W_{sr} + \sum_{k=P+1}^{N} W_{sk} A_{kr} \\ E_s &\triangleq X_s + \sum_{k=P+1}^{N} W_{sk} B_k \end{aligned}\right\} \quad (s = 1, \ldots, N;\ r = 1, \ldots, P) \tag{2.3}$$

we obtain from equations (2.1)–(2.3)

$$\dot{q}_s = \sum_{r=1}^{P} D_{sr} u_r + E_s \quad (s = 1, \ldots, N) \tag{2.4}$$

These are the *kinematical* equations, and, since D_{sr} and E_s are functions of $q_1, \ldots, q_N$ and t, they have the form of equation (I2.2).

[1] Equations numbers preceded by "I" refer to equations numbered correspondingly in Part I.

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

Discussion of this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August 22, 1986; final revision, December 1, 1986.

The *dynamical* equations can be deduced from the relationship

$$K_r + K_r^* = 0 \quad (r = 1, \ldots, P) \tag{2.5}$$

where K_r is the rth nonholonomic generalized active force, and K_r^* is the rth nonholonomic generalized inertia force. By definition (Kane, 1985, Sec. 4.11), the latter is given by

$$K_r^* = -\sum_{i=1}^{\nu} m_i \mathbf{a}^{P_i} \cdot \tilde{\mathbf{v}}_r^{P_i} \quad (r = 1, \ldots, P) \tag{2.6}$$

where $\mathbf{a}^{P_i}$ is the acceleration of the ith particle P_i of S, m_i is the mass of P_i, $\tilde{\mathbf{v}}_r^{P_i}$ is the rth partial velocity of P_i (Kane, 1985, Sec. 2.14), and ν is the number of particles in S. Now, $\mathbf{a}^{P_i}$ is given by [Kane, 1985, equation (5.6.10)]

$$\mathbf{a}^{P_i} = \sum_{s=1}^{P} (\dot{\tilde{\mathbf{v}}}_s^{P_i} u_s + \tilde{\mathbf{v}}_s^{P_i} \dot{u}_s) + \dot{\tilde{\mathbf{v}}}_t^{P_i} \quad (i = 1, \ldots, \nu) \tag{2.7}$$

where $\tilde{\mathbf{v}}_s^{P_i}$ and $\tilde{\mathbf{v}}_t^{P_i}$ are functions of $q_1, \ldots, q_N$ and t, and dots over $\tilde{\mathbf{v}}_s^{P_i}$ and $\tilde{\mathbf{v}}_t^{P_i}$ denote time differentiation. Hence, from equations (2.6) and (2.7),

$$K_r^* = -\sum_{s=1}^{P} M_{rs}\dot{u}_s + L_r(q_1, \ldots, q_N, u_1, \ldots, u_P, t) \quad (r = 1, \ldots, P) \tag{2.8}$$

where L_r and M_{rs} are defined as

$$L_r \triangleq -\sum_{i=1}^{\nu}\sum_{s=1}^{P} m_i \dot{\tilde{\mathbf{v}}}_s^{P_i} \cdot \tilde{\mathbf{v}}_r^{P_i} u_s - \sum_{i=1}^{\nu} m_i \dot{\tilde{\mathbf{v}}}_t^{P_i} \cdot \tilde{\mathbf{v}}_r^{P_i} \quad (r = 1, \ldots, P) \tag{2.9}$$

and

$$M_{rs} \triangleq \sum_{i=1}^{\nu} m_i \tilde{\mathbf{v}}_r^{P_i} \cdot \tilde{\mathbf{v}}_s^{P_i} \quad (r,s = 1, \ldots, P) \tag{2.10}$$

so that

$$M_{rs} = M_{sr}; \quad \mathrm{Det}(M_{rs}) > 0 \quad (r,s = 1, \ldots, P) \tag{2.11}$$

The generalized active force K_r is defined as [Kane, 1985, equation (4.4.1)]

$$K_r \triangleq \sum_{i=1}^{\nu} \tilde{\mathbf{v}}_r^{P_i} \cdot \mathbf{R}_i \quad (r = 1, \ldots, P) \tag{2.12}$$

where $\mathbf{R}_i$ is the resultant of all contact and body forces acting on P_i. Treating $\mathbf{R}_i$ as a function of $q_1, \ldots, q_N, u_1, \ldots, u_P$ and t, we thus find that K_r can be expressed as a function of the same variables; and substitution from equations (2.12) and (2.8) into equations (2.5) yields

$$\sum_{s=1}^{P} M_{rs}\dot{u}_s = L_r - K_r \quad (r = 1, \ldots, P) \tag{2.13}$$

which can be solved for $\dot{u}_r$ $(r = 1, \ldots, P)$ to obtain

$$\dot{u}_r = Y_r(q_1, \ldots, q_N, u_1, \ldots, u_p, t) \quad (r = 1, \ldots, P) \tag{2.14}$$

These are the *dynamical* equations cast into the form of equation (I2.2).

Before we can apply the theorems stated in Part I, we must identify the quantities x, F, and G that appear in equations (I2.2) and (I2.3). To this end, we first consider the requirement imposed by equation (I2.1) and define a *nominal* motion as one such that

$$q_s = \bar{q}_s \text{ , a constant} \quad (s = 1, \ldots, N) \tag{2.15}$$

$$u_r = 0 \quad (r = 1, \ldots, P) \tag{2.16}$$

Such a motion is possible if and only if[2]

$$E_s(\bar{q}_1, \ldots, \bar{q}_N, t) \underset{(2.4)(2.15)(2.16)}{=} 0 \quad (s = 1, \ldots, N) \tag{2.17}$$

and

$$Y_r(\bar{q}_1, \ldots, \bar{q}_N, \underbrace{0, \ldots, 0}_{P}, t) \underset{(2.4)(2.15)(2.16)}{=} 0 \quad (r = 1, \ldots, P) \tag{2.18}$$

Accordingly, we let

$$x_s \triangleq q_s - \bar{q}_s \quad (s = 1, \ldots, N) \tag{2.19}$$

$$x_{N+r} \triangleq u_r \quad (r = 1, \ldots, P) \tag{2.20}$$

and define $\mathfrak{D}_{sr}$, $\mathfrak{E}_s$, and $\mathfrak{Y}_r$ as

$$\mathfrak{D}_{sr}(x_1, \ldots, x_N, t) \triangleq D_{sr}(x_1 + \bar{q}_1, \ldots, x_N + \bar{q}_N, t) \tag{2.21}$$
$$(2.19)$$

$$\mathfrak{E}_s(x_1, \ldots, x_N, t) \triangleq E_s(x_1 + \bar{q}_1, \ldots, x_N + \bar{q}_N, t) \tag{2.22}$$
$$(2.19)$$

$$\mathfrak{Y}_r(x_1, \ldots, x_{N+P}, t) \triangleq Y_r(x_1 + \bar{q}_1, x_N + \bar{q}_N, x_{N+1}, \ldots, x_{N+P}, t) \tag{2.23}$$
$$(2.19) \qquad (2.20)$$

whereupon the kinematical and dynamical equations lead to

$$\dot{x}_s \underset{(2.19)}{=} \dot{q}_s \underset{(2.4)}{=} \sum_{r=1}^{P} \mathfrak{D}_{sr} x_{N+r} + \mathfrak{E}_s \quad (s = 1, \ldots, N) \tag{2.24}$$
$$(2.21)(2.20) \qquad (2.22)$$

$$\dot{x}_{N+r} \underset{(2.20)}{=} \dot{u}_r \underset{(2.14)}{=} \mathfrak{Y}_r \underset{(2.23)}{} \quad (r = 1, \ldots, P) \tag{2.25}$$

Hence, letting

$$F_s(x_1, \ldots, x_{N+P}) \triangleq \sum_{r=1}^{P} \mathfrak{D}_{sr} x_{N+r} + \mathfrak{E}_s \quad (s = 1, \ldots, N) \tag{2.26}$$

$$F_{N+r}(x_1, \ldots, x_{N+P}) \triangleq \mathfrak{Y}_r \quad (r = 1, \ldots, P) \tag{2.27}$$

we have

$$\dot{x}_s \underset{(2.24)}{=} F_s \underset{(2.26)}{} \quad (s = 1, \ldots, N) \tag{2.28}$$

$$\dot{x}_{N+r} \underset{(2.25)}{=} F_{N+r} \underset{(2.27)}{} \quad (r = 1, \ldots, P) \tag{2.29}$$

with

$$\bar{F}_s \triangleq F_s(0, \ldots, 0, t) \underset{(2.26)}{=} 0 + E_s(\bar{q}_1, \ldots, \bar{q}_N, t) \underset{(2.22)}{} \underset{(2.17)}{=} 0$$

$$(s = 1, \ldots, N) \tag{2.30}$$

and

$$\bar{F}_{N+r} \triangleq F_{N+r}(0, \ldots, 0, t) \underset{(2.27)}{=}$$

$$= Y_r(\bar{q}_1, \ldots, \bar{q}_N, \underbrace{0 \ldots, 0}_{P}, t) \underset{(2.18)}{=} 0 \quad (r = 1, \ldots, P) \tag{2.31}$$
$$(2.23)$$

[2] Numbers beneath signs of equality or terms of an equation refer to corresponding equations.

Finally, we note that, when $I(q_1, \ldots, q_N, u_1, \ldots, u_p, t)$ is an IGF for equations (2.4) and (2.14), then $G(x_1, \ldots, x_{N+P}, t)$, defined as

$$G(x_1, \ldots, x_{N+P}, t) \underset{(2.19)}{\triangleq} \mathcal{I}(x_1 + \bar{q}_1, \ldots, x_N + \bar{q}_N, x_{N+1}, \ldots, x_{N+P}, t) \underset{(2.20)}{} \tag{2.32}$$

is an IGF for equations (2.28) and (2.29). Now we are in a position to reformulate the three theorems of Part I in terms of the variables and functions appearing in the equations of motion.

3 Total Linearization and Reduction

Considering motions such that

$$q_s = \bar{q}_s + \hat{q}_s \qquad (s = 1, \ldots, N) \tag{3.1}$$

$$u_r = \hat{u}_r \qquad (r = 1, \ldots, P) \tag{3.2}$$

where $\hat{q}_s$ and $\hat{u}_r$ are time-dependent perturbations, we substitute $q_1, \ldots, q_N, u_1, \ldots, u_p$ from equation (3.1) and (3.2) into equations (2.4) and (2.14), expand all functions of the perturbations in power series, and drop all terms of second and higher degree to arrive at the *totally linearized* equations of motion

$$\dot{\hat{q}} = \sum_{r=1}^{P} \bar{D}_{sr} \hat{u}_r + \sum_{i=1}^{N} \frac{\overline{\partial E_s}}{\partial q_i} \hat{q}_i \qquad (s = 1, \ldots, N) \tag{3.3}$$

$$\dot{\hat{u}}_r = \sum_{s=1}^{N} \frac{\overline{\partial Y_r}}{\partial q_s} \hat{q}_s + \sum_{k=1}^{P} \frac{\overline{\partial Y_r}}{\partial u_k} \hat{u}_k \qquad (r = 1, \ldots, P) \tag{3.4}$$

where, as before,

$$\bar{E}_s = E_s(\bar{q}_1, \ldots, \bar{q}_N, t) \underset{(2.17)}{=} 0 \quad (s = 1, \ldots, N) \tag{3.5}$$

and

$$\bar{Y}_r = Y_r(\bar{q}_1, \ldots, \bar{q}_N, 0, \ldots, 0, t) \underset{(2.18)}{=} 0 \quad (r = 1, \ldots, P) \tag{3.6}$$

In view of equations (I2.6), (I3.1), (2.19), and (2.20), the linearized form of I and the reduced form of I are, respectively,

$$\hat{I} = \bar{I} + \sum_{s=1}^{N} \frac{\overline{\partial I}}{\partial q_s} \hat{q}_s + \sum_{r=1}^{P} \frac{\overline{\partial I}}{\partial u_r} \hat{u}_r \tag{3.7}$$

$$\tilde{I} = \hat{I} + \frac{1}{2}\left(\sum_{r=1}^{N} \sum_{s=1}^{N} \frac{\overline{\partial^2 I}}{\partial q_r \partial q_s} \hat{q}_r \hat{q}_s + 2 \sum_{r=1}^{P} \sum_{s=1}^{N} \frac{\overline{\partial^2 I}}{\partial q_s \partial u_r} \hat{q}_s \hat{u}_r + \sum_{r=1}^{P} \sum_{s=1}^{P} \frac{\overline{\partial^2 I}}{\partial u_r \partial u_s} \hat{u}_r \hat{u}_s \right) \tag{3.8}$$

Moreover, if $\hat{x}_1, \ldots, \hat{x}_{N+P}$ are defined as

$$\hat{x}_s \triangleq \hat{q}_s \qquad (s = 1, \ldots, N) \tag{3.9}$$

$$\hat{x}_{N+r} \triangleq \hat{u}_r \qquad (r = 1, \ldots, P) \tag{3.10}$$

then it follows from the definitions given in equations (2.21)–(2.13) and (2.26)–(2.27) that equations (3.3) and (3.4) can be represented by equation (I2.5) with $n = N + P$. Similarly, equations (3.7) and (3.8) correspond to equations (I2.6) and (I3.1), respectively. Consequently, the *total linearization theorem* of Part I, applied in conjunction with equations (2.4), (2.14), (3.3), and (3.4), leads to the following conclusion: when

$$I = C, \text{ a constant} \tag{3.11}$$

is an integral of equations (2.4) and (2.14), and when equations (3.5) and (3.6) are satisfied, then

$$\hat{I} = \hat{C}, \text{ a constant} \tag{3.12}$$

is an integral of equations (3.3) and (3.4).

The *reduction theorem* of Part I comes into play when H, the Hamiltonian of S, is an IGF for equations (2.4) and (2.14). To show this, we first recall that $T(q_1, \ldots, q_N, u_1, \ldots, u_P, t)$, the kinetic energy of S, is given by [Kane, 1985, equation (5.5.6)]

$$T = M(q_1, \ldots, q_N, t) + \sum_{r=1}^{P} M_r(q_1, \ldots, q_N, t) u_r + \frac{1}{2} \sum_{r=1}^{P} \sum_{s=1}^{P} M_{rs}(q_1, \ldots, q_N, t) u_r u_s \tag{3.13}$$

where M_{rs} is defined in equation (2.10), and that, if $V(q_1, \ldots, q_N, t)$ stands for the potential energy of S, then $H(q_1, \ldots, q_N, u_1, \ldots, u_p, t)$ can be written [Kane, 1985, equation (7.22)]

$$H(q_1, \ldots, q_N, u_1, \ldots, u_p, t) = V + \frac{1}{2} \sum_{r=1}^{P} \sum_{s=1}^{P} M_{rs} u_r u_s - M \tag{3.14}$$

Hence, when H is an IGF for equations (2.4) and (2.14), then it follows from equations (2.19) and (2.20) that $G(x_1, \ldots, x_{N+P}, t)$, defined as

$$G(x_1, \ldots, x_{N+P}, t) \triangleq H(x_1 + \bar{q}_1, \ldots, x_N + \bar{q}_N, x_{N+1}, \ldots, x_{N+P}, t) \tag{3.15}$$

is an IGF for equations (2.28) and (2.29). Next, we define $\mathcal{V}$, $\mathfrak{M}_{rs}$, and $\mathfrak{M}$ as

$$\mathcal{V}(x_1, \ldots, x_N, t) \underset{(2.19)}{\triangleq} V(x_1 + \bar{q}_1, \ldots, x_N + \bar{q}_N, t) \tag{3.16}$$

$$\mathfrak{M}_{rs}(x_1, \ldots, x_N, t) \underset{(2.19),(3.13)}{\triangleq} M_{rs}(x_1 + \bar{q}_1, \ldots, x_N + \bar{q}_N, t) \tag{3.17}$$

$$\mathfrak{M}(x_1, \ldots, x_N, t) \underset{(2.19),(3.13)}{\triangleq} M(x_1 + \bar{q}_1, \ldots, x_N + \bar{q}_N, t) \tag{3.18}$$

so that

$$G \underset{(3.15),(3.14)}{=} \mathcal{V} \underset{(3.16)}{} + \frac{1}{2} \sum_{r=1}^{P} \sum_{s=1}^{P} \mathfrak{M}_{rs} x_{N+r} x_{N+s} \underset{(3.17)(2.20)}{} - \mathfrak{M} \underset{(3.18)}{} \tag{3.19}$$

and

$$\frac{\overline{\partial G}}{\partial x} \underset{(I2.7)}{=} \left| \frac{\overline{\partial G}}{\partial x_1} \cdots \frac{\overline{\partial G}}{\partial x_N} \frac{\overline{\partial G}}{\partial x_{N+1}} \cdots \frac{\overline{\partial G}}{\partial x_{N+P}} \right|$$

$$\underset{(3.19)}{=} \left[\frac{\overline{\partial \mathcal{V}}}{\partial x_1} - \frac{\overline{\partial \mathfrak{M}}}{\partial x_1} \cdots \frac{\overline{\partial \mathcal{V}}}{\partial x_N} - \frac{\overline{\partial \mathfrak{M}}}{\partial x_N} \underbrace{0 \ldots 0}_{P} \right] \tag{3.20}$$

Moreover, $\partial^2 F / \partial x_s \partial x$ can be found as follows:

$$\frac{\partial F}{\partial x}=\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_N} & \frac{\partial F_1}{\partial x_{N+1}} & \cdots & \frac{\partial F_1}{\partial x_{N+P}} \\ \vdots & & & & & \vdots \\ \frac{\partial F_N}{\partial x_1} & \cdots & \frac{\partial F_N}{\partial x_N} & \frac{\partial F_N}{\partial x_{N+1}} & \cdots & \frac{\partial F_N}{\partial x_{N+P}} \\ & & & & & \vdots \\ \frac{\partial F_{N+1}}{\partial x_1} & \cdots & \frac{\partial F_{N+1}}{\partial x_N} & \frac{\partial F_{N+1}}{\partial x_{N+1}} & \cdots & \frac{\partial F_{N+1}}{\partial x_{N+P}} \\ \vdots & & & & & \vdots \\ \frac{\partial F_{N+P}}{\partial x_1} & \cdots & \frac{\partial F_{N+P}}{\partial x_N} & \frac{\partial F_{N+P}}{\partial x_{N+1}} & \cdots & \frac{\partial F_{N+P}}{\partial x_{N+P}} \end{bmatrix} \tag{3.21}$$

$$\frac{\partial F}{\partial x} \underset{(3.21),(2.26),(2.27)}{=} \begin{bmatrix} \sum_{r=1}^{P} \frac{\partial \mathfrak{D}_{1r}}{\partial x_1} x_{N+r} + \frac{\partial \mathcal{E}_1}{\partial x_1} & \cdots & \sum_{r=1}^{P} \frac{\partial \mathfrak{D}_{1r}}{\partial x_N} x_{N+r} + \frac{\partial \mathcal{E}}{\partial x_N} & \mathfrak{D}_{11} & \cdots & \mathfrak{D}_{1P} \\ \vdots & & & & & \\ \sum_{r=1}^{P} \frac{\partial \mathfrak{D}_{Nr}}{\partial x_1} x_{N+r} + \frac{\partial \mathcal{E}_N}{\partial x_1} & \cdots & \sum_{r=1}^{P} \frac{\partial \mathfrak{D}_{Nr}}{\partial x_N} x_{N+r} + \frac{\partial \mathcal{E}_N}{\partial x_N} & \mathfrak{D}_{N1} & \cdots & \mathfrak{D}_{NP} \\ \frac{\partial \mathcal{Y}_1}{\partial x_1} & \cdots & \frac{\partial \mathcal{Y}_1}{\partial x_N} & \frac{\partial \mathcal{Y}_1}{\partial x_{N+1}} & \cdots & \frac{\partial \mathcal{Y}_1}{\partial x_{N+P}} \\ \vdots & & & & & \\ \frac{\partial \mathcal{Y}_P}{\partial x_1} & \cdots & \frac{\partial \mathcal{Y}_\mathcal{P}}{\partial x_N} & \frac{\partial \mathcal{Y}_P}{\partial x_{N+1}} & \cdots & \frac{\partial \mathcal{Y}_P}{\partial x_{N+P}} \end{bmatrix} \tag{3.22}$$

$$\frac{\partial^2 F}{\partial x_s \partial x} \underset{(3.22)}{=} \begin{bmatrix} \sum_{r=1}^{P} \frac{\partial^2 \mathfrak{D}_{1r}}{\partial x_s \partial x_1} x_{N+r} + \frac{\partial^2 \mathcal{E}_1}{\partial x_s \partial x_1} & \cdots & \sum_{r=1}^{P} \frac{\partial^2 \mathfrak{D}_{1r}}{\partial x_s \partial x_N} x_{N+r} + \frac{\partial^2 \mathcal{E}_1}{\partial x_s \partial x_N} & \frac{\partial \mathfrak{D}_{11}}{\partial x_s} & \cdots & \frac{\partial \mathfrak{D}_{1P}}{\partial x_s} \\ \vdots & & & & & \\ \sum_{r=1}^{P} \frac{\partial^2 \mathfrak{D}_{\mathfrak{N}_r}}{\partial x_s \partial x_1} x_{N+r} + \frac{\partial^2 \mathcal{E}_N}{\partial x_s \partial x_1} & \cdots & \sum_{r=1}^{P} \frac{\partial^2 \mathfrak{D}_{Nr}}{\partial x_s \partial x_N} x_{N+r} + \frac{\partial^2 \mathcal{E}_N}{\partial x_s \partial x_N} & \frac{\partial \mathfrak{D}_{N1}}{\partial x_s} & \cdots & \frac{\partial \mathfrak{D}_{NP}}{\partial x_s} \\ \frac{\partial^2 \mathcal{Y}_1}{\partial x_s \partial x_1} & \cdots & \frac{\partial^2 \mathcal{Y}_1}{\partial x_s \partial x_N} & \frac{\partial^2 \mathcal{Y}_1}{\partial x_s \partial x_{N+1}} & \cdots & \frac{\partial^2 \mathcal{Y}_1}{\partial x_s \partial x_{N+P}} \\ \vdots & & & & & \\ \frac{\partial^2 \mathcal{Y}_\mathcal{P}}{\partial x_s \partial x_1} & \cdots & \frac{\partial^2 \mathcal{Y}_P}{\partial x_s \partial x_N} & \frac{\partial^2 \mathcal{Y}_P}{\partial x_s \partial x_{N+1}} & \cdots & \frac{\partial^2 \mathcal{Y}_P}{\partial x_s \partial x_{N+P}} \end{bmatrix} \tag{3.23}$$

$(s=1, \ldots, N)$

$$\frac{\partial^2 F}{\partial x_{N+r}\partial x} \underset{(3.22)}{=} \begin{bmatrix} \dfrac{\partial \mathfrak{D}_{1r}}{\partial x_1} & \cdots & \dfrac{\partial \mathfrak{D}_{1r}}{\partial x_N} & 0 \cdots 0 \\ \vdots & & & \\ \dfrac{\partial \mathfrak{D}_{Nr}}{\partial x_1} & \cdots & \dfrac{\partial \mathfrak{D}_{Nr}}{\partial x_N} & 0 \cdots 0 \\ \dfrac{\partial^2 \mathfrak{Y}_1}{\partial x_{N+1}\partial x_1} & \cdots & & \dfrac{\partial^2 \mathfrak{Y}_1}{\partial x_{N+r}\partial x_{N+P}} \\ \vdots & & & \\ \dfrac{\partial^2 \mathfrak{Y}_p}{\partial x_{N+r}\partial x_1} & \cdots & & \dfrac{\partial^2 \mathfrak{Y}_P}{\partial x_{N+r}\partial x_{N+P}} \end{bmatrix} \quad (r=1, \ldots, P) \tag{3.24}$$

Hence, substitution from equations (3.20), (3.23), and (3.24) into equation (I3.3) yields:

$$\left(\frac{\overline{\partial \mathfrak{V}}}{\partial x_1} - \frac{\overline{\partial \mathfrak{M}}}{\partial x_1}\right)\frac{\overline{\partial^2 \mathfrak{E}_1}}{\partial x_s \partial x_r} + \cdots + \left(\frac{\overline{\partial \mathfrak{V}}}{\partial x_N} - \frac{\overline{\partial \mathfrak{M}}}{\partial x_N}\right)\frac{\overline{\partial^2 \mathfrak{E}_N}}{\partial x_s \partial x_r} = 0 \quad (s,r=1, \ldots, N) \tag{3.25}$$

$$\left(\frac{\overline{\partial \mathfrak{V}}}{\partial x_1} - \frac{\overline{\partial \mathfrak{M}}}{\partial x_1}\right)\frac{\overline{\partial \mathfrak{D}_{1r}}}{\partial x_s} + \cdots + \left(\frac{\overline{\partial \mathfrak{V}}}{\partial x_N} - \frac{\overline{\partial \mathfrak{M}}}{\partial x_N}\right)\frac{\overline{\partial \mathfrak{D}_{Nr}}}{\partial x_s} = 0 \quad (r=1, \ldots, P; s=1, \ldots, N) \tag{3.26}$$

or, in view of equations (2.21), (2.22), (3.16), (3.18), and (2.19),

$$\sum_{k=1}^{N}\left(\frac{\overline{\partial V}}{\partial q_k} - \frac{\overline{\partial M}}{\partial q_k}\right)\frac{\overline{\partial^2 E_k}}{\partial q_r \partial q_s} = 0 \quad (s,r=1 \ldots, N) \tag{3.27}$$

$$\sum_{k=1}^{N}\left(\frac{\overline{\partial V}}{\partial q_k} - \frac{\overline{\partial M}}{\partial q_k}\right)\frac{\overline{\partial D}_{kr}}{\partial q_s} = 0 \quad (r=1, \ldots, P;\ s=1, \ldots, N) \tag{3.28}$$

and now we can apply the *reduction theorem* to assert that, when

$$H = C, \text{ a constant} \tag{3.29}$$

is an integral of equations (2.4) and (2.14), and when equations (3.5) and (3.6) are satisfied, then

$$\tilde{H} = \tilde{C}, \text{ a constant} \tag{3.30}$$

where $\tilde{H}$ is defined similarly to $\tilde{I}$ in equation (3.8) (with $\hat{H}$ analogous to $\hat{I}$ in equation (3.7)), is an integral of equations (3.3) and (3.4) if and only if equations (3.27) and (3.28) are satisfied.

A special case of frequent interest arises when S is a holonomic system, that is, $P = N$, and when

$$u_r = \dot{q}_r \qquad (r=1, \ldots, N) \tag{3.31}$$

Then, in accordance with equations (2.4),

$$D_{sr} = \delta_{sr}, E_s = 0 \qquad (s,r=1, \ldots, N) \tag{3.32}$$

and equations (3.27) and (3.28) are satisfied automatically. Under these circumstances, $\tilde{H}$ is guaranteed to be an IGF of the totally linearized equations of motion.

4 Partial Linearization

Defining Q_s as

$$Q_s \triangleq \sum_{r=1}^{P} D_{sr}u_r + E_s \qquad (s=1, \ldots, N) \tag{4.1}$$

and, accordingly, replacing equations (2.4) with

$$\dot{q}_s = Q_s \qquad (s=1, \ldots, N) \tag{4.2}$$

we consider motions such that

$$q_s = \bar{q}_s + \hat{q}_s \qquad (s=1, \ldots, i_N) \tag{4.3}$$

$$u_r = \hat{u}_r \ (r=1, \ldots, i_P) \tag{4.4}$$

where i_N and i_P are integers smaller than N and P, respectively. Next, we rewrite equations (2.4) and (2.14) as

$$\dot{q}_s = Q_s \ (s=1, \ldots, i_N) \tag{4.5}$$

$$\dot{u}_r = Y_r \ (r=1, \ldots, i_P) \tag{4.6}$$

$$\dot{q}_s = Q_s \ (s=i_N+1, \ldots, N) \tag{4.7}$$

$$\dot{u}_r = Y_r \ (r=i_P+1, \ldots, P) \tag{4.8}$$

and substitute $q_1, \ldots, q_{i_N}, u_1, \ldots, u_{i_P}$ from equations (4.3) and (4.4) into equations (4.5)–(4.8). Finally, we expand functions appearing in these equations in power series in all perturbations, and drop terms of second or higher degree to obtain the following partially linearized equations of motion:

$$\dot{\hat{q}}_s = \bar{Q}_s + \sum_{i=1}^{i_N}\frac{\overline{\partial Q_s}}{\partial q_i}\hat{q}_i + \sum_{i=1}^{i_P}\frac{\overline{\partial Q_s}}{\partial u_i}\hat{u}_i \ (s=1, \ldots, i_N) \tag{4.9}$$

$$\dot{\hat{u}}_r = \bar{Y}_r + \sum_{i=1}^{i_N}\frac{\overline{\partial Y_r}}{\partial q_i}\hat{q}_i + \sum_{i=1}^{i_P}\frac{\overline{\partial Y_r}}{\partial u_i}\hat{u}_i \ (r=1, \ldots, i_P) \tag{4.10}$$

$$\dot{q}_s = \bar{Q}_s + \sum_{i=1}^{i_N}\frac{\overline{\partial Q_s}}{\partial q_i}\hat{q}_i + \sum_{i=1}^{i_P}\frac{\overline{\partial Q_s}}{\partial u_i}\hat{u}_i \ (s=i_N+1, \ldots, N) \tag{4.11}$$

$$\dot{u}_r = \bar{Y}_r + \sum_{i=1}^{i_N}\frac{\overline{\partial Y_r}}{\partial q_i}\hat{q}_i + \sum_{i=1}^{i_P}\frac{\overline{\partial Y_r}}{\partial u_i}\hat{u}_i \ (r=i_P+1, \ldots, P) \tag{4.12}$$

where overbars denote evaluation at $q_s = \bar{q}_s$ $(s=1, \ldots, i_N)$ and $u_r = 0$ $(r=1, \ldots, i_p)$. The *partially linearized form* of I is

$$\hat{I} = \bar{I} + \sum_{i=1}^{i_N}\frac{\overline{\partial I}}{\partial q_i}\hat{q}_i + \sum_{i=1}^{i_P}\frac{\overline{\partial I}}{\partial u_i}\hat{u}_i \tag{4.13}$$

Moreover, if $\hat{x}_1, \ldots, \hat{x}_{i_N+i_P}$ are defined as

$$\hat{x}_s \triangleq \hat{q}_s \ (s=1, \ldots, i_N) \tag{4.14}$$

$$\hat{x}_{i_N+r} = \hat{u}_r \ (r=1, \ldots, i_P) \tag{4.15}$$

then it follows from the definitions given in equations (4.1), (2.21)–(2.23), and (2.26)–(2.27) that equations (4.9)–(4.12) can be represented by equation (I4.5) with

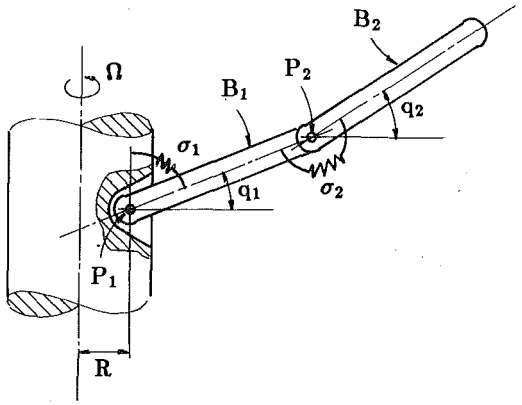


Fig. 1 System S

$$n \triangleq N+P \qquad i \triangleq i_N + i_P \tag{4.16}$$

$$y \triangleq [q_1, \ldots, q_{i_N} | u_1, \ldots, u_{i_p}]^T$$

$$z \triangleq [q_{i_N+1}, \ldots, q_N | u_{i_P+1}, \ldots, u_P]^T \tag{4.17}$$

$$I \triangleq [Q_1, \ldots, Q_{i_N} | Y_1, \ldots, Y_{i_P}]^T$$

$$J \triangleq [Q_{i_N+1}, \ldots, Q_N | Y_{i_P+1}, \ldots, Y_P]^T \tag{4.18}$$

Similarly, equation (4.13) corresponds to equation (I4.6). Consequently, the *partial linearization theorem* applied in conjunction with equations (4.2), (2.14), and (4.9)–(4.12), leads to the following conclusion: when

$$I = C, \text{ a constant} \tag{4.19}$$

is an integral of equations (4.2) and (2.14), then

$$\hat{I} = \hat{C}, \text{ a constant} \tag{4.20}$$

is an integral of equations (4.9)–(4.12), if and only if

$$[\bar{Q}_1 \ldots \bar{Q}_{i_N} | \bar{Y}_1 \ldots Y_{i_P}] \begin{bmatrix} \overline{\frac{\partial^2 I}{\partial q_1^2}} \cdots & \overline{\frac{\partial^2 I}{\partial q_1 \partial q_{i_N}}} & \overline{\frac{\partial^2 I}{\partial q_1 \partial u_1}} \cdots & \overline{\frac{\partial^2 I}{\partial q_1 \partial u_{i_P}}} \\ \vdots & & & \\ \overline{\frac{\partial^2 I}{\partial u_{i_P} \partial q_1}} \cdots & \overline{\frac{\partial^2 I}{\partial u_{i_P} \partial q_{i_N}}} & \overline{\frac{\partial^2 I}{\partial u_{i_P} \partial u_1}} \cdots & \overline{\frac{\partial^2 I}{\partial u_{i_P}}} \end{bmatrix} = 0 \tag{4.21}$$

(4.21)

$$\begin{bmatrix} \overline{\frac{\partial Q_{i_N+1}}{\partial q_1}} \cdots & \overline{\frac{\partial Y_P}{\partial q_1}} \\ \vdots & \\ \overline{\frac{\partial Q_{i_N+1}}{\partial q_{i_N}}} \cdots & \overline{\frac{\partial Y_P}{\partial q_{i_N}}} \\ \overline{\frac{\partial Q_{i_N+1}}{\partial u_1}} \cdots & \overline{\frac{\partial Y_p}{\partial u_1}} \\ \vdots & \\ \overline{\frac{\partial Q_{i_N+1}}{\partial u_{i_P}}} \cdots & \overline{\frac{\partial Y_p}{\partial u_{i_p}}} \end{bmatrix} \begin{bmatrix} \overline{\frac{\partial^2 I}{\partial q_{i_N+1} \partial q_1}} \cdots & \overline{\frac{\partial^2 I}{\partial q_{i_N+1} \partial u_{i_P}}} \\ \vdots & \\ \overline{\frac{\partial^2 I}{\partial q_N \partial q_1}} \cdots & \overline{\frac{\partial^2 I}{\partial q_N \partial u_{i_P}}} \\ \overline{\frac{\partial^2 I}{\partial u_{i_P+1} \partial q_1}} \cdots & \overline{\frac{\partial^2 I}{\partial u_{i_P+1} \partial u_{i_P}}} \\ \vdots & \\ \overline{\frac{\partial^2 I}{\partial u_P \partial q_1}} \cdots & \overline{\frac{\partial^2 I}{\partial u_P \partial u_{i_P}}} \end{bmatrix} = S \tag{4.22}$$

where S is a skew symmetric matrix and where overbars denote evaluation at $q_s = \bar{q}_s$ $(s = 1, \ldots, i_N)$ and $u_r = 0$ $(r = 1, \ldots, i_p)$. Since equations (4.21) and (4.22) cannot be simplified, it may be easier to check directly whether or not equation (4.19) is an integral of the partially linearized equations than to use equations (4.21) and (4.22).

5 Example

Figure 1 shows a system S consisting of two bars, ,B_1 and B_2, each of mass M and length L. At point P_1, B_1 is connected by means of a revolute joint and a linear torsion spring σ_1 of modulus k_1 to a vertical shaft that is made to rotate at a constant angular speed Ω; and B_1 and B_2 are connected to each other at point P_2 by a revolute joint and a linear torsion spring σ_2 of modulus k_2.

If u_1 and u_2 are defined as $\dot{q}_1$ and $\dot{q}_2$, respectively, where q_1 and q_2 are angles between the horizontal and B_1 and B_2, and if σ_1 and σ_2 are undeformed when $q_1 = q_2 = \pi/2$ rad, then all motions of S are governed by the four first-order differential equations (see Fig. 1 for R)

$$\dot{q}_1 = u_1 \tag{5.1}$$

$$\dot{q}_2 = u_2 \tag{5.2}$$

$$\frac{4}{3}\dot{u}_1 + \frac{1}{2}\cos(q_1 - q_2)\dot{u}_2 = -\frac{1}{2}\sin(q_1 - q_2)u_2^2$$

$$-s_1\left(\frac{3}{2}R/L + \frac{4}{3}c_1 + \frac{1}{2}c_2\right)\Omega^2 - (k_1/ML^2)q_1$$

$$+ (k_2/ML^2)(q_2 - q_1) - \frac{3}{2}(g/L)c_1 \tag{5.3}$$

$$\frac{1}{2}\cos(q_2 - q_1)\dot{u}_1 + \frac{1}{3}\dot{u}_2 = -\frac{1}{2}\sin(q_2 - q_1)u_1^2$$

$$-\frac{1}{2}s_2\left(R/L + c_1 + \frac{2}{3}c_2\right)\Omega^2 - (k_2/ML^2)(q_2 - q_1)$$

$$-\frac{1}{2}(g/L)c_2 \tag{5.4}$$

where g is the local gravitational acceleration and s_i and c_i denote sin q_i and cos q_i ($i=1,2$), respectively.

The Hamiltonian of S, given by

$$H = \frac{4}{3}u_1^2 + \frac{1}{3}u_2^2 + \cos(q_2-q_1)u_1u_2 - \left[\frac{1}{12}(c_1^2+c_2^2) + (R/L+\frac{1}{2}c_1)^2 + \left(R/L+c_1+\frac{1}{2}c_2\right)^2\right]\Omega^2 + (k_1/ML^2)q_1^2 + (k_2/ML^2)(q_2-q_1)^2 + (g/L)(3s_1+s_2) \tag{5.5}$$

is an IGF for equations (5.1)–(5.4); that is, the equation $H=C$ where C is a constant, is an integral of equations (5.1)–(5.4).

Particular solutions of equations (5.1)–(5.4) can be found by setting $u_1=u_2=0$ and letting $\bar{q}_1$ and $\bar{q}_2$ be constants such that

$$\bar{s}_1\left(\frac{3}{2}R/L+\frac{4}{3}\bar{c}_1+\frac{1}{2}\bar{c}_2\right)\Omega^2 + [k_1/(ML^2)]\bar{q}_1 - [k_2/(ML^2)](\bar{q}_2-\bar{q}_1) + \frac{3}{2}(g/L)\bar{c}_1 = 0 \tag{5.6}$$

$$\frac{1}{2}\bar{s}_2\left(R/L+\bar{c}_1+\frac{2}{3}\bar{c}_2\right)\Omega^2 + [k_2/(ML^2)](\bar{q}_2-\bar{q}_1) + \frac{1}{2}(g/L)\bar{c}_2 = 0 \tag{5.7}$$

where $\bar{s}_i$ and $\bar{c}_i$ stand for sin $\bar{q}_i$ and cos $\bar{q}_i$ ($i=1,2$), respectively.

To study motions differing only slightly from those corresponding to such particular solutions, we introduce perturbations u_i and q_i by setting

$$u_i=\hat{u}_i,\ q_i=\bar{q}_i+\hat{q}_i \qquad (i=1,2) \tag{5.8}$$

in equations (5.1)–(5.4) and then linearize in the perturbations. This yields, with the aid of equations (5.6) and (5.7), the linear differential equations

$$\dot{\hat{q}}_1=\hat{u}_1 \tag{5.9}$$

$$\dot{\hat{q}}_2=\hat{u}_2 \tag{5.10}$$

$$\frac{4}{3}\dot{\hat{u}}_1+\frac{1}{2}\cos(\bar{q}_1-\bar{q}_2)\dot{\hat{u}}_2 = -\left[\left(\frac{3}{2}R/L\bar{c}_1+\frac{4}{3}\cos2\bar{q}_1+\frac{1}{2}\bar{c}_1\bar{c}_2\right)\hat{q}_1 - \frac{1}{2}\bar{s}_1\bar{s}_2\hat{q}_2\right]\Omega^2 - [k_1/(ML^2)]\hat{q}_1 + [k_2/(ML^2)](\hat{q}_2-\hat{q}_1) + \frac{3}{2}(g/L)\bar{s}_1\hat{q}_1 \tag{5.11}$$

$$\frac{1}{2}\cos(\bar{q}_1-\bar{q}_2)\dot{\hat{u}}_1+\frac{1}{3}\dot{\hat{u}}_2 = -\left[-\frac{1}{2}\bar{s}_1\bar{s}_2\hat{q}_1+\left(\frac{1}{2}R/L\bar{c}_2+\frac{1}{3}\cos2\bar{q}_2+\frac{1}{2}\bar{c}_1\bar{c}_2\right)\hat{q}_2\right]\Omega^2 - [k_2/(ML^2)](\hat{q}_2-\hat{q}_1) + \frac{1}{2}(g/L)\bar{s}_2\hat{q}_2 \tag{5.12}$$

To construct an integral of this set of equations, we substitute from equations (5.8) into equations (5.5) and then expand H in a power series in $\hat{u}_1$, $\hat{u}_2$, $\hat{q}_1$, $\hat{q}_2$, thus obtaining

$$\tilde{H}=H_0+H_1+H_2 \tag{5.13}$$

where H_i contains solely terms of degree i in $\hat{u}_1$, $\hat{u}_2$, $\hat{q}_1$, $\hat{q}_2$, with

$$H_1 = 2\left\{\bar{s}_1\left(\frac{3}{2}R/L+\frac{4}{3}\bar{c}_1+\frac{1}{2}\bar{c}_2\right)\Omega^2 + [k_1/(ML^2)]\bar{q}_1 - [k_2/(ML^2)](\bar{q}_2-\bar{q}_1) + \frac{3}{2}(g/L)\bar{c}_1\right\}\hat{q}_1 + 2\left\{\frac{1}{2}\bar{s}_2\left(R/L+\bar{c}_1+\frac{2}{3}\bar{c}_2\right)\Omega^2 + [k_2/(ML^2)](\bar{q}_2-\bar{q}_1) + \frac{1}{2}(g/L)\bar{c}_2\right\}\hat{q}_2 \tag{5.14}$$

and

$$H_2 = \frac{4}{3}\hat{u}_1^2+\frac{1}{3}\hat{u}_2^2+\cos(\bar{q}_2-\bar{q}_1)\hat{u}_1\hat{u}_2 - \left\{\frac{1}{12}(-\cos2\bar{q}_1\hat{q}_1^2 - \cos2\bar{q}_2\hat{q}_2^2) - \frac{1}{2}\left(R/L\bar{c}_1+\frac{1}{2}\cos2\bar{q}_1\right)\hat{q}_1^2 - (R/L\bar{q}_1 + \frac{1}{2}\bar{c}_1\bar{c}_2+\cos2\bar{q}_1)\hat{q}_1^2 - \left[\frac{1}{2}(R/L+\bar{c}_1)\bar{c}_2 + \frac{1}{4}\cos2\bar{q}_2\right]\hat{q}_2^2 + \bar{s}_1\bar{s}_2\hat{q}_1\hat{q}_2\right\}\Omega^2 + [k_1/(ML^2)]\hat{q}_1^2 + [k_2/(ML^2)](\hat{q}_2-\hat{q}_1)^2 + g/L\left(-\frac{3}{2}\bar{s}_1\hat{q}_1^2-\frac{1}{2}\bar{s}_2\hat{q}_2^2\right) \tag{5.15}$$

Noting that H_1 vanishes by virtue of equations (5.6) and (5.7), that the system under consideration is holonomic, and that equations (3.27) and (3.28) are satisfied (see equations (5.1) and (5.2)), we now take advantage of the fact that H is an IGF for equations (5.1)–(5.4) to express an integral for equations (5.9–(5.12) as

$$H_2=C, \text{ a constant} \tag{5.16}$$

where H_2 is given by equation (5.15).

References

Kailath, T., 1980, *Linear Systems*, Prentice-Hall, New Jersey, p. 59.

Kane, T. R., and Levinson, D. A., 1985, *Dynamics: Theory and Application*, McGraw-Hill, New York.

Eigenproperties of Nonclassically Damped Primary Structure and Oscillator Systems

L. E. Suarez
Assistant Professor.

M. P. Singh
Professor.

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

The calculation of the combined eigenproperties of a nonclassically damped structure and a supported equipment is of practical interest. Herein an approach is developed whereby these properties can be obtained, in terms of the eigenproperties of the structure and equipment, without a conventional eigenvalue analysis of the combined system. The eigenvalues are obtained as the solutions of a nonlinear characteristic equation, easily solvable by a simple Newton-Raphson scheme. Once the eigenvalues are known, the corresponding eigenvectors can be obtained from closed-form expressions. The approach can also be used effectively to obtain exact eigenproperties for very light as well as very heavy equipment supported on structures.

Introduction

The problem of the dynamic response of equipment or internal structures attached to massive structures, also referred to as primary structures, is of practical importance in structural dynamics. The proper functioning of certain equipment when the structure is subjected to dynamic forces is essential in facilities like power plants, hospitals, chemical factories, etc. Other physical systems where the response of mounted equipment or secondary systems is of interest are vehicular structures, where guidance and control devices must always remain operational and thus require a careful design.

If the system response is to be obtained through modal analysis, the modal properties of the combined structure and equipment system must be obtained. Some problems, however, may arise when the combined system is analyzed to obtain its eigenvalues and eigenvectors. For very light equipment, the combined system matrices may be ill-conditioned due to the large differences in the numerical values of the equipment and structure mass and stiffness properties. Furthermore, the combined analysis, though maybe possible with increased computational precision, becomes impractical in the design situation where several alternative locations and properties for the equipment need to be considered, each requiring the solution of a large eigenvalue problem. For example, for the generation of seismic floor response spectra incorporating equipment-structure interaction effects, such repeated analysis will be required with different characteristics of the oscillator placed at different locations. Thus the approach in which such repeated large eigenvalue analyses can be avoided and the individual modal characteristics of the structure and equipment can be used to obtain combined properties are preferred.

For light equipment, perturbation methods have been proposed by Sackman and Kelly (1979), Sackman et al. (1983), and Suarez and Singh (1986) to obtain the eigenproperties of the combined system. Alternative approaches based on the synthesis of modes, which can be used with light as well as heavy equipment, have also been proposed by the authors (Suarez and Singh, 1987a and 1987b). In all of these works, however, the supporting structure was assumed to be classically or proportionally damped (Caughey, 1960). There are certain important cases where the primary system must be regarded as nonproportionally damped. This is usually the case when the primary system is composed of two or more components with widely different energy dissipation characteristics.

In this paper a method is presented to obtain the exact complex eigenproperties of a nonclassically damped structure supporting a single degree of freedom oscillator. The eigenvalues are obtained as the solution of a simple nonlinear equation with real coefficients. The corresponding eigenvectors are calculated from the closed-form expressions once the eigenvalues are known. To implement this method it is only necessary to know the complex eigenvalues and eigenvectors of the primary structure and, of course, the natural frequency, mass, and damping ratio of the equipment.

Eigenproperties of the Combined Structure-Equipment System

It is assumed that the damped eigenproperties of the primary structure are available. For a primary system, mo-

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13-18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 12, 1986.

Paper No. 87-WA/APM-8.

delled as an n degree of freedom system with stiffness matrix $[K_p]$, damping matrix $[C_p]$, and mass matrix $[M_p]$, the damped eigenvalues λ_j and eigenvectors ϕ_j are obtained as a solution of the following eigenvalue problem (Meirovitch, 1980):

$$\begin{bmatrix} M_p & 0 \\ 0 & -K_p \end{bmatrix} \phi_j = \lambda_j \begin{bmatrix} 0 & M_p \\ M_p & C_p \end{bmatrix} \phi_j;\ j=1, \ldots .2n \tag{1}$$

These eigenvalues and eigenvectors occur in complex conjugate pairs. Also, the upper half of the eigenvector can be written as a product of the eigenvalue and the lower half. Thus, we can partition the $(2n \times 2n)$ eigenvector matrix as follows

$$[\Phi_p] = \begin{bmatrix} U\Lambda_p & \bar{U}\bar{\Lambda}_p \\ U & \bar{U} \end{bmatrix} \tag{2}$$

where the matrix $[U]$ is a submatrix composed of the last n rows and first n columns of $[\Phi_p]$; $[\Lambda_p]$ is a diagonal matrix whose elements are the eigenvalues λ_j; and a bar over a quantity denotes its complex conjugate value. If the modal shape vectors are normalized with respect to the matrix on the right-hand side of equation (1), the orthonormality condition of the eigenvectors give:

$$\begin{aligned} &\Lambda_p U^T M_p U \Lambda_p - U^T K_p U = \Lambda_p \\ &U^T M_p U \Lambda_p + \Lambda_p U^T M_p U + U^T C_p U = I \\ &\bar{\Lambda}_p \bar{U}^T M_p U \Lambda_p - \bar{U}^T K_p U = 0 \\ &\bar{U}^T M_p U \Lambda_p + \bar{\Lambda}_p \bar{U}^T M_p U + \bar{U}^T C U = 0 \end{aligned} \tag{3}$$

Analogous damped eigenvalues and eigenvectors can be defined for the oscillator. For an oscillator of mass m_e, stiffness, k_e, and damping constant c_e, we define the following (2×2) eigenvalue problem

$$\begin{bmatrix} m_e & 0 \\ 0 & -k_e \end{bmatrix} \phi_{s_j} = \lambda_{s_j} \begin{bmatrix} 0 & m_e \\ m_e & c_e \end{bmatrix} \phi_{s_j};\ j=1,2 \tag{4}$$

from which we obtain the equipment eigenvalues and eigenvectors, in terms of the equipment (undamped) natural frequency $\omega_e = \sqrt{k_e/m_e}$, and critical damping ratio $\beta_e = c_e/(2m_e\omega_e)$, as follows

$$\lambda_{s_1} = \bar{\lambda}_{s_2} = \lambda_e = -\beta_e\omega_e + i\omega_e\sqrt{1-\beta_e^2} \tag{5}$$

$$\phi^T{}_{s_1} = \bar{\phi}^T{}_{s_2} = [\phi_e\lambda_e,\ \phi_e] \tag{6}$$

For the equipment also we normalize the eigenvectors with respect to the matrix on the right-hand side of equation (4), and thus get

$$\phi_e = \frac{1-i}{2\sqrt{\omega_e m_e (1-\beta_e^2)^{1/2}}} \tag{7}$$

The properties described in equations (3) also hold for this case if we substitute $[\Lambda_p]$ by λ_e and $[U]$ by ϕ_e.

We now examine the combined structure and equipment system. The equations of motion for the combined system subjected to the dynamic excitation $\mathbf{F}(t)$ can be written in the state vector form as follows:

$$\begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix} \mathbf{z} + \begin{bmatrix} 0 & M \\ M & C \end{bmatrix} \dot{\mathbf{z}} = \mathbf{F}(t) \tag{8}$$

where $\mathbf{z}$ is the state vector

$$\mathbf{z}^T = [\dot{\mathbf{x}}^T\ \mathbf{x}^T] \tag{9}$$

and $\mathbf{x}$ is the displacement vector of the combined system of dimension $m = n+1$. The matrices $[M]$ and $[K]$ are

$$[M] = \begin{bmatrix} M_p & \mathbf{0} \\ \mathbf{0} & m_e \end{bmatrix};\ [K] = \begin{bmatrix} K_p & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} + k_e \mathbf{v}\,\mathbf{v}^T \tag{10}$$

Matrix $[C]$ is of similar form as $[K]$ and is obtained by replacing $[K_p]$ by $[C_p]$ and k_e by c_e in equation (10). Assuming that the oscillator is attached to the Kth degree of freedom of the primary system, vector $\mathbf{v}$ is

$$\mathbf{v}^T = [0, \ldots, 1, \ldots, -1] \tag{11}$$

where the nonzero entries are at the Kth and mth locations.

The eigenvalue problem associated with the combined system (8) is as follows:

$$\begin{bmatrix} M & 0 \\ 0 & -K \end{bmatrix} \hat{\psi}_j = p_j \begin{bmatrix} 0 & M \\ M & C \end{bmatrix} \hat{\psi}_j;\ j=1, \ldots, 2m \tag{12}$$

It is desired to obtain the combined system eigenproperties $\hat{\psi}_j$ and p_j. However, we would like to avoid a direct solution of equation (12) and would prefer to obtain these eigenproperties in terms of the known eigenproperties of the systems defined in equations (1) and (4). To achieve this objective we introduce the following transformation in equation (12)

$$\hat{\psi}_j = [T]\psi_j = \begin{bmatrix} T_u & \bar{T}_u \\ T_\ell & \bar{T}_\ell \end{bmatrix} \psi_j \tag{13}$$

where the submatrices $[T_u]$ and $[T_\ell]$ are

$$[T_u] = \begin{bmatrix} U\Lambda_p & \mathbf{0} \\ \mathbf{0} & \phi_e\lambda_e \end{bmatrix};\ [T_\ell] = \begin{bmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \phi_e \end{bmatrix} \tag{14}$$

We then premultiply the resulting equation by $[T]^T$ and obtain the following transformed eigenvalue problem

$$[A]\psi_j = p_j[B]\psi_j;\ j=1, \ldots, 2m \tag{15}$$

where

$$[A] = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & \bar{A}_{11} \end{bmatrix};\ [B] = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}^T & \bar{B}_{11} \end{bmatrix} \tag{16}$$

with

$$A_{11} = T_u^T M T_u - T_\ell^T K T_\ell;\ A_{12} = T_u^T M \bar{T}_u - T_\ell^T K \bar{T}_\ell \tag{17}$$

$$B_{11} = T_\ell^T M T_u + T_u^T M T_\ell + T_\ell^T C T_\ell;$$

$$B_{12} = T_\ell^T M \bar{T}_u + T_u^T M \bar{T}_\ell + T_\ell^T C \bar{T}_\ell \tag{18}$$

Using the definitions of $[M]$, $[C]$, and $[K]$ from equations (10), and the orthogonality properties listed in equations (3), it can be shown that the above submatrices become

$$[A_{11}] = [\Lambda_c] - k_e[N_1];\ [A_{12}] = -k_e[N_2] \tag{19}$$

$$[B_{11}] = [I] + c_e[N_1];\ [B_{12}[= c_e[N_2] \tag{20}$$

where

$$[\Lambda_c] = \begin{bmatrix} \Lambda_p & \mathbf{0} \\ \mathbf{0} & \lambda_e \end{bmatrix};\ [N_1] = \begin{bmatrix} \boldsymbol{\alpha}\,\boldsymbol{\alpha}^T & -\phi_e\boldsymbol{\alpha} \\ -\phi_e\boldsymbol{\alpha}^T & 0 \end{bmatrix};$$

$$[N_2] = \begin{bmatrix} \boldsymbol{\alpha}\,\bar{\boldsymbol{\alpha}}^T & -\bar{\phi}_e\boldsymbol{\alpha} \\ -\phi_e\bar{\boldsymbol{\alpha}}^T & 0 \end{bmatrix} \tag{21}$$

$$\boldsymbol{\alpha}^T = [\phi_{n+K,1},\ \phi_{n+K,2}, \ldots, \phi_{n+K,n}] \tag{22}$$

and $\phi_{n+K,i}$ is the $(n+K)$th element of the ith eigenvector of the primary system $\boldsymbol{\phi}_i$. With the substitution of equations (19) and (20), the eigenvalue problem (15) can now be expressed as follows

$$[[\Lambda]-k_e[D]]\psi_j=p_j[[I]+c_e[D]]\psi_j;\ j=1,\ \ldots,2m \tag{23}$$

where

$$[\Lambda]=\begin{bmatrix}\Lambda_c & 0\\ 0 & \bar{\Lambda}_c\end{bmatrix};[D]=\begin{bmatrix}N_1 & N_2\\ \bar{N}_2 & \bar{N}_1\end{bmatrix}. \tag{24}$$

The eigenvalue problem (23) is in such a form now that with some simple algebraic manipulations it can be expanded to define its characteristic equation by a simple closed-form equation. For this, we rearrange the terms of equation (23) and obtain

$$[p_jI-\Lambda]\psi_j=-(k_e+c_ep_j)[D]\psi_j \tag{25}$$

The matrix on the left-hand side of equation (25) is a diagonal matrix, and thus can be easily inverted to give the following

$$\psi_j=-(k_e+c_ep_j)[\ \delta_i\ ]^{-1}[D]\psi_j \tag{26}$$

where the diagonal elements δ_i are defined as

$$\delta_i=p_j-\lambda_i;\ \delta_{i+m}=p_j-\bar{\lambda}_i;\ i=1,\ \ldots,n \tag{27}$$

$$\delta_m=p_j-\lambda_e;\ \delta_{2m}=p_j-\bar{\lambda}_e \tag{28}$$

Furthermore, from the definition of $[D]$ in terms of $[N_1]$ and $[N_2]$, we note that it is possible to express this in terms of the products of two vectors as follows:

$$[D]=[\nu\ \nu^T-\mu\ \mu^T] \tag{29}$$

where

$$\nu^T=(\alpha^T,\ -\phi_e,\ \bar{\alpha}^T,\ -\bar{\phi}_e)$$
$$\mu^T=(\mathbf{0}^T,\ \phi_e,\mathbf{0}^T,\ \bar{\phi}_e) \tag{30}$$

The nonzero elements of μ^T are at the mth and $2m$th locations.

We now substitute equation (29) into (26), and premultiply by ν^T and obtain

$$\nu^T\Psi_j=-(k_e+c_ep_j)\nu^T[\ \delta_i\ ]^{-1}(\nu\ \nu^T\psi_j-\mu\ \mu^T\psi_j) \tag{31}$$

We also denote the last vector product in the parenthesis of equation (31) by β. This vector can be expanded as

$$\beta^T=[\mu\ \mu^T\psi_j]^T=(0,\ \ldots,\phi_e\Delta,\ \ldots.0,\ \ldots.\bar{\phi}_e\Delta) \tag{32}$$

in which the only nonzero elements $\phi_e\Delta$ and $\bar{\phi}_e\Delta$ are in the mth and $(2m)$th rows. The term Δ is defined as:

$$\Delta=\phi_e\psi_{m,j}+\bar{\phi}_e\psi_{2m,j} \tag{33}$$

Realizing that we are free to normalize an eigenvector any way we choose, here we adopt a rather uncommon but simplifying normalization of the eigenvector ψ_j such that $\nu^T\psi_j=1$. With this normalization and some rearrangements we now obtain the following for equation (31):

$$-\frac{1}{k_e+c_ep_j}=\nu^T[\ \delta_i\ ]^{-1}\nu-\nu^T[\ \delta_i\ ]^{-1}\beta \tag{34}$$

The right-hand side of equation (34), involving the product of vectors with a diagonal matrix, can be easily expanded to give the following:

$$\sum_{i=1}^{2m}\frac{\nu_i^2}{\delta_i}-\left(\frac{\nu_m}{\delta_m}\phi_e+\frac{\nu_{2m}}{\delta_{2m}}\bar{\phi}_e\right)\Delta+\frac{1}{k_e+p_jc_e}=0 \tag{35}$$

Substituting for δ_m and ν from equations (28) and (30), we obtain

$$\sum_{i=1}^{2m}\frac{\nu_i^2}{\delta_i}+\left(\frac{\phi_e^2}{p_j-\lambda_e}+\frac{\bar{\phi}_e^2}{p_j-\bar{\lambda}_e}\right)\Delta+\frac{1}{k_e+p_jc_e}=0 \tag{36}$$

With λ_e and ϕ_e defined as in equations (5) and (7), it is straightforward to show that

$$\frac{\phi_e^2}{p_j-\lambda_e}+\frac{\bar{\phi}_e^2}{p_j-\bar{\lambda}_e}=\frac{1}{m_e}\,\frac{1}{p_j^2+2\beta_e\omega_ep_j+\omega_e^2} \tag{37}$$

We note that all the terms in equation (36) except Δ and the combined system eigenvalues p_j are known. However, Δ can also be expressed in terms of the unknowns p_j as follows. We note that equation (31), with the normalization $\nu^T\psi_j=1$ and β defined by equation (32), can also be written as:

$$-\frac{1}{k_e+p_jc_e}[\ \delta_i\ ]\psi_j+\beta=\nu \tag{38}$$

From the mth and $(2m)$th rows of the above set of equations we obtain

$$-\frac{p_j-\lambda_e}{k_e+c_ep_j}\psi_{m,j}+\phi_e\Delta=-\phi_e \tag{39}$$

$$-\frac{p_j-\bar{\lambda}_e}{k_e+c_ep_j}\psi_{2m,j}+\bar{\phi}_e\Delta=-\bar{\phi}_e \tag{40}$$

We solve equations (39) and (40) for $\psi_{m,j}$ and $\psi_{2m,j}$ in terms of Δ and substitute in equation (33) to obtain the following:

$$\Delta=(k_e+c_ep_j)\left(\frac{\phi_e}{p_j-\lambda_e}+\frac{\bar{\phi}_e}{p_j-\bar{\lambda}_e}\right)(1+\Delta) \tag{41}$$

Utilizing equation (37) in equation (41) and solving for Δ, we obtain

$$\Delta=\frac{2\beta_e\omega_ep_j+\omega_e^2}{p_j^2} \tag{42}$$

With equations (37) and (42), equation (36) now becomes

$$m_e\sum_{i=1}^{2m}\frac{\nu_i^2}{\delta_i}+\frac{2\beta_e\omega_ep_j+\omega_e^2}{p_j^4+2\beta_e\omega_ep_j^3+\omega_ep_j^2}+\frac{1}{2\beta_e\omega_ep_j+\omega_e^2}=0 \tag{43}$$

The summation over $2m$ terms in the above equation can be written as a summation over n complex and conjugate terms as follows

$$m_e\sum_{i=1}^{2m}\frac{\nu_i^2}{\delta_i}=\sum_{i=1}^{n}\frac{a_ip_j+b_i}{p_j^2+c_ip_j+d_i}+\frac{1}{p_j^2+2\beta_e\omega_ep_j+\omega_e^2} \tag{44}$$

where

$$a_i=2m_e\text{Real}(\nu_i^2);\ b_i=-2m_e\text{Real}(\nu_i^2\overline{\lambda_i})$$
$$c_i=-2\text{Real}(\lambda_i);\ d_i=|\lambda_i|^2 \tag{45}$$

Substituting equation (44) into equation (43) and with some simplifications, we finally obtain the following characteristic equation:

$$f(p_j)=\sum_{i=1}^{n}\frac{a_ip_j+b_i}{p_j^2+c_ip_j+d_i}+\frac{1}{2\beta_e\omega_ep_j+\omega_e^2}+\frac{1}{p_j^2}=0 \tag{46}$$

This equation is defined in terms of the known eigenproperties of the primary system and equipment. The only unknown is p_j. The solution of this characteristic equation will provide the combined system eigenvalues p_j. This equation can be solved by any standard iterative technique. We find the simple Newton-Raphson technique quite adequate for solving equation (46). Initial values for the roots of this equation, required in the Newton-Raphson iteration process are provided in Appendix I.

Combined Eigenvectors

Once the first m roots of $f(p_j)$ are found, the eigenvectors can be obtained directly from equation (38). For each p_j, equation (42) defines Δ, which in turn defines β through equation (32). Thus knowing β and ν, the eigenvector ψ_j can be obtained from equation (38) as

$$\psi_j=(k_e+p_jc_e)[\ \delta_i\ ]^{-1}(\beta-\nu) \tag{47}$$

Recalling equation (32), we can simplify equation (47) to ob-

tain the closed-form expression for the elements of this eigenvector as:

$$\psi_{i,j} = -(k_e + c_e p_j)\frac{\nu_i}{p_j - \lambda_i}; \qquad i = 1, \ldots, n$$

$$\psi_{i+m,j} = -(k_e + c_e p_j)\frac{\bar{\nu}_i}{p_j - \bar{\lambda}_i} \qquad j = 1, \ldots, m \tag{48}$$

The values of the remaining elements of this eigenvector $\psi_{m,j}$ and $\psi_{2m,j}$ are also obtained from equation (47), by substituting from Δ from equation (42) and ν_m and ν_{2m} from equation (30), as follows:

$$\psi_{m,j} = (k_e + c_e p_j)\frac{p_j^2 + 2\beta_e\omega_e p_j + \omega_e^2}{p_j^2}\frac{\phi_e}{p_j - \lambda_e} \quad ; j = 1, \ldots m$$

$$\psi_{2m,j} = (k_e + c_e p_j)\frac{p_j^2 + 2\beta_e\omega_e p_j + \omega_e^2}{p_j^2}\frac{\bar{\phi}_e}{p_j - \bar{\lambda}_e} \tag{49}$$

The remaining m eigenvectors can be obtained with:

$$\begin{aligned} \psi_{i,j+m} &= \bar{\psi}_{i+m,j} \qquad i = 1, \ldots, m \\ \psi_{i+m,j+m} &= \bar{\psi}_{i,j} \qquad j = 1, \ldots, m \end{aligned} \tag{50}$$

The solution of the characteristic equation (46) and equations (48)–(50) provide a complete solution of the eigenvalue problem (23) without going through a conventional, computationally expensive, eigenvalue analysis. This efficiency is of special significance where repeated analyses with different oscillator characteristics are required to be performed, such as in the process of generation of seismic floor response spectra which are used as inputs for design of equipment supported on primary structures.

The eigenvectors can be rendered orthonormal with respect to the matrix in the right-hand side of equation (23). It can be shown that this normalization is achieved by multiplying them by the complex constant

$$\theta_j = \left\{2\beta_e\omega_e m_e\left[1 - \frac{(2\beta_e\omega_e p_j + \omega_e^2)^2}{p_j^4}\right] + \sum_{i=1}^{2m}\psi_{i,j}^2\right\}^{-1/2} \tag{51}$$

For light equipment with the mass ratio less than 1/10, the eigenvalues calculated from the equations provided in Appendix I are very accurate. In such a case one may not want to refine these estimates any further by solving equation (46). For heavier equipment, however, such refinements are necessary but can be easily carried out by simply solving equation (46) by a Newton-Raphson scheme with the help of the initial estimates of the roots of the equations given in the appendix. As for the calculation of the eigenvectors, equations (48) and (49) can still be used with no numerical difficulties for light as well as heavy equipment.

If one is interested in evaluating the limiting value of the eigenvector elements for the case of $m_e = 0$, equations (*I*1) and (*I*2) can be used. It is straightforward to show that in such a case of $m_e = 0$,

$$\psi_{i,j} = \begin{cases} 0 & ; \quad i \neq j,\ i = 1, \ldots 2m,\ j = 1, \ldots, m \\ \dfrac{(2\beta_e\omega_e\lambda_j + \omega_e^2)(\lambda_j^2 + 2\beta_e\omega_e\lambda_j + \omega_e^2)}{\omega_e^2\nu_j\lambda_j(\lambda_j + 2\beta_e\omega_e)}; & i = j,\ j = 1, \ldots, n \end{cases} \tag{52}$$

which when normalized according to equation (51) will give

$$\psi_{i,j} = \delta_{ij} \tag{53}$$

Equation (53) implies that the original eigenvectors of the structure and equipment will remain unchanged, an obvious conclusion.

To obtain the eigenvectors $\hat{\psi}_i$ of the original system from the eigenvectors of the transformed system ψ_i, we use the transformation of equation (13). It is straightforward to show

Table 1 Damping matrix of the nonclassically damped primary structure

$$[C_p] = \begin{bmatrix} 20.0 & -4.0 & -0.4 & -0.1 & -0.08 & -0.06 \\ & 9.0 & -4.0 & -0.3 & -0.2 & -0.15 \\ & & 8.0 & -4.0 & -0.3 & -0.2 \\ & & & 7.0 & -2.0 & -0.6 \\ & \text{SYMM} & & & 5.0 & -3.0 \\ & & & & & 4.0 \end{bmatrix} \times 10^8 \text{[Kg/sec]}$$

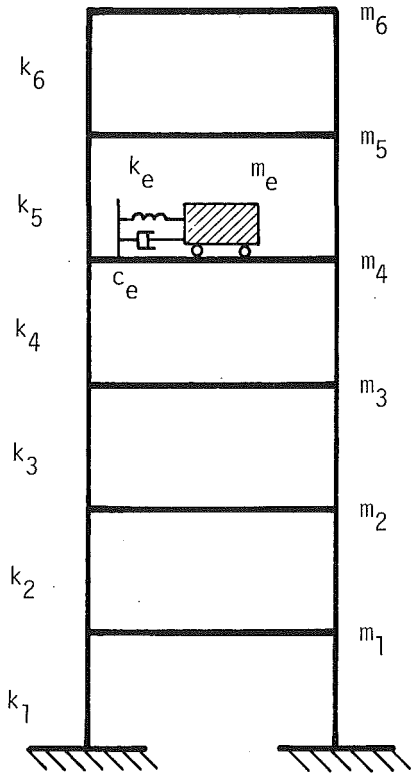


Fig. 1 A six degrees of freedom primary structure with oscillator

that the lower and upper halfs of the original system eignvectors are:

$$\hat{\psi}_j^\ell = [T_\ell]\psi_j^u + [\bar{T}_\ell]\psi_j^\ell \quad ; \quad j = 1, \ldots, m$$

$$\hat{\psi}_j^u = p_j\hat{\psi}_j^\ell \tag{54}$$

where the superscripts u and ℓ refer to the upper and lower halves, respectively. Substituting for ψ_j^u and ψ_j^ℓ from equations (48) and (49), and with $[T_\ell]$ defined by equation (14), we obtain the following closed-form expressions for the elements of the lower part of the original system eigenvectors:

$$\hat{\psi}_{i+m,j} = \begin{cases} = (k_e + c_e p_j)\theta_j\displaystyle\sum_{k=1}^{n}\frac{e_k p_j + f_k}{p_j^2 + c_k p_j + d_k}; & i = 1, \ldots, n \\ \dfrac{k_e + c_e p_j}{m_e p_j}\theta_j; & i = m \end{cases} \tag{55}$$

Table 2 Complex eigenvalues of the primary structure

Eigenvalues of primary system		
No.	Real	Imaginary
1	-0.2923	23.8724
2	-3.7472	61.5101
3	-7.2993	97.1178
4	-11.7546	132.5499
5	-11.7583	153.0566
6	-12.1126	170.3976

Table 3 Eigenvalues of the combined damped structure-oscillator system. Undamped oscillator frequency = 23.87 rad/s—Mass ratio = 1/2.

Eigenvalue No.	Initial Eigenvalues		Final Eigenvalues		Number of Iterations
	Real	Imaginary	Real	Imaginary	
1	-0.2184	20.3492	-0.1830	20.0146	4
2	-0.3242	28.6534	-0.3831	28.2502	4
3	-3.7434	61.7358	-3.7565	61.7378	3
4	-7.3007	97.1200	-7.3009	97.1199	3
5	-11.7768	132.6637	-11.7888	132.6584	4
6	-11.7791	153.3001	-11.8112	153.2927	4
7	-12.0584	170.6659	-12.0989	170.6761	4

Table 4 Eigenvalues of the combined damped structure-oscillator system. Undamped oscillator frequency = 40.0 rad/s—Mass ratio = 1.

Eigenvalue No.	Initial Eigenvalues		Final Eigenvalues		Number of Iterations
	Real	Imaginary	Real	Imaginary	
1	-0.3590	18.7342	-0.1996	20.2035	4
2	-0.8381	43.6057	-0.8169	42.1339	4
3	-3.4283	67.2862	-3.7923	66.6982	4
4	-7.1231	99.2579	-7.2265	99.4323	4
5	-11.7759	133.2061	-11.8177	133.2468	4
6	-11.7834	153.5581	-11.8305	153.5939	4
7	-12.0566	170.6857	-12.0810	170.7184	4

where the constants e_k and f_k are:

$$e_k = 2\ \text{Real}(U_{ik}\nu_k);\ f_k = -2\ \text{Real}(U_{ik}\nu_k\bar{\lambda}_k) \tag{56}$$

in which U_{ik} is the (i, k)th element of matrix $[U]$. If necessary, the elements of the upper half of the eigenvector can be obtained from the second part of equation (54).

It can be shown that when the eigenvectors ψ_j are rendered orthonormal with respect to the matrix on the right-hand side of equation (23) with the help of the constant in equation (51) the eigenvectors $\hat{\psi}_j$ of the original system will be orthonormal in the following sense

$$\hat{\psi}_i^T \begin{bmatrix} 0 & M \\ M & C \end{bmatrix} \hat{\psi}_j = \delta_{ij};\ i, j = 1, \ldots, 2m \tag{57}$$

Once the eigenproperties of the combined system are known, the response of the system described by equation (8) can be obtained for any arbitrary forcing function.

An Example Problem

A six degree of freedom primary structure modelled as a shear building shown in Fig. 1 is considered to present some numerical results. Its mass and stiffness properties are: $k_1 = k_2 = 5\times10^{11}$ N/m; $k_3 = k_4 = 4/5\,k_1$; $k_5 = k_6 = 7/10\,k_1$; $m_1 = m_2 = 7\times10^7$ kg; $m_3 = m_4 = 5/7\,m_1$; $m_5 = m_6 = 4/7\,m_1$. The nonclassicality of damping in the structure was introduced by arbitrarily selecting a damping matrix. The damping matrix used in the calculations is shown in Table 1. The complex eigenvalues of the primary structure obtained with these structural parameters are shown in Table 2.

Table 3 shows the case when an oscillator is attached to the fifth floor with mass equal to 1/2 of the supporting floor mass. The equipment natural frequency is tuned to the first frequency (modulus of the damped eigenvalue) of the primary system. The equipment damping ratio is 0.01. Columns 2 and 3 show the initial approximations to the eigenvalues obtained according to Appendix I, and columns 4 and 5 show the final values obtained by solving equation (46). The number of iterations required to achieve a desired accuracy in a straightforward Newton-Raphson root finding scheme are given under column 6. The final values are identical up to 12 significant figures to the values obtained by a direct application of a conventional eigenvalue-solver subroutine applied to the combined system represented by equation (12). Similar results are shown in Table 4 for an oscillator of mass equal to the floor mass, supported on the top floor, but not tuned to any structural frequency. Again when compared up to 12 significant figures, the final values are identical to the values obtained by a direct eigenvalue analysis of equation (12).

To demonstrate the computational efficiency of the proposed approach, several cases of primary structure and equipment systems with increasing number of degrees of freedom were analyzed. The complex eigenproperties of these systems were obtained by the proposed characteristic equation approach as well as by a direct solution of the algebraic eigenvalue problem by a standard IMSL subroutine. The CPU time taken by the two approaches, on an IBM 3090 Class VI Supercomputer, is plotted against the number of degrees of freedom of the system, m, in Fig. 2. It is seen that the CPU time taken by the proposed approach increases linearly with the degrees-of-freedom with a rather flat slope. The CPU time taken by

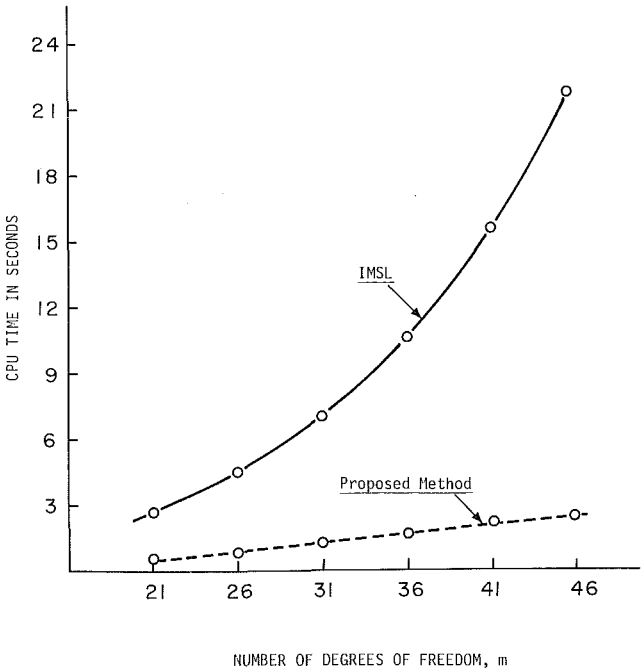


Fig. 2 Comparison of CPU time taken by the proposed approach and direct eigenvalue analysis on IBM 3090, class VI, super-computer

the direct eigenvalue analysis, on the other hand, is seen to increase approximate as m^3. Thus the difference in the CPU time taken by the two methods increases as m increases. It is also noted that the CPU time difference shown in Fig. 2 is for only one oscillator with a given set of parameters. If it is necessary to consider several oscillators with different frequency and damping characteristics (for example in generation of seismic floor response spectrum), this difference in computation cost can be very large indeed. Furthermore, the differences shown in Fig. 2 are for a Class VI computer; for an earlier generation computer or microcomputers the difference in CPU time will be even larger.

Conclusions

An efficient approach is developed for calculating the complex eigenproperties of a combined system, composed of a nonclassically damped primary structure and an oscillator, in terms of the eigenproperties of the individual systems. The eigenvalues can be obtained as a solution of a nonlinear characteristic equation by a simple Newton-Raphson scheme. The initial estimates of the eigenvalue required in the Newton-Raphson scheme are provided. Once the eigenvalues are known, the calculation fo the eigenvectors is straightforward; the closed form expressions are given for calculating the eigenvectors elements. These eigenproperties can be used in any linear response analysis for an arbitrarily varying forcing function.

Acknowledgments

This research is supported by the National Science Foundation through Grant No. CEE-8208897 with Dr. S.C. Liu as its Program Director. This support is gratefully acknowledged.

References

Caughey, T. K., 1960, "Classical Normal Modes in Damped Linear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 27, pp. 269-271.

Householder, A. S., 1970, *The Numerical Treatment of a Single Nonlinear Equation*, McGraw-Hill, New York, NY.

Meirovitch, L., 1980, *Computational Methods in Structural Dynamics*, Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands.

Sackman, J. L., and Kelly, J. M., 1979, "Seismic Analysis of Internal Equipments and Components in Structures," *Engineering Structures*, Vol. 1, No. 4, pp. 179-190.

Sackman, J. L., Der Kiureghian, A., and Nour-Omid, B., 1983, "Dynamic Analysis of Light Equipment in Structures: Modal Properties of the Combined System," *ASCE J. of Engr. Mech.*, Vol. 109, pp. 73-89.

Singh, M. P., and Suarez, L. E., 1986, "A Perturbation Analysis of the Eigenproperties of Equipment-Structure Systems," *Nuclear Engr. & Design*, Vol. 97, pp. 167-186.

Suarez, L. E., and Singh, M. P., 1987a, "Seismic Response of SDF Equipment-Structure Systems," *ASCE J. of Engr. Mech.*, Vol. 113, pp. 16-30.

Suarez, L. E., and Singh, M. P., 1987b, "Floor Response Spectra with Structure-Equipment Interaction Effects by a Mode Synthesis Approach," *Earthquake Engr. and Structural Dynamics*, Vol. 15, pp. 141-158.

Suarez, L. E., and Singh, M. P., 1987c, "Eigenproperties of Nonclassically Damped Primary Structure and Equipment Systems," *Earthquake Engr. and Structural Dynamics*, Vol. 16, to appear.

APPENDIX I

The roots of the function $f(p)$, equation (46), can be obtained with any complex roots-finding technique (Householder, 1970). In particular the common Newton-Raphson method was found to be very convenient because of its simplicity and quadratic convergence rate. It is well known that the main drawback of the Newton-Raphson method is that it requires good initial approximations of the roots. However, for our case, these approximations can be obtained by solving the eigenvalue problem (23) by a perturbation technique (Suarez and Singh, 1987c).

As shown by Suarez and Singh (1987c), two different cases must be considered: (1) when the complex equipment eigenvalue λ_e is not close to any primary structure eigenvalue, and, (2) when both eigenvalues have equal or nearly equal numerical values. It is shown that for case (1), the eigenvalues of the combined system can be approximately obtained by the following expressions

$$p_j = \lambda_j \left(1 - m_e \omega_e^2 \nu_j^2 \frac{\lambda_j + 2\beta_e \omega_e}{\lambda_j^2 + 2\beta_e \omega_e \lambda_j + \omega_e^2}\right); \; j = 1, \ldots, n \tag{I1}$$

$$p_m = \lambda_e + m_e \omega_e^4 \phi_e^2 \sum_{i=1}^{n} \frac{a_i \lambda_e + b_i}{\lambda_e^2 + c_i \lambda_e + d_i} \tag{I2}$$

For case (2) when the equipment eigenvalue λ_e and one of the structure eigenvalues, say the ℓth, are such that the following condition is satisfied

$$\frac{|\lambda_\ell - \lambda_e|}{\omega_e^2} \leq \frac{1}{\sqrt{2}} m_e |\nu_\ell| \; |\phi_e| \tag{I3}$$

the ℓth and mth eigenvalues are given by the following equations

$$p_\ell = \lambda_\ell + \frac{1}{2}[\delta\lambda_e - m_e\omega_e^2\nu_\ell(2\phi_e + \nu_\ell) - 2\beta_e\omega_e m_e \phi_e \nu_\ell \lambda_\ell + m_e \omega_e^4 \phi_e^2 \sigma] \tag{I4}$$

$$p_m = \lambda_\ell + \frac{1}{2}[\delta\lambda_e + m_e\omega_e^2\nu_\ell(2\phi_e + \nu_\ell) + 2\beta_e\omega_e m_e \phi_e \nu_\ell \lambda_\ell + m_e \omega_e^4 \phi_e \sigma \tag{I5}$$

where

$$\delta = 1 - \frac{\lambda_\ell}{\lambda_e} \tag{I6}$$

$$\sigma = \sum_{\substack{i=1 \\ i \neq \ell}}^{n} \frac{a_i \lambda_\ell + b_i}{\lambda_\ell^2 + c_i \lambda_\ell + d_i} \tag{I7}$$

With these initial estimates, the convergence of the Newton-Raphson method has been quite fast.

Recursive Simulation of Stationary Multivariate Random Processes—Part I

M. P. Mignolet[1]
Research Assistant,
Student Mem. ASME

P. D. Spanos
Professor,
Mem. ASME

Brown School of Engineering,
Rice University,
Houston, Texas 77251

A unified approach is presented in determining autoregressive moving average (ARMA) algorithms for simulating realizations of multivariate random processes with a specified (target) spectral matrix. The ARMA algorithms are derived by relying on a prior autoregressive (AR) approximation of the target matrix. Several AR to ARMA procedures are formulated by minimizing a frequency domain error. Equations which can lead to a convenient computation of the ARMA matrix coefficients for a particular problem are given. Finally, the features of the various procedures are critically assessed.

Introduction

Over a period of several years significant interest has developed in the techniques of signal processing and spectral analysis in particular (Kay and Marple, 1981). The fields of application are numerous. These techniques are applicable both to system identification and system response simulation. Historically, the system identification applications preceded the system response simulation. However, with the advent of numerical quadrature techniques and the inevitability of nonlinear behavior of many engineering systems, the simulation problem has received rapidly increasing attention.

Within the scope of structural dynamics applications, significant efforts have been devoted to the problem of simulating realizations of a random process which are compatible with a specified (target) spectrum. (Mignolet and Spanos, 1987). The generation of these time histories for multivariate random processes has traditionally been achieved by relying on the superposition of several harmonic components with random phases (Shinozuka, 1970, 1972). Recently, new and computationally more efficient algorithms based on the development of multiple input-multiple output autoregressive (AR) and autoregressive moving average (ARMA) discrete systems have been suggested by Spanos and Hansen (1981), Samii and Vandiver (1984), Samaras et al. (1985), and Spanos and Schultz (1985 and 1986). These efforts reflect the feasibility of adapting system identification techniques (Mullis and Roberts, 1976; Gersch and Liu, 1976; Wang and Fang, 1986) to system simulation techniques. Furthermore, they illuminate some of the intrinsic features of their applicability to vibration problems.

In the present series of articles a unified approach is presented for developing ARMA simulation algorithms which are based on a prior approximation of a target spectral matrix by the response of an AR discrete dynamic system to white noise excitation. Existing procedures are briefly reviewed and new ones are presented. Further, the properties of these procedures are thoroughly analyzed. Finally, their applicability is exemplified by considering spectra of a variety of natural processes.

The first part of this series focuses on the formulation of various procedures. The system of equations leading to the determination of the ARMA coefficients are derived through the minimization of frequency domain errors. The interrelationships and advantages of these procedures are investigated as well.

In the second part, the procedures are further analyzed by assessing the matching of the auto and cross-correlations of the target and the simulated processes. Further, properties such as stability and invertibility are investigated. Finally, implementation aspects of these procedures are reviewed and examples of applications are given.

Autoregressive Approximation

The ARMA synthesis procedures described in the ensuing section require that the target process be first approximated in terms of an autoregressive process. For this reason and for the sake of completeness, the determination and the properties of an AR system are reviewed.

An n-variate autoregressive (AR) process $\hat{\mathbf{Y}}$ of order m is a discrete stochastic vector process whose rth sample can be computed from the m previous ones in the following manner

$$\hat{\mathbf{Y}}_r = -\sum_{k=1}^{m} \hat{A}_k \hat{\mathbf{Y}}_{r-k} + \hat{B}_o \mathbf{W}_r, \tag{1}$$

where $\hat{A}_k$ and $\hat{B}_o$ are real $n \times n$ matrices. The symbol $\mathbf{W}$ denotes an n-variate band limited in the interval $[-\omega_b, \omega_b]$ white noise vector process. The autocorrelation matrix of $\mathbf{W}$ is

[1]Currently, Assistant Professor, Dept. of Mechanical and Aerospace Engineering, Arizona State University, Tempe, AZ 85287.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13-18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 1, 1986; final revision March 27, 1987.

Paper No. 87-WA/APM-12.

$$E[\mathbf{W}_i \mathbf{W}_j^\dagger] = 2\omega_b I_n \delta_{ij}, \tag{2}$$

where $E[.]$ and $[.]^\dagger$ are the operators of mathematical expectation and transposition. The symbols I_n and δ_{ij} denote the $n \times n$ identity matrix and the Kronecker delta, respectively. The sampling period T and the cutoff frequency ω_b are related through the Nyquist relation

$$T = \frac{\pi}{\omega_b}. \tag{3}$$

Clearly, the process defined by equation (1) can be considered as the response vector to white noise excitation of a multi-degree-of-freedom discrete dynamic system. Its transfer function matrix in terms of z-transform notation is

$$H(z) = \hat{D}^{-1}(z)\hat{B}_o, \tag{4}$$

where

$$\hat{D}(z) = I_n + \sum_{k=1}^{m} \hat{A}_k z^{-k}. \tag{5}$$

The spectral matrix of $\hat{\mathbf{Y}}$ is given by the equation

$$S_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(\omega) = H^*(e^{j\omega T}) H^\dagger(e^{j\omega T}) \tag{6}$$

where $[.]^*$ designates complex conjugation.

Given an arbitrary spectral matrix $S_{\mathbf{YY}}(\omega)$ (target), the AR simulation problem involves the determination of the coefficients $\hat{A}_k$ and $\hat{B}_o$ so that $S_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(\omega)$ is close in some sense to $S_{\mathbf{YY}}(\omega)$. It has been shown (Hannan, 1970), that a meaningful measure of the error is

$$\epsilon_{AR} = \frac{1}{2\omega_b} \int_{-\omega_b}^{\omega_b} |\hat{B}_o^{-1} \hat{D}(e^{j\omega T}) Q(\omega)|^2 d\omega \tag{7}$$

where $Q(\omega)$ is a causal transfer function such that

$$S_{\mathbf{YY}}(\omega) = Q^*(\omega) Q^\dagger(\omega). \tag{8}$$

The symbol $|U|$ signifies the Euclidean norm of an arbitrary matrix U of elements u_{ij}. That is,

$$|U|^2 = tr(U^* U^\dagger) = \sum_{i=1}^{n} \sum_{j=1}^{n} u_{ij} u_{ij}^*. \tag{9}$$

Introducing the autocorrelation function of the target process

$$R_{\mathbf{YY}}(k) = E[\mathbf{Y}_r \mathbf{Y}_{r+k}^\dagger] = \int_{-\omega_b}^{\omega_b} S_{\mathbf{YY}}(\omega) e^{jk\omega T} d\omega \tag{10}$$

it is readily shown that equation (7) can be rewritten as

$$\epsilon_{AR} = \frac{1}{2\omega_b} tr\left\{ \hat{B}_o^{-1} \left(\sum_{k=0}^{m} \sum_{\ell=0}^{m} \hat{A}_k R_{\mathbf{YY}}(k-\ell) \right) \hat{A}_\ell^\dagger / \hat{B}_o^{-\dagger} \right\} \tag{11}$$

where $\hat{A}_o = I_n$, and the symbol $(\bullet^{-\dagger})$ denotes the transpose of the inverse of a matrix. The minimum error is obtained when the coefficients A_k satisfy the following equations (Yule-Walker equations)

$$R_{\mathbf{YY}}^\dagger(\ell) + \sum_{k=1}^{m} \hat{A}_k R_{\mathbf{YY}}(k-\ell) = 0 \quad \ell = 1, \ldots, m. \tag{12}$$

The matrix $\hat{B}_o$ is obtained by equating the total "energies" of the target and AR processes. That is,

$$\int_{-\omega_b}^{\omega_b} \hat{D}^{-*}(e^{j\omega T}) \hat{B}_o \hat{B}_o^\dagger \hat{D}^{-\dagger}(e^{j\omega T}) d\omega = \int_{-\omega_b}^{\omega_b} S_{\mathbf{YY}}(\omega) d\omega \tag{13}$$

or

$$R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(0) = R_{\mathbf{YY}}(0). \tag{14}$$

From equation (1) it is readily shown that the autocorrelation function of the AR process satisfies the following recurrence relation

$$R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}^\dagger(\ell) + \sum_{k=1}^{m} \hat{A}_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k-\ell) = \hat{B}_o R_{\hat{\mathbf{Y}}\mathbf{W}}^\dagger(\ell) \quad \ell = 0, \pm 1, \pm 2, \ldots \tag{15}$$

where $R_{\hat{\mathbf{Y}}\mathbf{W}}$ is the input-output crosscorrelation function defined by

$$R_{\hat{\mathbf{Y}}\mathbf{W}}(\ell) = 0 \qquad \ell > 0 \tag{16a}$$

$$R_{\hat{\mathbf{Y}}\mathbf{W}}(\ell) = 2\omega_b \hat{B}_o \qquad \ell = 0 \tag{16b}$$

$$R_{\hat{\mathbf{Y}}\mathbf{W}}(\ell) = -\sum_{k=1}^{\min(m, -\ell)} \hat{A}_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k+\ell) \qquad \ell < 0 \tag{16c}$$

It is readily seen that equations (12), (14), and (15) admit the solution

$$R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k) = R_{\mathbf{YY}}(k) \quad k = 0, 1, \ldots, m. \tag{17}$$

It will be shown in Part II that this solution is unique. The matrix $\hat{B}_o$ is then obtained from equations (15) and (16b)

$$\hat{B}_o \hat{B}_o^\dagger = \frac{1}{2\omega_b} \left\{ R_{\mathbf{YY}}(0) + \sum_{k=1}^{m} \hat{A}_k R_{\mathbf{YY}}(k) \right\}. \tag{18}$$

Note that the right-hand side of this equation must be positive definite. This relation does not yield a unique solution for $\hat{B}_o$. Indeed, it can be seen as a set of $n(n+1)/2$ independent nonlinear equations in the n^2 elements of $\hat{B}_o$. It can be shown (Mignolet, 1987) that the quality of matching of the target spectral matrix by the AR and the subsequent ARMA approximations is independent of the form chosen for $\hat{B}_o$, as long as equation (18) is satisfied. To ease the computations, $\hat{B}_o$ is assumed to be lower triangular so that it is obtained through the Cholesky decomposition of equation (18).

It should be noted that the corresponding minimum value of ϵ_{AR}, always equal to n, does not provide a readily discernible measure of the quality of the AR approximation over the entire interval $[-\omega_b, \omega_b]$. Such a measure can, however, be introduced. It is proposed to monitor the closeness of the matrix $\hat{B}_o$ to its asymptotic, $m = \infty$, form. Specifically, since

$$\epsilon_m = \det(\hat{B}_o \hat{B}_o^\dagger) - \exp\left\{ \frac{1}{2\omega_b} \int_{-\omega_b}^{\omega_b} tr[\log S_{\mathbf{YY}}(\omega)] d\omega \right\} \tag{19}$$

decreases monotonically to zero as $m \to \infty$ (Hannan 1970) it reflects in a natural manner the quality of approximation by becoming equal to zero when the spectral matching is exact. It should be noted that if

$$\int_{-\omega_b}^{\omega_b} tr[\log S_{\mathbf{YY}}(\omega)] d\omega = -\infty, \tag{20}$$

then

$$\det(\hat{B}_o) \to 0 \text{ as } m \to \infty. \tag{21}$$

If, in addition to satisfying equation (20), $S_{\mathbf{YY}}(\omega)$ is such that

$$\det(S_{\mathbf{YY}}(\omega)) \neq 0 \text{ on } \Omega \subseteq [-\omega_b, \omega_b] \tag{22}$$

then

$$\det[\hat{D}(e^{j\omega T})] = 0 \text{ almost everywhere on } \Omega \tag{23}$$

This can be seen from the following equivalent form of equation (18)

$$2\omega_b \hat{B}_o \hat{B}_o^\dagger = \oint_{-\omega_b}^{\omega_b} \hat{D}^*(e^{j\omega T}) S_{\mathbf{YY}}(\omega) \hat{D}^\dagger(e^{j\omega T}) d\omega. \tag{24}$$

That is, the AR spectrum possesses densely spaced poles on the circumference of the unit circle of the complex domain, so that the spectrum of any finite AR approximation involves a series of peaks modulated by a small value of $\det \hat{B}_o$. This phenomenon and related problems have already been noted in the context of univariate simulation (Spanos and Mignolet, 1986).

The AR system is a special case of the ARMA class of systems whose transfer function is a rational function of z^{-1}.

Clearly, the transfer function of an ARMA system has both zeros and poles and offers the most versatility in generating recursively realizations of random processes with specified spectral matrices. Several procedures of determining low-order ARMA approximations by relying on an initial high order AR approximation of a given spectral matrix will be discussed in the ensuing sections.

Autoregressive Moving Average (ARMA) Approximation—Original Spectral Matrix

Preliminary Remarks. An n-variate ARMA (p, q) process $\hat{\mathbf{Y}}$ is a discrete random vector process whose rth sample can be obtained from the previous ones in the following manner

$$A_o\hat{\mathbf{Y}}_r = -\sum_{k=1}^{p} A_k\hat{\mathbf{Y}}_{r-k} + \sum_{\ell=0}^{q} B_\ell \mathbf{W}_{r-\ell} \tag{25}$$

where $\mathbf{W}$ is defined by equation (2). The process $\hat{\mathbf{Y}}$ can also be considered as the output of a multi-degree-of-freedom discrete dynamic system whose transfer function matrix is

$$H(z) = D^{-1}(z)N(z) \tag{26}$$

with

$$D(z) = \sum_{k=0}^{p} A_k z^{-k} \tag{27}$$

and

$$N(z) = \sum_{\ell=0}^{q} B_\ell z^{-\ell}. \tag{28}$$

The following relations involving the input-output crosscorrelations $R_{\hat{\mathbf{Y}}\mathbf{W}}(k)$ and the output autocorrelations $R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k)$ are readily established

$$\sum_{k=0}^{p} A_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k-i) = \sum_{\ell=0}^{q} B_\ell R^{\dagger}_{\hat{\mathbf{Y}}W}(i-\ell) \quad i=0, \pm 1, \pm 2, \ldots \tag{29}$$

and

$$R_{\hat{\mathbf{Y}}\mathbf{W}}(-i) = 0 \qquad i<0 \tag{30a}$$

$$\sum_{k=0}^{\min(i,p)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i) = 2\omega_b B_i \qquad i=0, \ldots, q \tag{30b}$$

$$\sum_{k=0}^{\min(i,p)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i) = 0 \qquad i>q. \tag{30c}$$

For simulation, the corresponding spectral matrix

$$S_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(\omega) = H^*(e^{j\omega T})H^{\dagger}(e^{j\omega T}) \tag{31}$$

must represent a good approximation of a target expression $S_{\mathbf{YY}}(\omega)$. The quality of the matching of $S_{\mathbf{YY}}(\omega)$ by $S_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(\omega)$ is quantified by the error

$$\epsilon = \frac{1}{2\omega_b}\int_{-\omega_b}^{\omega_b} |D(e^{j\omega T})Q(\omega) - N(e^{j\omega T})|^2 d\omega, \tag{32}$$

where $Q(\omega)$ has been defined by equation (8). Two procedures to determine the coefficients A_k and B_ℓ by minimizing ϵ under various conditions are presented. These procedures will be subsequently recognized as special cases of a general class of solutions.

First, $Q(\omega)$ is approximated as

$$Q(\omega) = H_{AR}(e^{j\omega T}), \tag{33}$$

where $H_{AR}(e^{j\omega T})$ is the transfer function of the AR system determined previously.

Equation (32) can then be rewritten as

$$\epsilon = tr(E), \tag{34}$$

where

$$E = \frac{1}{2\omega_b}\Bigg\{\sum_{k=0}^{p}\sum_{i+0}^{p} A_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k-i)A_i^{\dagger} - \sum_{k=0}^{p}\sum_{\ell=0}^{q} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell)B_\ell^{\dagger} - \sum_{k=0}^{p}\sum_{\ell=0}^{q} B_\ell R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell)A_k^{\dagger} + 2\omega_b\sum_{\ell=0}^{q} B_\ell B_\ell^{\dagger}\Bigg\}, \tag{35}$$

and

$$R_{\hat{\mathbf{Y}}\mathbf{W}}(k) = \int_{-\omega_b}^{\omega_b} H_{AR}(e^{j\omega T})e^{-jk\omega T}d\omega \tag{36}$$

is the input-output crosscorrelation of the AR system; it is readily computed by using equations (16a)–(16c).

Auto Cross-Correlations Matching Procedure (ACM). This procedure involves a constrained minimization of ϵ with respect to the coefficients A_k and B_ℓ. Clearly, the minimization of ϵ with respect to A_k and B_ℓ for all values of k and ℓ would yield the trivial solution with

$$A_k = B_\ell = 0 \quad k=0, \ldots, p \text{ and } \ell=0, \ldots, q. \tag{37}$$

Thus, a nonhomogeneous constraint must be added to the problem. An obvious choice is

$$A_o = I_n. \tag{38}$$

The minimization of the error requires setting the derivatives of ϵ with respect to A_k $(k = 1, \ldots, p)$ and $B_\ell(\ell = 0, \ldots, q)$ equal to zero. This condition yields the following set of linear equations

$$\sum_{k=0}^{p} A_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k-i) - \sum_{\ell=0}^{q} B_\ell R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(i-\ell) = 0 \quad i=1, \ldots, p \tag{39}$$

and

$$2\omega_b B_i = \sum_{k=0}^{\min(i,p)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i) \quad i=0,1, \ldots, q. \tag{40}$$

Combining equations (35) and (38)–(40), the corresponding minimum value of ϵ becomes

$$\hat{\epsilon}_{\min} = tr(\hat{E}_{\min}), \tag{41}$$

where

$$\hat{E}_{\min} = \frac{1}{2\omega_b}\Bigg\{\sum_{k=0}^{p} A_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k) - \sum_{\ell=0}^{q} B_\ell R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(-\ell)\Bigg\}. \tag{42}$$

The choice expressed by equation (38) is convenient but arbitrary and might not give an absolute minimum for ϵ. Another possibility is

$$\bar{B}_o = B_o = \hat{B}_o = \frac{1}{2\omega_b} R_{\hat{\mathbf{Y}}\mathbf{W}}(0). \tag{43}$$

Then, the optimal coefficients $\bar{A}_k$ and $\bar{B}_\ell$ can be computed by setting the derivatives of ϵ with respect to A_k $(k = 0, \ldots, p)$ and $B_\ell(\ell = 1, \ldots, q)$ to zero. The following equations are obtained

$$\sum_{k=0}^{p} \bar{A}_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k-i) - \sum_{\ell=0}^{q} \bar{B}_\ell R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(i-\ell) = 0 \quad i=0, \ldots, p \tag{44}$$

and

$$2\omega_b \bar{B}_i = \sum_{k=0}^{\min(i,p)} \bar{A}_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i) \quad i=1, \ldots, q. \tag{45}$$

Note that the matrix $\bar{B}_o$ appears only when $i = 0$ in equation (44). Thus, the systems (39)–(40) and (44)–(45) are identical except in the $i = 0$ case. A solution of the form

$$\bar{A}_k = \bar{A}_o A_o^{-1} A_k = \bar{A}_o A_k \qquad k = 0, \ldots, p \tag{46}$$

$$\bar{B}_\ell = \bar{A}_o A_o^{-1} B_\ell = \bar{A}_o B_\ell \qquad \ell = 1, \ldots, q \tag{47}$$

to equations (44)–(45) is then possible. The unknown $\bar{A}_o$ is readily obtained by satisfying equation (44) for $i = 0$. Namely

$$\bar{A}_o \Big\{ \sum_{k=0}^{p} A_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k) - \sum_{\ell=1}^{q} B_\ell R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(-\ell) \Big\} = \bar{B}_o R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(0). \tag{48}$$

Taking into account equations (42) and (43), equation (48) can be rewritten as

$$\bar{A}_o (\hat{E}_{\min} + B_o B_o^{\dagger}) = B_o B_o^{\dagger}. \tag{49}$$

The corresponding minimum value of ϵ is $\bar{\epsilon}_{\min} = tr(\bar{E}_{\min})$, where $\bar{E}_{\min}$ is readily obtained from equation (35)

$$\bar{E}_{\min} = B_o B_o^{\dagger} - \frac{1}{2\omega_b} \sum_{k=0}^{p} \bar{A}_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k) B_o^{\dagger}. \tag{50}$$

Combining this expression and the causality condition, equation (16*a*), yields

$$\bar{E}_{\min} = (I_n - \bar{A}_o) B_o B_o^{\dagger}. \tag{51}$$

It will be shown in Part II that the preceding procedure yields an ARMA process which exhibits certain auto-cross correlation matching properties with regard to the target process. This observation accounts for the proposed abbreviation, ACM.

Power Order Matching (POM) Procedure. Note the similarity between ϵ as defined by equation (32) and the error

$$\epsilon_F = \frac{1}{2\omega_b} \int_{-\omega_b}^{\omega_b} |F(\omega) - \sum_{n=-\infty}^{\infty} C_n e^{-jn\omega T}|^2 d\omega \tag{52}$$

which is involved in the Fourier representation of an arbitrary matrix function $F(\omega)$ in terms of its coefficients C_n. In fact, from equation (40), the parameters B_ℓ appear as the Fourier coefficients of $D(e^{j\omega T})\, H_{AR}(e^{j\omega T})$. Note also that coefficients corresponding to positive powers of $e^{j\omega T}$ vanish due to the causality of $D(e^{j\omega T})\, H_{AR}(e^{j\omega T})$. On the basis of this remark, an alternative ARMA synthesis procedure can be conceived. It treats $N(e^{j\omega T})$ as a Fourier approximation of $D(e^{j\omega T})\, H_{AR}(e^{j\omega T})$ of highest order possible $p + q$, whose coefficients

$$B_{q+1} = \ldots = B_{q+p} = 0. \tag{53}$$

In other words, A_K are selected so that $D(e^{j\omega T})\, H_{AR}(e^{j\omega T})$ has Fourier coefficients of order $q + 1$ to $q + p$ equal to zero, and B_ℓ are selected so that $D(e^{j\omega T})\, H_{AR}(e^{j\omega T})$ and $N(e^{j\omega T})$ have identical Fourier coefficients of order 0 to q. For the previous interpretation to hold, the coefficients B_ℓ must vary independently so that the traditional nonhomogeneous constraint, equation (38), is enforced.

It is readily shown that the coefficients A_k and B_ℓ must satisfy the following equations

$$\sum_{k=0}^{\min(p,\ell)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell) = 2\omega_b B_\ell \qquad \ell = 0, \ldots, q \tag{54}$$

and

$$\sum_{k=0}^{\min(p,\ell)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell) = 0 \qquad \ell = q+1, \ldots, q+p \tag{55}$$

The error associated with this procedure can be computed from equation (52) as

$$\overset{*}{\epsilon}_{\min} = \frac{1}{4\omega_b^2} \sum_{\ell=q+p+1}^{\infty} |\sum_{k=0}^{p} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell)|^2. \tag{56}$$

This procedure can equivalently be considered as a minimization of ϵ, defined by equation (32), subject to the constraints specified by equations (38) and (55). Thus, using equations (35) and (54), the corresponding minimum value of ϵ can be rewritten as

$$\overset{*}{\epsilon}_{\min} = \frac{1}{2\omega_b} \sum_{k=0}^{p} tr\Big\{ A_k \Big[\sum_{i=0}^{p} R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k-i) A_i^{\dagger} - \sum_{\ell=0}^{q} R_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell) B_\ell^{\dagger} \Big] \Big\}. \tag{57}$$

Note that the equivalence of the two forms (56) and (57) is readily shown by relying on the identity

$$R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k) = \frac{1}{2\omega_b} \sum_{\ell=0}^{\infty} R_{\hat{\mathbf{Y}}\mathbf{W}}(-\ell) R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(-k-\ell) \text{ for all } k. \tag{58}$$

It will be shown in Part II that the preceding procedure yields an ARMA system whose transfer function can be obtained from the corresponding description of the AR system by equating the coefficients of powers of z^{-1}. This observation accounts for the proposed abbreviation, POM.

Generalization. In this section a general class of procedures including as special cases ACM and POM is introduced.

A class of ARMA systems synthesis procedures can be obtained by minimizing ϵ as defined by equation (32) under the constraints

$$A_o = I_n \tag{59}$$

and

$$\sum_{k=0}^{\min(p,\ell)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell) = 0 \qquad \ell = q+1, \ldots, q+r. \tag{60}$$

Clearly, if $r = 0$ no constraint of the kind specified by equation (60) is to be satisfied, and the traditional ACM procedure is represented. Further, the POM procedure is derived for $r = p$.

The problem can be restated as the minimization of

$$\tilde{\epsilon} = \epsilon + \frac{1}{\omega_b} tr\Big\{ \sum_{l=q+1}^{q+r} \Big[\sum_{k=0}^{\min(p,\ell)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell) \Big] \Lambda_\ell \Big\} \tag{61}$$

with respect to the coefficients A_k and B_ℓ with the exception of A_o. In minimizing $\tilde{\epsilon}$ it is readily found that the associated $n \times n$ Lagrange multiplier matrices Λ_ℓ and the matrices A_k and B_ℓ are solutions of the system of equations

$$\sum_{k=0}^{p} A_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k-i) - \sum_{\ell=0}^{q} B_\ell R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(i-\ell) + \sum_{\ell=q+1}^{q+r} \Lambda_\ell^{\dagger} R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(i-\ell) = 0 \qquad i = 1, \ldots, p \tag{62}$$

$$\sum_{k=0}^{\min(p,\ell)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-\ell) = 2\omega_b B_\ell \qquad \ell = 0, \ldots, q \tag{63}$$

with equations (59) and (60) appended. The corresponding minimum value of the error is

$$\tilde{\epsilon}_{\min} = tr(\tilde{E}_{\min}) \tag{64}$$

where $\tilde{E}_{\min}$ is the symmetric matrix

$$\tilde{E}_{\min} = \frac{1}{2\omega_b} \Big\{ \sum_{k=0}^{p} A_k R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k) - \sum_{\ell=0}^{q} B_\ell R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(-\ell) + \sum_{\ell=q+1}^{q+r} \Lambda_\ell^{\dagger} R^{\dagger}_{\hat{\mathbf{Y}}\mathbf{W}}(-\ell) \Big\}. \tag{65}$$

Autoregressive Moving Average (ARMA) Approximation—Inverse Spectral Matrix

Preliminary Remarks. Clearly, the transfer function matrix of an ARMA (p, q) system exhibits the interesting property that its inverse is also the transfer function of an ARMA (q, p) system. Thus, alternative ARMA approximation procedures based on the inverse transfer function can be obtained as duals of the previous ones.

In this section, the quality of the ARMA approximation will be measured by the error

$$\delta=\frac{1}{2\omega_b}\int_{-\omega_b}^{\omega_b} |N(e^{j\omega T})Q^{-1}(\omega)-D(e^{j\omega T})|^2 d\omega. \tag{66}$$

Auto/Cross-Correlation Matching Method. This procedure is based on the minimization of δ with respect to the parameters A_k and B_ℓ subject to the nonhomogeneous constraint

$$B_o=\hat{B}_o. \tag{67}$$

Equation (66) can be rewritten in the form

$$\delta=tr(\Delta), \tag{68}$$

where

$$\Delta=\sum_{\ell=0}^{q}\sum_{i=0}^{q} B_\ell\hat{B}_o^{-1}\Big(\sum_{u=\max(\ell,i)}^{m+\min(\ell,i)}\hat{A}_{u-\ell}\hat{A}_{u-i}^{\dagger}\Big)\hat{B}_o^{-\dagger}\mathbf{B}_i^{\dagger}$$

$$-\sum_{k=0}^{p}\sum_{\ell=0}^{\min(q,k)} B_\ell\hat{B}_o^{-1}\hat{A}_{k-\ell}A_k^{\dagger}$$

$$-\sum_{k=0}^{p}\sum_{\ell=0}^{\min(q,k)} A_k\hat{A}_{k-\ell}^{\dagger}\hat{B}_o^{-\dagger}B_\ell^{\dagger}+\sum_{k=0}^{p}A_kA_k^{\dagger}. \tag{69}$$

It is readily shown that the minimum of δ is attained when the matrices A_k and B_ℓ are solutions of the following system of equations

$$\sum_{\ell=0}^{\min(q,k)} B_\ell\hat{B}_o^{-1}\hat{A}_{k-\ell}=A_k \quad k=0,\ldots,p \tag{70}$$

and

$$\sum_{\ell=0}^{q} B_\ell\hat{B}_o^{-1}\Big(\sum_{u=\max(\ell,i)}^{m+\min(\ell,i)}\hat{A}_{u-\ell}\hat{A}_{u-i}^{\dagger}\Big) = \sum_{k=i}^{p}A_k\hat{A}_{k-i}^{\dagger} \quad i=1,\ldots,q. \tag{71}$$

The corresponding minimum value of δ is

$$\hat{\delta}_{\min}=tr(\hat{\Delta}_{\min}) \tag{72}$$

where

$$\hat{\Delta}_{\min}=\sum_{\ell=0}^{q} B_\ell\hat{B}_o^{-1}\Big(\sum_{u=\ell}^{m}\hat{A}_{u-\ell}\hat{A}_u^{\dagger}\Big)-\sum_{k=0}^{p}A_k\hat{A}_k^{\dagger}. \tag{73}$$

As previously, a modified approach based on imposing a constraint on $\bar{A}_o$ can be developed. In fact, the value of $\bar{A}_o$ will equal the value A_o of the previous approach. Using equation (70) for $k = 0$, this requirement is equivalent to

$$\bar{A}_o=A_o=I_n. \tag{74}$$

The equations for the coefficients $\bar{A}_k$ and $\bar{B}_\ell$ are then derived by setting equal to zero the derivatives of δ with respect to A_k $(k = 1, \ldots, p)$ and B_ℓ $(\ell = 0, \ldots, q)$; one obtains

$$\sum_{\ell=0}^{\min(q,k)} \bar{B}_\ell\hat{B}_o^{-1}\hat{A}_{k-\ell}=\bar{A}_k \quad k=1,\ldots,p \tag{75}$$

and

$$\sum_{\ell=0}^{q} \bar{B}_\ell\hat{B}_o^{-1}\Big(\sum_{u=\max(\ell,i)}^{m+\min(\ell,i)}\hat{A}_{u-\ell}\hat{A}_{u-i}^{\dagger}\Big) = \sum_{k=i}^{p}\bar{A}_k\hat{A}_{k-i}^{\dagger} \quad i=0,\ldots,q. \tag{76}$$

The resulting minimum value of the error δ is readily computed as

$$\bar{\delta}_{\min}=tr(\bar{\Delta}_{\min}) \tag{77}$$

where

$$\bar{\Delta}_{\min}=\bar{A}_o\bar{A}_o^{\dagger}-\bar{B}_o\hat{B}_o^{-1}\hat{A}_o\bar{A}_o^{\dagger}=I_n-\bar{B}_o\hat{B}_o^{-1}. \tag{78}$$

The two sets of coefficients can be computed from one another, as in the approach based on the original spectral matrix, by noting that equations (75) and (76) admit the solution

$$\bar{A}_k=\bar{B}_o\hat{B}_o^{-1}A_k \qquad k=1,\ldots,p \tag{79}$$

$$\bar{B}_\ell=\bar{B}_o\hat{B}_o^{-1}B_\ell \qquad \ell=0,\ldots,q \tag{80}$$

provided that

$$\bar{B}_o\hat{B}_o^{-1}\Big[\sum_{\ell=0}^{q}B_\ell\hat{B}_o^{-1}\Big(\sum_{u=\ell}^{m}\hat{A}_{u-\ell}\hat{A}_u^{\dagger}\Big)-\sum_{k=1}^{p}A_k\hat{A}_k^{\dagger}\Big]=I_n. \tag{81}$$

This condition can be rewritten in terms of $\hat{\Delta}_{\min}$ and $\bar{\Delta}_{\min}$ as follows

$$\bar{B}_o\hat{B}_o^{-1}\hat{\Delta}_{\min}=\bar{\Delta}_{\min}. \tag{82}$$

Power Order Matching (POM) Method. This ARMA system synthesis procedure involves the minimization of δ as defined by equation (66) with respect to the coefficients of $D(e^{j\omega T})$. The coefficients of $N(e^{j\omega T})$ are such that equation (67) is satisfied and

$$A_{p+i}=0 \quad i=1,\ldots,q. \tag{83}$$

It is readily shown that the coefficients A_k and B_ℓ must satisfy the following equations

$$\sum_{\ell=\max(0,k-m)}^{\min(q,k)} B_\ell\hat{B}_o^{-1}\hat{A}_{k-\ell}=A_k \quad k=0,\ldots,p \tag{84}$$

and

$$\sum_{\ell=\max(0,k-m)}^{\min(q,k)} B_\ell\hat{B}_o^{-1}\hat{A}_{k-\ell}=0 \quad k=p+1,\ldots,p+q. \tag{85}$$

Note that taking into account equation (67), equation (84) for $k = 0$ renders equation (38). Next equations (85) are solved to yield the values of the parameters B_ℓ. Finally, the coefficients A_k are computed by relying on equations (84). The corresponding value of the error can be computed as in the procedure involving the original target matrix,

$$\overset{*}{\delta}_{\min}=tr(\overset{*}{\Delta}_{\min}), \tag{86}$$

where

$$\overset{*}{\Delta}_{\min}=\sum_{k=p+q+1}^{m+q}\Big[\Big(\sum_{\ell=\max(0,k-m)}^{q}B_\ell\hat{B}_o^{-1}\hat{A}_{k-\ell}\Big)$$

$$\Big(\sum_{\ell=\max(0,k-m)}^{q}B_\ell\hat{B}_o^{-1}\hat{A}_{k-\ell}\Big)^{\dagger}\Big]. \tag{87}$$

Postmultipyling equations (84) and (85) by $R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i)$ and summing over k from 0 to i yield

$$\sum_{k=0}^{i}\sum_{\ell=\max(0,k-m)}^{\min(q,k)} B_\ell \hat{B}_o^{-1}\hat{A}_{k-\ell}R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i)$$

$$= \sum_{k=0}^{\min(i,p)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i) \tag{88}$$

or

$$\sum_{\ell=0}^{\min(i,q)} B_\ell \hat{B}_o^{-1}\left\{\sum_{u=0}^{\min(m,i-\ell)} \hat{A}_u R_{\hat{\mathbf{Y}}\mathbf{W}}[u-(i-\ell)]\right\}$$

$$= \sum_{k=0}^{\min(i,p)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i). \tag{89}$$

Relying on equations (16), these equations can be rewritten as

$$\sum_{k=0}^{\min(i,p)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i) = 2\omega_b B_i \quad i=0, \ldots, q \tag{90}$$

and

$$\sum_{k=0}^{\min(i,p)} A_k R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i) = 0 \quad i=q+1, \ldots, p+q \tag{91}$$

which are identical to equations (54) and (55). Thus, the POM original spectral matrix and inverse spectral matrix approaches yield the identical systems.

Generalization. The two ACM and POM procedures for the inverse spectral matrix can be seen as particular cases of the general class of algorithms obtained by minimizing

$$\tilde{\delta} = \frac{1}{2\omega_b}\int_{-\omega_b}^{\omega_b} |N(e^{j\omega T})\hat{B}_o^{-1}\hat{D}(e^{j\omega T}) - D(e^{j\omega T})|^2 d\omega$$

$$+ 2\, tr\left\{\sum_{k=p+1}^{p+r}\left[\sum_{\ell=\max(0,k-m)}^{\min(q,k)} B_\ell \hat{B}_o^{-1}\hat{A}_{k-\ell}\right]\Lambda_k\right\} \tag{92}$$

with respect to the ARMA coefficients A_k, B_ℓ and the Lagrange multipliers Λ_k. Imposing the nonhomogeneous condition specified by equation (67), it is readily shown that the unknown coefficients are solutions of

$$\sum_{\ell=0}^{q} B_\ell \hat{B}_o^{-1}\left(\sum_{u=\max(\ell,i)}^{m+\min(\ell,i)} \hat{A}_{u-\ell}\hat{A}_{u-i}^{\dagger}\right) - \sum_{k=i}^{\min(p,m+i)} A_k \tilde{A}_{k-i}^{\dagger}$$

$$+ \sum_{k=\max(p+1,i)}^{\min(p+r,m+i)} \Lambda_k^{\dagger}\hat{A}_{k-i}^{\dagger} = 0$$

$$i=1, \ldots, q \tag{93}$$

$$\sum_{\ell=\max(0,k-m)}^{\min(q,k)} B_\ell \hat{B}_o^{-1}\hat{A}_{k-\ell} = 0 \quad k=p+1, \ldots, p+r \tag{94}$$

with equations (67) and (84) appended. The corresponding minimum value of $\tilde{\delta}$ is

$$\tilde{\delta}_{\min} = tr(\tilde{\Delta}_{\min}) \tag{95}$$

where

$$\tilde{\Delta}_{\min} = \sum_{\ell=0}^{q} B_\ell \hat{B}_o^{-1}\left(\sum_{u=\ell}^{m} \hat{A}_{u-\ell}\hat{A}_u^{\dagger}\right) - \sum_{k=0}^{\min(p,m)} A_k \hat{A}_k^{\dagger}$$

$$+ \sum_{k=p+1}^{\min(p+r,m)} \Lambda_k^{\dagger}\hat{A}_k^{\dagger}. \tag{96}$$

Critical Assessment

Comparison of Original and Inverse Spectral Matrix Procedures. The described procedures which are based on the inverse spectral matrix, have some computational advantages over the corresponding procedures which rely on the original spectral matrix. Specifically, the original ACM and POM procedures require the solution of a set of $p \times n$ equations, while their inverse counterparts involve systems of size $q \times n$. Note that since the first approximation of the target process is provided by a pure autoregression, q is expected to be lower or equal to p. In addition to the reduction of the linear system size, it can be seen that the recursive computation of the crosscorrelations by means of equations (16) is not required. When selecting the appropriate procedure, original or inverse spectral matrix, for a particular problem, the POM procedure requires the solution of only $\min(p, q) \times n$ simultaneous equations.

Comparison of the A_o or B_o Constrained ACM Procedure. It can be shown from equations (49) and (51) that the errors $\bar{\epsilon}_{\min}$ and $\hat{\epsilon}_{\min}$ satisfy the inequality (Mignolet, 1987)

$$\hat{\epsilon}_{\min} - \bar{\epsilon}_{\min} - tr[\hat{E}_{\min}(\hat{E}_{\min} + B_o B_o^{\dagger})^{-1}\hat{E}_{\min}] \geq 0. \tag{97}$$

That is, the improvement of the B_o constrained procedure over the one retaining $A_o = I_n$, namely $\hat{\epsilon}_{\min} - \bar{\epsilon}_{\min}$, is a second order term in $\hat{E}_{\min}$. Thus, when a good matching, $\hat{\epsilon}_{\min} \to 0$, is obtained, the two procedures give almost identical results. If, however, the A_o based procedure is not satisfactory, the alternative procedure represents an improvement quantified by $tr[\hat{E}_{\min}(\hat{E}_{\min} + B_o B_o^{\dagger})^{-1}\hat{E}_{\min}]$. Note that a similar relation between the errors $\hat{\Delta}_{\min}$ and $\bar{\Delta}_{\min}$ corresponding to the inverse spectral matrix procedure can be derived.

In the context of simulation, the quality of the matching between the target and the approximate spectral matrices can be better described by the "visual" error e_b defined as (Mignolet, 1987)

$$e_b = \frac{1}{2\omega_b}\left\{\int_{-\omega_b}^{\omega_b} |A_o^{-1}[D(e^{j\omega T})Q(\omega) - N(e^{j\omega T})]|^2 d\omega\right\}$$

$$\left\{\max_{|\omega|\leq\omega_b} |D^{-1}(e^{j\omega T})A_o|^2\right\}. \tag{98}$$

It can be shown (Mignolet, 1987) that the values $\hat{e}_b$ and $\bar{e}_b$ corresponding to the A_o and B_o constrained procedures satisfy the inequality

$$\bar{e}_b - \hat{e}_b = \beta\, tr[\hat{E}_{\min}\bar{E}_{\min}^{-1}(\hat{E}_{\min} - \bar{E}_{\min})] \geq 0. \tag{99}$$

where

$$\beta = \max_{|\omega|\leq\omega_b} |D^{-1}(e^{j\omega T})|^2 = \max_{|\omega|\leq\omega_b} |\bar{D}^{-1}(e^{j\omega T})\bar{A}_o|^2. \tag{100}$$

That is, the B_o constrained procedure yields a larger visual error than the one which is based on constraining A_o. Thus, in the context of the present procedures, there exists a trade-off between an absolute minimum of ϵ and a small value of e_b as defined by equations (32) and (98), respectively.

Concluding Remarks

The simulation of a multivariate random process with a specified (target) spectral matrix as the output to white noise input of autoregressive (AR) and autoregressive moving average (ARMA) discrete systems has been studied. The contributions of this investigation may be summarized as follows.

1. A meaningful measure of the quality of the matching between the target and the AR spectra was introduced by equation (19). Furthermore, the analysis of this relation called attention to the existence of some pathological target spectra,

namely those satisfying equations (20) and (22). Therefore, given an arbitrary spectral matrix, it should not unreservedly be assumed that a reliable AR approximation can always be constructed. This point has been demonstrated in the context of ocean waves spectra by Spanos and Mignolet (1986).

2. Two procedures, ACM and POM, for determining the coefficients of an ARMA approximation from an initial AR approximation were presented. They are applicable either in connection with the target matrix or with its inverse. The set of equations for the unknown ARMA coefficients was derived through the minimization of a frequency domain error subject to a certain set of constraints. Further, it was shown that the two procedures are special cases of a general minimization procedure.

3. It was shown that the spectral matrices of the various ARMA and AR systems depend on $\hat{B}_o$ only through the product $\hat{B}_o\hat{B}_o^{\dagger}$. Thus for computational convenience, $\hat{B}_o$ can be taken as the lower triangular matrix satisfying equation (18).

4. The arbitrariness of the standard choice of the nonhomogeneous condition, $A_o = I_n$, was pointed out. An alternative procedure based on a constrained B_o was developed. The set of equations for the remaining coefficients was obtained through the minimization of the same frequency error. It was proved that the minimum of this quantity in the case of the new procedure is lower or equal to the corresponding value for the standard procedure. A quite efficient algorithm to compute the corresponding sets of ARMA coefficients from one another was given.

5. Attention was called to the computational advantages of the POM procedure over the ACM procedure in terms of the size of the system of linear equations which must be solved to determine the requisite ARMA coefficients.

Acknowledgments

The financial support of this work by a PYI-84 grant for the National Science Foundation is gratefully acknowledged. Industrial Matching funds to the PYI program provided by Brown and Root, Inc.; Conoco, Inc.; Dow Chemical Co.; N. L. Industries; and Tracor, Inc., are appreciated.

References

Gersch, W., and Liu, R. S-Z., 1976, "Time Series Methods for the Synthesis of Random Vibration Systems," ASME JOURNAL OF APPLIED MECHANICS, pp. 159–165.

Hannan, E. J., 1970, *Multiple Time Series*, Wiley, New York.

Kay, S. M., and Marple, S. L., Jr., 1981, "Spectrum Analysis—A Modern Perspective," *Proceedings of the IEEE*, Vol. 69, No. 11, pp. 1380–1419.

Mignolet, M. P., 1987, "ARMA Simulation of Multivariate and Multidimensional Random Processes," Ph.D. Dissertation, Rice University, Houston, Texas.

Mignolet, M. P., and Spanos, P. D., 1987, "ARMA Monte Carlo Simulation in Probabilistic Structural Analysis," AIAA Dynamics Specialist Conference, Monterey, CA, Part 2B, pp. 800–808.

Mullis, C. T., and Roberts, R. A., 1976, "The Use of Second-Order Information in the Approximation of Discrete-Time Linear Systems," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. ASSP-24, No. 3, pp. 226–238.

Samaras, E., Shinozuka, M., and Tsurui, A., 1985, "ARMA Representation of Random Processes," *ASCE Journal of Engineering Mechanics*, Vol. 111, No. 3, pp. 449–461.

Samii, K., and Vandiver, J. K., 1984, "A Numerically Efficient Technique for the Simulation of Random Wave Forces on Offshore Structures," *Proceedings of the Sixteenth Annual Offshore Technology Conference in Houston, TX*, OTC 4811, pp. 301–305.

Shinozuka, M., 1970, "Simulation of Multivariate and Multidimensional Random Processes," *The Journal of the Acoustical Society of America*, Vol. 49, No. 1 (Part 2), pp. 357–367.

Shinozuka, M., 1972, "Monte Carol Solution of Structural Dynamics," *Computers and Structures*, Vol. 2, pp. 855–874.

Spanos, P-T. D., and Hansen, J. E., 1981, "Linear Prediction Theory for Digital Simulation of Sea Waves," *Journal of Energy Resources Technology*, Vol. 103, pp. 243–249.

Spanos, P-T. D., and Mignolet, M. P., 1986, "Z-Transform Modeling of the P-M Wave Spectrum," *ASCE Journal of Engineering Mechanics*, Vol. 112, No. 8, pp. 745–759.

Spanos, P-T. D., and Schultz, K. P., 1985, "Two-Stage Order-of-Magnitude Matching for the von Karman Turbulence Spectrum," *Proceedings of the Fourth International Conference on Structural Safety and Reliability*, ICOSSAR, Vol. 1, pp. 211–218.

Spanos, P-T. D., and Schultz, K. P., 1986, "Numerical Synthesis of Trivariate Velocity Realizations of Turbulence," *International Journal of Nonlinear Mechanics*, Vol. 21, No. 4, pp. 269–277.

Wang, Z., and Fang, T., 1986, "A Time-Domain Method for Identifying Modal Parameters," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 28–32.

P. D. Spanos
Professor,
Mem. ASME

M. P. Mignolet[1]
Research Assistant,
Student Mem. ASME

Brown School of Engineering,
Rice University,
Houston, TX 77251

Recursive Simulation of Stationary Multivariate Random Processes—Part II

Stability and invertibility aspects of the AR to ARMA procedures developed in Part I in connection with simulation of multivariate random processes are addressed. A general criterion is proved for this purpose. Furthermore, several properties regarding the matching of the correlations at various time lags of the target and the simulated processes are shown. Finally, the reliability and efficiency of the discussed procedures are demonstrated by application to spectra encountered in earthquake engineering, offshore engineering, and wind engineering.

Introduction

In the first part of this series the usefulness of efficient algorithms for simulation of realizations of multivariate stochastic processes with specified (target) spectral matrices is discussed in context with random vibration analyses of multi-degree-of-freedom systems. Various procedures for determining appropriate autoregressive moving average (ARMA) algorithms are presented. All of the presented ARMA procedures are based on a prior AR approximation of the target spectral matrix. Equations for the unknown matrix coefficients of the various algorithms are obtained through the minimization of frequency domain errors. These equations can be readily used for a particular practical problem. In this regard a critical assessment of the features of these procedures is provided as well.

In this part of the series, the various procedures are examined from the perspective of stability and invertibility of the generated ARMA system, and the auto and cross-correlations matching properties of the target and the simulated processes. Further, some implementation aspects are discussed. Finally, the reliability and the efficiency of these procedures are demonstrated by producing ARMA approximations of physical processes encountered in various structural dynamics applications.

In order to avoid needless duplication, all equations marked by an asterisk in this part will refer to the equations with the same number given in Part I.

Stability and Invertibility

Preliminary Remarks. Ensuring stability of the discrete system which is associated with a particular simulation algorithm is of fundamental importance. Indeed, it is a necessary condition for the stationarity of the system response, that is, of the time series generated. In the case of an ARMA system, stability is ensured if and only if the poles of the transfer function, defined as the roots z_i of

$$\det\left[\sum_{k=0}^{p} A_k z^{-k}\right] = 0 \tag{1}$$

have a modulus smaller than one. Clearly, the computation of z_i from equation (1) can represent a burden, especially for large values of the product $p \times n$. A more practical criterion is, therefore, desirable. In the following a sufficient condition for the stability of the ARMA system obtained by the generalized original matrix method will be established. It is based on an extension of a theorem on the Lyapunov equation (Mullis and Roberts, 1976). The condition is applicable to the ARMA systems developed by the ACM and POM procedures, as well as to the AR systems.

Note that the ARMA algorithms obtained by using the original spectral matrix can be interpreted as duals of the algorithms obtained from the inverse spectral matrix. Thus, the ensuing mathematical developments can be applied, as well, to examine the invertibility of the latter algorithms.

Two Auxiliary Block Matrices. The matrices K and K' with (u, i) $n \times n$ block elements defined by the equations

$$K_{ui} = R_{\hat{Y}\hat{Y}}(u-i) - \frac{1}{2\omega_b}\sum_{\ell=0}^{q-1} R_{\hat{Y}W}(u-\ell)R_{\hat{Y}W}^{\dagger}(i-\ell)$$

$$u,i = 0, \ldots, p-1 \tag{2}$$

$$K'_{ui} = R_{\hat{Y}\hat{Y}}(u-i) - \frac{1}{2\omega_b}\sum_{\ell=0}^{q} R_{\hat{Y}W}(u-\ell)R_{\hat{Y}W}^{\dagger}(i-\ell)$$

$$u,i = 0, \ldots, p \tag{3}$$

are crucial in the forthcoming proof. Thus, their properties are first presented.

[1]Currently Assistant Professor, Department of Mechanical and Aerospace Engineering, Arizona State University, Tempe, AZ 85287.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 1, 1986; final revision March 27, 1987.

Paper No. 87-WA/APM-13

It is readily seen from their definition that the matrices K and K' are symmetric. That is,

$$K = K^{\dagger} \tag{4}$$

$$K' = K'^{\dagger}. \tag{5}$$

Next, note that K is the covariance matrix of p random vectors $\mathbf{Z}_i$, each of n components. That is

$$K_{ui} = E[\mathbf{Z}_u \mathbf{Z}_i^{\dagger}] \; u,i = 0, \ldots, p-1, \tag{6}$$

where

$$\mathbf{Z}_u = \hat{\mathbf{Y}}_{r-u} - \frac{1}{2\omega_b} \sum_{\ell=0}^{q-u-1} R_{\hat{\mathbf{Y}}\mathbf{W}}(-\ell)\mathbf{W}_{r-\ell-u} \tag{7}$$

for $u = 0, \ldots, p-1$ and any given value of r.

Thus, K is at least positive semidefinite. The pathological case, det $K = 0$, will not be considered here. With this restriction, the matrix K is positive definite.

The previous reasoning can be repeated to show that K' is also positive semidefinite. From the definition of the matrices K and K', equations (2) and (3), it is clear that

$$K_{ui} = K'_{ui} + \frac{1}{2\omega_b} R_{\hat{\mathbf{Y}}\mathbf{W}}(u-q) R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(i-q) \; u,i = 0, \ldots, p-1 \tag{8}$$

and

$$K_{ui} = K'_{(u+1)(i+1)} \; u,i = 0, \ldots, p-1. \tag{9}$$

Equations (5) and (9) imply that

$$K' = \begin{bmatrix} K_1 & K_2^{\dagger} \\ K_2 & K \end{bmatrix} \tag{10}$$

where K_1 and K_2 are $n \times n$ and $np \times n$ matrices, respectively.

The Companion Matrix. The companion matrix A of an ARMA (p, q) system with leading coefficient $A_o = I_n$ is the square block matrix defined as follows

$$A = \begin{bmatrix} -A_1 & -A_2 & \ldots & -A_{p-1} & -A_p \\ I_n & 0 & & 0 & 0 \\ 0 & I_n & & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & & \cdot & \cdot \\ 0 & 0 & \ldots & I_n & 0 \end{bmatrix} \tag{11}$$

Using the Frobenius-Schur formula for determinants, it is readily shown that the eigenvalues λ_i $(i = 1, \ldots, n \times p)$ of the matrix A are solutions of the equation

$$\det\left[\sum_{k=0}^{p} A_k \lambda^{-k}\right] = 0. \tag{12}$$

Upon comparing equations (1) and (12), it becomes obvious that the poles of the ARMA transfer function are the eigenvalues of A.

Next the matrices A, K, K' are combined to yield additional results. First, eliminating the coefficients B_ℓ from equations (62*) and (63*) yields

$$\sum_{u=0}^{p} A_u \left[R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(u-i) - \frac{1}{2\omega_b} \sum_{\ell=0}^{q} R_{\hat{\mathbf{Y}}\mathbf{W}}(u-\ell) R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(i-\ell)\right] + \sum_{\ell=q+1}^{q+r} \Lambda_\ell^{\dagger} R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(i-\ell) = 0$$

$$\text{for } i = 1, \ldots, p \tag{13}$$

where according to equation (26a*), the crosscorrelations $R_{\hat{\mathbf{Y}}\mathbf{W}}$ are zero for positive lags. Further, combining equations (3), (13), and (65*) gives

$$\sum_{u=0}^{p} A_u K'_{ui} + \sum_{\ell=q+1}^{q+r} \Lambda_\ell^{\dagger} R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(i-\ell) = 2\omega_b \tilde{E}_{\min} \delta_{i0} \; i = 0, \ldots, p. \tag{14}$$

Then, it is readily shown that

$$[KA^{\dagger}]_{ui} = K_{u(i-1)} \; i = 1, \ldots, p-1; \; u = 0, \ldots, p-1 \tag{15}$$

and

$$[KA^{\dagger}]_{u0} = K'_{(u+1)0} + \sum_{\ell=q+1}^{q+r} R_{\hat{\mathbf{Y}}\mathbf{W}}(u+1-\ell)\Lambda_\ell \; u = 0, \ldots, p-1. \tag{16}$$

Furthermore, it can be proved that

$$[AKA^{\dagger}]_{ui} = K'_{ui} + L_{ui} - 2\omega_b \tilde{E}_{\min} \delta_{u0} \delta_{i0} \; u,i = 0, \ldots, p-1 \tag{17}$$

where

$$L_{ui} = \left[\sum_{\ell=q+1}^{q+r} \Lambda_\ell^{\dagger} R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(i-\ell)\right]\delta_{u0} + \left[\sum_{\ell=q+1}^{q+r} R_{\hat{\mathbf{Y}}\mathbf{W}}(u-\ell)\Lambda_\ell\right]\delta_{i0}. \tag{18}$$

Finally, equation (17) can be rewritten as

$$[AKA^{\dagger}]_{ui} = K_{ui} + L_{ui} - 2\omega_b \tilde{E}_{\min} \delta_{u0} \delta_{i0} - \frac{1}{2\omega_b} R_{\hat{\mathbf{Y}}\mathbf{W}}(u-q) R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(i-q) \tag{19}$$

by relying on equation (8).

Stability Criterion. The derivation of a sufficient stability criterion will be based on the following theorem.

If the block matrix

$$G = [F, AF, \ldots, A^{p-1}F] \tag{20}$$

is nonsingular and there exist symmetric matrices K and C, with K positive definite and C positive semidefinite (possibly zero) such that

$$K = A \; K \; A^{\dagger} + \gamma F \; F^{\dagger} + C \tag{21}$$

for some $\gamma > 0$, then the eigenvalues of A all lie within the unit disk of the complex plane.

The proof of this theorem for A and F being $p \times p$ and $p \times 1$ matrices, respectively, appears in Mullis and Roberts (1976). Its extension to A and F being $np \times np$ and $np \times n$ matrices, respectively, is straightforward and will not be presented here. It should, however, be noted that the matrix A_{MR} appearing in that article is related to the one defined by equation (11) through the equation

$$A_{MR} = Q \; A \; Q^{\dagger} \tag{22}$$

where

$$Q = \begin{bmatrix} 0 & \ldots & I_n \\ 0 & \ldots & 0 \\ I_n & \ldots & 0 \end{bmatrix}. \tag{23}$$

Clearly, such a transformation does not alter the values of the eigenvalues of the matrix A.

Introduce first the square root $\tilde{E}_{1/2}$ of the error matrix such that

$$\tilde{E}_{1/2} \tilde{E}_{1/2}^{\dagger} = \tilde{E}_{\min}. \tag{24}$$

Define the following block matrices

$$F_1 = [R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(-q), \ldots, R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(p-q-1)]^{\dagger} \tag{25}$$

$$F_2 = [\tilde{E}_{1/2}^{\dagger}, 0, \ldots, 0]^{\dagger} \tag{26}$$

$$G_1 = [F_1, A \; F_1, A^2 F_1, \ldots, A^{p-1} F_1] \tag{27}$$

$$G_2 = [F_2, A \; F_2, A^2 F_2, \ldots, A^{p-1} F_2]. \tag{28}$$

Also introduce the matrices

$$C_1 = 2\omega_b F_2 F_2^\dagger + \alpha\, F_1 F_1^\dagger - L \tag{29}$$

$$C_2 = \alpha\, F_2 F_2^\dagger + \frac{1}{2\omega_b} F_1 F_1^\dagger - L \tag{30}$$

where L is the matrix with blocks defined by equation (18), and the scalars α, γ_1, γ_2 satisfy the equations

$$\gamma_1 = \frac{1}{2\omega_b} - \alpha \tag{31}$$

$$\gamma_2 = 2\omega_b - \alpha. \tag{32}$$

Clearly, equation (19) can be written in either of the forms

$$K = A\, K\, A^\dagger + \gamma_1 F_1 F_1^\dagger + C_1 \tag{33}$$

or

$$K = A\, K\, A^\dagger + \gamma_2 F_2 F_2^\dagger + C_2. \tag{34}$$

Upon applying the aforementioned theorem to equations (33) and (34), and noting the equality of the poles of the transfer function to the eigenvalues of A, one obtains the following sufficient criterion for the stability of the ARMA system.

If, for either triplet (G_1, C_1, γ_1) or (G_2, C_2, γ_2), $\det G \neq 0$ and there exists a value of α such that C and γ are a positive semidefinite $np \times np$ matrix and a positive scalar, respectively, then the ARMA system whose coefficients are computed from equations (59*), (60*), (62*), and(63*) is stable.

It should be noted that the nonsingularity of G_2 is equivalent to

$$\det(\tilde{E}_{\min}) \neq 0 \tag{35}$$

which is computationally easier to check than $\det(G_1) \neq 0$.

Indeed, only the first $2n$ rows of $\alpha\, F_2 F_2^\dagger - L$ may be linearly independent. The following ones are linear combinations of the rows $n+1$ through $2n$. The inclusion of $1/2\omega_b\, F_1 F_1^\dagger$ can increase the rank of C by at most n.

Clearly, the matrices C_1 and C_2 can be written in the form

$$C = \alpha_1 F_1 F_1^\dagger + \alpha_2 E_1 E_1^\dagger M - U E_1^\dagger - E_1 U^\dagger \tag{39}$$

where M, E_1, and U are the following block matrices and block vectors

$$M = \begin{bmatrix} \tilde{E}_{\min} & & & 0 \\ & \tilde{E}_{\min} & & \\ 0 & & \ddots & \\ & & & \tilde{E}_{\min} \end{bmatrix} \tag{40}$$

$$E_1 = [I_n, 0, \ldots, 0]^\dagger \tag{41}$$

and

$$U = \left[\left(\sum_{\ell=q+1}^{q+r} \Lambda_\ell^\dagger R_{\hat{\mathbf{Y}}\mathbf{W}}^\dagger(-\ell) \right), \ldots, \left(\sum_{\ell=q+1}^{q+r} \Lambda_\ell^\dagger R_{\hat{\mathbf{Y}}\mathbf{W}}^\dagger(p-\ell-1) \right) \right]^\dagger. \tag{42}$$

and $\alpha_1 \alpha_2$ are appropriate scalars assuming different values for C_1 and C_2. The eigenvectors of C can be written as linear combinations of the column vectors of F_1, E_1, and U, the coefficients of which are the blocks of eigenvectors of the $3n \times 3n$ matrix $\tilde{C}$

$$\tilde{C} = \begin{bmatrix} \alpha_1 F_1^\dagger F_1 & \alpha_1 F_1^\dagger E_1 & \alpha_1 F_1^\dagger U \\ \alpha_2 E_1^\dagger M F_1 - U^\dagger F_1 & \alpha_2 E_1^\dagger M E_1 - U^\dagger E_1 & \alpha_2 E_1^\dagger M U - U^\dagger U. \\ -E_1^\dagger F_1 & -E_1^\dagger E_1 & -E_1^\dagger U \end{bmatrix} \tag{43}$$

AR System. The previous mathematical development can be repeated to treat the stability of the AR system as a special case. The criterion still holds provided that

$$R_{\hat{\mathbf{Y}}\mathbf{W}}(k) = 0 \quad \text{for all } k \tag{36}$$

and

$$\tilde{E}_{\min} = \hat{B}_o \hat{B}_o^\dagger. \tag{37}$$

The corresponding K matrix is positive semidefinite. Assuming again that $\det K \neq 0$ and taking (G, C, γ) to be (G_2, C_2, γ_2), the criterion simply reduces to the nonsingularity of $\hat{B}_o$.

ACM Procedure. In the special case of the ACM procedure, the previous criterion can be further simplified. Indeed, if no additional constraint is enforced, $r = 0$ and $L = 0$. The positive semidefiniteness of $\tilde{E}_{\min}$ implies that the choice $0 < \alpha < \min(2\omega_b, 1/2\omega_b)$ yields C_1, C_2 positive semidefinite and $\gamma_1, \gamma_2 > 0$. Thus, the criterion reduces to $\det(G_1) \neq 0$ or $\det(\tilde{E}_{\min}) \neq 0$. Note from equation (46*), that the denominators of the ARMA transfer functions obtained by the procedures based on constraining A_o or B_o are proportional. Thus, the corresponding ARMA systems have identical stability characteristics.

POM Procedure. Clearly, the stability of the ARMA system derived by the POM procedure can be analyzed by relying on the general criterion and selecting $r = p$.

Practical Verification of the Stability Condition. The computational complexity of the verification of the positive semidefiniteness of C can be reduced by noting that the rank ρ_c of C is at most $3n$. Specifically,

$$\rho_c \leq [\max(3,p)] \times n. \tag{38}$$

Moreover, the eigenvalues of $\tilde{C}$ are the $3n$ nonzero eigenvalues of C. Thus, a condition sufficient to ensure stability is the positive semidefiniteness of $\tilde{C}$.

A Consequence of Stability. Note that if the developed ARMA system is stable, the equation

$$K = A\, K\, A^\dagger + Q \tag{44}$$

has a unique solution K for every $np \times np$ matrix Q. This can be proved by relying on a general stability theorem (Kailath, 1980). In particular, there is one and only one matrix K solution of equations (33) or (34) for every matrix L and vectors F_1 and F_2.

Next, the matrices $K_2^\dagger$ and K_1 are computed by combining equations (5), (9), and (14) in the form

$$K'_{0i} = -\sum_{u=0}^{p-1} A_{u+1} K_{u(i-1)} - \sum_{\ell=q+1}^{q+r} \Lambda_\ell^\dagger R_{\hat{\mathbf{Y}}\mathbf{W}}^\dagger(i-\ell) \quad i = 1, \ldots, p \tag{45}$$

$$K'_{00} = 2\omega_b \tilde{E}_{\min} - \sum_{\ell=q+1}^{q+r} \Lambda_\ell^\dagger R_{\hat{\mathbf{Y}}\mathbf{W}}^\dagger(-\ell) - \sum_{u=1}^{p} A_u K'^\dagger_{0u}. \tag{46}$$

Finally, K' is obtained from equation (10). Thus, for any given matrices $\tilde{E}_{\min}$, Λ_ℓ and $R_{\hat{\mathbf{Y}}\mathbf{W}}(i-\ell)$ there exists a unique solution K' to equation (14).

This conclusion applies also in the case of the AR system and implies that equations (12*) and (14*)–(16*) admit only the trivial solution given by equation (17*).

Invertibility. As it was alluded to at the beginning of this section, the invertibility of the ARMA systems derived through the inverse spectral matrix can be deduced from the

previous developments by permuting the roles of A_k,p and B_ℓ,q, respectively.

Time Domain Analysis

Crosscorrelation Matching—Original Spectral Matrix. The two generalized procedures can also be investigated in the time domain by analyzing the connection between the auto and cross-correlations of the AR and ARMA processes.

Combining equations (60*), (63*), and (30*) yields

$$\sum_{k=0}^{\min(i,p)} A_k[R_{\tilde{\mathbf{Y}}\mathbf{W}}(k-i)-R_{\hat{\mathbf{Y}}\mathbf{W}}(k-i)]=0 \quad i=0,\ldots,q+r \tag{47}$$

which represent a set of $(q + r)$ matrix equations in the $(q + r)$ unknowns

$$\Delta R_{\mathbf{Y}\mathbf{W}}(k)=R_{\hat{\mathbf{Y}}\mathbf{W}}(k)-R_{\tilde{\mathbf{Y}}\mathbf{W}}(k) \quad k=-q-r,\ldots,0. \tag{48}$$

This system has only the trivial solution $\Delta R_{\mathbf{YW}}(k) = 0$ provided that the matrix A_o is nonsingular. This condition is always satisfied. Thus,

$$R_{\tilde{\mathbf{Y}}\mathbf{W}}(k)=R_{\hat{\mathbf{Y}}\mathbf{W}}(k) \quad k=-q-r,\ldots,0. \tag{49}$$

It is seen that the AR to ARMA procedures developed by relying on the original target spectral matrix yield processes which have the same excitation-response crosscorrelation as the AR process which approximates the target process. The proof of this property further elucidates the features of these procedures and offers a formulation which is an alternative to the minimization of the frequency domain error defined by equation (32*).

Autocorrelation Matching—Original Spectral Matrix. Combining equations (29*), (62*), (65*), and (49) yields the following relations between the autocorrelations of the two processes

$$\sum_{k=0}^{p} A_k[R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k-i)-R_{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}}(k-i)]=2\omega_b\tilde{E}_{\min}\delta_{i0} - \sum_{\ell=q+1}^{q+r} \Lambda_\ell^\dagger R_{\hat{\mathbf{Y}}\mathbf{W}}^{\dagger}(i-\ell) \tag{50}$$

$$i=0,\ldots,p.$$

It was shown previously that if the ARMA system is stable, this set of equations yields a unique value to the difference of the autocorrelations

$$\Delta R_{\mathbf{YY}}(k)=R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(k)-R_{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}}(k) \quad k=0,\ldots,p. \tag{51}$$

In the special case of the ACM procedure, $r = 0$, and equation (50) reduces to

$$\sum_{k=0}^{p} A_k\Delta R_{\mathbf{YY}}(k-i)=2\omega_b\tilde{E}_{1/2}\tilde{E}_{1/2}^{\dagger}\delta_{i0} \quad i=0,\ldots,p. \tag{52}$$

Then, $\Delta R_{\mathbf{YY}}$ can be considered as the autocorrelation function of the output of the AR system whose transfer function matrix is

$$H(z)=\left(\sum_{k=0}^{p} A_k z^{-k}\right)^{-1}\tilde{E}_{1/2}. \tag{53}$$

Thus, a possible solution of equation (52) is

$$\Delta R_{\mathbf{YY}}(\ell)=R_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(\ell)-R_{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}}(\ell)=\int_{-\omega_b}^{\omega_b}\left(\sum_{k=0}^{p} A_k e^{jk\omega T}\right)^{-1}\cdot \tilde{E}_{\min}\left(\sum_{k=0}^{p} A_k e^{-jk\omega T}\right)^{-\dagger} e^{j\ell\omega T}d\omega. \tag{54}$$

And, as before, the requirement of stability of the ARMA system renders this solution unique.

Thus, it is seen that if $\tilde{E}_{\min}$ approaches zero as p increases, the AR and the ARMA representations of the target process have identical autocorrelation values at time lags $\ell = 0, \ldots, p$. This property illuminates further the features of the AR to ARMA procedure and can be used as an alternative formulation to the one introduced by the frequency domain minimization.

Crosscorrelation Matching—Inverse Spectral Matrix. Clearly, the previous analysis applies also to the generalized inverse procedure provided that the symbols A_k,p and B_ℓ,q are interchanged. The quantities involved will be the auto and cross-correlations of the output $\hat{X}$ and $\tilde{X}$ of the inverse AR and ARMA systems, respectively. Some additional conclusions concerning the original systems can be drawn.

The output of a stable causal ARMA system $\tilde{\mathbf{Y}}$ can be computed from equation (25*) or equivalently from equation

$$\tilde{\mathbf{Y}}_r=\sum_{i=0}^{\infty} h(i)\mathbf{W}_{r-i}, \tag{55}$$

where the impulse response $h(i)$ is such that

$$H(z)=D^{-1}(z)N(z)=\sum_{i=0}^{\infty} h(i)z^{-i} \quad \text{for } |z|\geq 1. \tag{56}$$

Thus,

$$R_{\tilde{\mathbf{Y}}\mathbf{W}}(-s)=E[\tilde{\mathbf{Y}}_r\mathbf{W}_{r-s}^{\dagger}]=\sum_{i=0}^{\infty} h(i)E[\mathbf{W}_{r-i}\mathbf{W}_{r-s}^{\dagger}]=2\omega_b h(s) \tag{57}$$

is the relation between the impulse response and crosscorrelation sequences.

The product of the transfer functions of the ARMA system and its inverse is equal to I_n. Thus, the impulse responses, or equivalently the crosscorrelations $R_{\tilde{\mathbf{Y}}\mathbf{W}}$ and $R_{\tilde{\mathbf{X}}\mathbf{W}}$ of the ARMA system and its inverse, respectively, are related through the following convolution

$$\sum_{s=0}^{i} R_{\tilde{\mathbf{Y}}\mathbf{W}}(s-i)R_{\tilde{\mathbf{X}}\mathbf{W}}(-s)=4\omega_b^2 I_n\delta_{i0} \quad i=0,1,\ldots \tag{58}$$

This equation shows that the first i values of $R_{\tilde{\mathbf{X}}\mathbf{W}}$ are uniquely determined by the first i values of $R_{\tilde{\mathbf{Y}}\mathbf{W}}$ and vice versa.

Thus, the matching of $p + r + 1$ crosscorrelations of the inverse systems imply the same property for the original systems. That is,

$$R_{\hat{\mathbf{Y}}\mathbf{W}}(k)=R_{\hat{\mathbf{Y}}\mathbf{W}}(k) \quad k=-p-r,\ldots,0. \tag{59}$$

This matching property which can be seen as the dual of equation (49) elucidates, as before, the features of the AR to ARMA procedures developed by relying on the AR to ARMA inverse spectral matrix method.

The POM Procedure as Padé-Type Approximation. With the help of equation (59) the POM procedure can be seen as a Padé-type approximation of the AR transfer function. Indeed, it is readily shown that the lowest power of z^{-1} in the expansion of the function

$$N(z)H_{AR}^{-1}(z)-D(z)=\left(\sum_{\ell=0}^{q} B_\ell z^{-\ell}\right)\hat{B}_o^{-1}\left(\sum_{s=0}^{m}\hat{A}_s z^{-s}\right) - \left(\sum_{k=0}^{p} A_k z^{-k}\right) \tag{60}$$

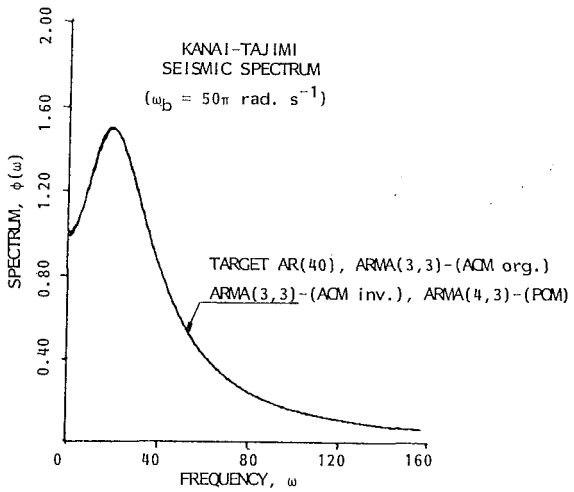


Fig. 1 Kanai-Tajimi seismic spectrum

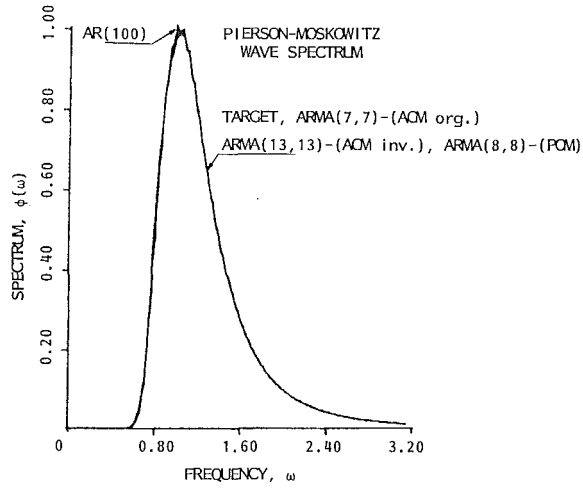


Fig. 2 Pierson-Moskowitz wave spectrum

is $z^{-(p+q+1)}$, when the coefficients A_k and B_ℓ satisfy equations (54*) and (55*) or equivalently, equations (84*) and (85*). Further, the function

$$\Delta H(z) = H_{\mathrm{ARMA}}(z) - H_{\mathrm{AR}}(z)$$

$$= \Big(\sum_{k=0}^{p} A_k z^{-k}\Big)^{-1} [N(z)\hat{B}_o^{-1}\hat{D}(z) - D(z)]\Big(\sum_{s=0}^{m} \hat{A}_s z^{-s}\Big)^{-1} \hat{B}_o \qquad (61)$$

is analytic in the domain Ω: $|z| > |z_1|$ where z_1 is the root of equation

$$\det\Big[\Big(\sum_{k=0}^{p} A_k z^{-k}\Big)\Big(\sum_{s=0}^{m} \hat{A}_s z^{-s}\Big)\Big] = 0 \qquad (62)$$

which has the greatest modulus. Thus, the Laurent expansion of $\Delta H(z)$ in the domain Ω has the following form

$$\Delta H(z) = \sum_{s=p+q+1}^{\infty} \delta(s) z^{-s}. \qquad (63)$$

That is, the first $p + q + 1$ coefficients of the Laurent expansions of $H_{\mathrm{AR}}(z)$ and $H_{\mathrm{ARMA}}(z)$ are equal. Thus the Padé-type approximation is established.

Implementation Aspects

In this section the choice of the initial conditions for the recursions of the AR and ARMA systems will be investigated.

Once the coefficients $\hat{B}_o$, $\hat{A}_k$, B_ℓ, and A_k have been computed from appropriate equations, the AR and ARMA models of the target process are available. These approximations can be used to simulate a time history of $\mathbf{Y}$. That is, values of $\mathbf{Y}_r = \hat{\mathbf{Y}}_r$ or $\tilde{\mathbf{Y}}_r$, $r = 0, \ldots, N$ can be generated by using equations (1*) or (25*), respectively. Clearly, to compute $\mathbf{Y}_o$, the values $\mathbf{Y}_{-1}, \ldots, \mathbf{Y}_{-s}$, where s is m or p for the AR or ARMA model, respectively, are needed. These initial values form a $n \times s$ random vector $\mathbf{Y}_{IC}$ with covariance matrix K_{IC}. Its $n \times n$ block element ui is

$$(K_{IC})_{ui} = E[\mathbf{Y}_{-u}\mathbf{Y}_{-i}^{\dagger}] = R_{\mathbf{YY}}(u-i) \quad u,i = 1, \ldots, s. \qquad (64)$$

The vector $\mathbf{Y}_{IC}$ can be conveniently generated from $n \times s$ independent normal $N(0, 1)$ deviates stacked in the vector $\mathbf{P}$, through the equation

$$\mathbf{Y}_{IC} = V\,\mathbf{P}. \qquad (65)$$

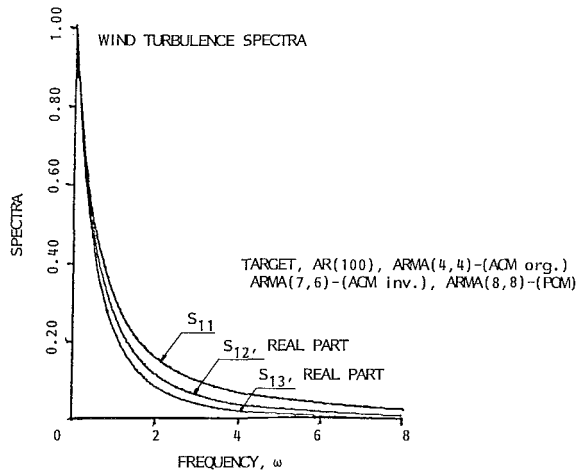


Fig. 3 Wind turbulence spectrum

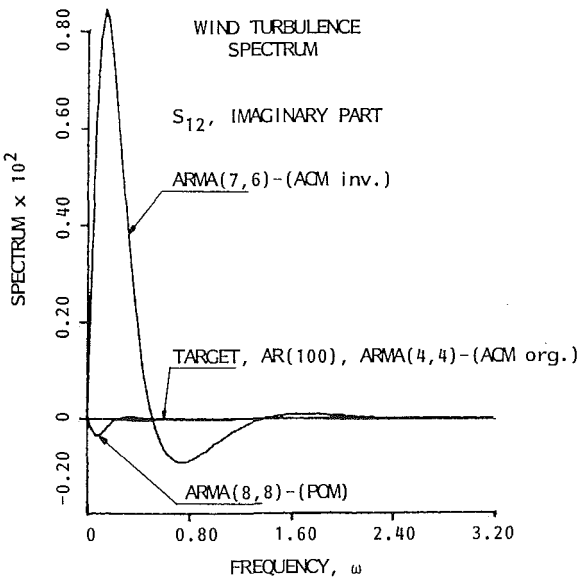


Fig. 4 Wind turbulence spectrum imaginary part of 1,3 component

The symbol V denotes the $ns \times ns$ lower triangular matrix obtained from the Cholesky decomposition of K_{IC}

$$K_{IC} = V\,V^{\dagger}. \qquad (66)$$

Clearly, $p \leq m$ so that the ARMA scheme exhibits a reduction of the computational complexity as compared to the AR algorithm.

Numerical Results

The ARMA system development procedures presented in the first part of this series have been successfully applied to various spectral shapes encountered in engineering.

For the first application, the Kanai-Tajimi spectrum from earthquake engineering defined as (Madsen et al., 1986)

$$\phi_1(\omega) = \frac{\omega_g^4 + 4\zeta_g^2\omega_g^2\omega^2}{(\omega_g^2 - \omega^2)^2 + 4\zeta_g^2\omega_g^2\omega^2} S_o \tag{67}$$

was selected as a target expression. The parameters ω_g, ζ_g, S_o of the spectrum were selected to be 8π, 0.8, and 1, respectively. The spectra of the different approximations are plotted on Fig. 1. Clearly, either of the AR(40), ARMA(3,3)-(ACM org.), ARMA(3,3)-(ACM inv.), and ARMA(4,3)-(POM) systems can be used for an accurate simulation of earthquake records. Note that the qualifiers (org) or (inv) denote ARMA systems produced by relying on the original or the inverse spectral matrices, respectively.

The simulation of a time record of wave elevation is an important problem of ocean engineering. The most commonly used statistical description of the elevation process involves the Pierson-Moskowitz (P-M) spectrum. Its dimensionless form (Spanos, 1983)

$$\phi_2(\omega) = \frac{e^{5/4}}{\omega^5} \exp\left(-\frac{5}{4\omega^4}\right) \tag{68}$$

was selected as a target expression. Figure 2 shows that a reasonable AR approximation of the P-M spectrum can lead to quite efficient ARMA models of the wave elevation. The comparatively large filter orders needed for the ACMinv. method can be attributed to the mathematical peculiarity of the target spectrum as studied by Spanos and Mignolet (1986).

Finally, a trivariate example was chosen from wind engineering. Specifically, the fluctuating velocities in the direction of the mean wind at 3 equidistant points can be modeled as normal random processes with mean 0 and the following dimensionless spectral matrix (see Madsen et al., 1986)

$$S(\omega) = \frac{1}{(1+|\omega|)^{5/3}} \begin{bmatrix} 1 & \exp\left(\frac{-\beta d|\omega|}{\eta}\right) & \exp\left(\frac{-2\beta d|\omega|}{\eta}\right) \\ \exp\left(\frac{-\beta d|\omega|}{\eta}\right) & 1 & \exp\left(\frac{-\beta d|\omega|}{\eta}\right) \\ \exp\left(\frac{-2\beta d|\omega|}{\eta}\right) & \exp\left(\frac{-\beta d|\omega|}{\eta}\right) & 1 \end{bmatrix} . \tag{69}$$

The symbols d and η represent the distance between 2 points and their common height above the ground. The constant β equals 0.163. Equation (69) with $d = \eta$ was chosen as a target spectrum. Figures 3-4 show the matching of the 1,1, 1,2, and 1,3 components of the AR, ARMA, and target spectral matrices. Clearly, the AR(100), ARMA(4,4)-(ACMorg.), ARMA(7,6)-(ACMinv.), and ARMA(8,8)-(POM) processes represent good approximations of the turbulent velocities.

Note that all the systems presented have been found to be stable.

It is noted that this series of articles aims primarily to unify and set on a concrete basis, from a perspective of frequency and time domain optimization, stability, and matching criteria, various AR to ARMA simulation procedures of multivariate stationary random processes. The presented numerical results reinforce their usefulness for random vibration problems. In this regard they should be viewed in context with other applications such as in Gersch and Yonemoto (1976).

Concluding Remarks

Properties and computational aspects of the ARMA system (algorithm) procedures developed in Part I have been studied in detail. The results of the present analysis may be summarized as follows.

1. The stability of the systems obtained by the generalized original spectral matrix procedure can be ensured if at least one of the triplets (G_1, C_1, γ_1) or (G_2, C_2, γ_2) as defined by equations (27)-(32) satisfies the conditions of nonsingular G, positive semi-definite C, and positive γ. When $p > 3$, this criterion can be simplified by noting that the $3n$ nonzero eigenvalues of C are also eigenvalues of the $3n \times 3n$ matrix $\tilde{C}$ defined by equation (43). These conditions are applicable to the special cases of the POM and ACM procedures. In the latter case, it was proved that the stability of the system is ensured by the condition $\det(G_1) \neq 0$ or $\det(\tilde{E}_{\min}) \neq 0$.
2. Relying on the duality of the procedures which involve the original or the inverse spectral matrix, it was shown that the stability criterion could be used to test the invertibility of the systems developed by the generalized inverse spectral matrix procedure.
3. It was proved that the AR and ARMA processes derived through the generalized original or inverse spectral matrix procedure have in common the same $p + q + 1$ first input-output crosscorrelations.
4. It was shown that the differences between the first $p + 1$ autocorrelations of the AR and ACM (original spectral matrix) processes are the autocorrelations of an AR process with known parameters.
5. It was discussed that the simulated ARMA processes would not display nonstationary characteristics if pertinent initial values were generated as jointly normal random variables with mean 0 and covariance matrix K_{IC} as defined by equation (64).
6. Finally, the discussed ARMA procedures were exemplified by application to a variety of spectral shapes encountered in different technical areas such as earthquake engineering (Kanai-Tajimi spectrum), ocean engineering (Pierson-Moskowitz spectrum), and wind engineering (spectral matrix of the turbulent velocities). Excellent matching of the target spectra and the ARMA approximations was observed in all cases.

It should be noted that the preceding developments could also serve as a basis for investigating properties of ARMA schemes used in connection with nonstationary random processes such as in Polhemus and Cakmak (1981) and Gersch and Kitagawa (1985), and multidimensional random processes (Naganuma et al., 1985).

References

Gersch, W., and Kitagawa, G., 1985, "A Time Varying AR Coefficient

Model for Modelling and Simulating Earthquake Ground Motion," *Earthquake Engineering and Structural Dynamics*, Vol. 13, pp. 243–254.

Gersch, W., and Yonemoto, J., 1977, "Synthesis of Multivariate Random Vibration Systems: A Two-Stage Least Squares AR-MA Model Approach," *Journal of Sound and Vibration*, Vol. 52, No. 4, pp. 553–565.

Hannan, E. J., 1970, *Multiple Time Series*, Wiley, New York.

Kailath, T., 1980, *Linear Systems*, Prentice-Hall, Englewood Cliffs, N.J.

Madsen, H. O., Krenk, S., and Lind, N. C., 1986, *Methods of Structural Safety*, Prentice-Hall, Englewood Cliffs, N.J.

Mignolet, M. P., 1987, "ARMA Simulation of Multivariate and Multidimensional Random Processes," Ph.D. Dissertation, Rice University, Houston, Texas.

Mullis, C. T., and Roberts, R. A., 1976, "The Use of Second-Order Information in the Approximation of Discrete-Time Linear Systems," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. ASSP-24, No. 3, pp. 226–238.

Naganuma, T., Deodatis, G., Shinozuka, M,. and Samaras, E., 1985, "Digital Generation of Multidimensional Random Fields," *4th International Conference on Structural Safety and Reliability, ICOSSAR '85*, Vol. I, pp. 251–260.

Polhemus, N. W., and Cakmak, A. S., 1981, "Simulation of Earthquake Ground Motion Using Autoregressive Moving Average (ARMA) Models," *Earthquake Engineering and Structural Dynamics*, Vol. 9, pp. 343–354.

Spanos, P. D., 1983, "ARMA Algorithms for Ocean Wave Modeling," ASME *Journal of Energy Resources Technology*, Vol. 105, pp. 300–309.

Spanos, P. D., and Mignolet, M. P., 1986, "z-Transform Modeling of the P-M Wave Spectrum," *ASCE Journal of Engineering Mechanics*, Vol. 112, No. 8, pp. 745–759.

W. D. Iwan
Professor,
Division of Applied Mechanics
and Civil Engineering.
Fellow ASME

K. S. Smith
Member of the Technical Staff,
Jet Propulsion Laboratory,
Lecturer in Applied Mechanics.

California Institute of Technology,
Pasadena, CA 91125

On the Nonstationary Response of Stochastically Excited Secondary Systems

The envelope response of a secondary system is derived for the case where the primary system is subjected to nonstationary stochastic excitation. An approximate closed form expression for the mean square envelope response is obtained for the case of transient response to stationary excitation when the primary and secondary systems are noninteracting. When the combined system is classically damped, the effect of the interaction is described by the introduction of an equivalent noninteracting system. The analytical results are compared with results of numerical simulations.

Introduction

The design of secondary systems to withstand seismic and other loads has recently received increased attention. A secondary system may be a piece of equipment in a primary structure or a substructure which is separate from its supporting primary structure. Secondary systems are usually characterized by having a mass which is small in comparison with the mass of the structure on which they are supported. The excitation to such systems is frequently stochastic in nature as in the case of earthquake, wave, and wind excitation. Secondary systems are often essential to the safety and wellbeing of the primary structure and their failure may have very serious consequences. This is particulary true in critical facilities such as nuclear power plants.

Several researchers have investigated the response of secondary systems to stochastic excitation. Primarily, attention has been focused on determining the stationary response to stationary excitation. It may be argued that this information will adequately represent the response to a long duration excitation with nearly stationary statistics. Using the assumption of stationarity, Singh (1975) undertook to simplify the computation of the mean square response of a single-degree-of-freedom secondary system attached to a multiple-degree-of-freedom primary system. More recently, Igusa and Der Kiureghian (1983, 1985) used perturbation methods to simplify expressions for the stationary response statistics of a multiple-degree-of-freedom secondary system attached arbitrarily to a multiple-degree-of-freedom primary system.

Relatively little work has been done on the response of secondary systems to transient excitation. Perhaps most notable among the work reported is that of Chakravorty and Vanmarcke (1973) who obtained the mean square relative displacement of a single-degree-of-freedom secondary system attached to a single-degree-of-freedom primary system, in response to suddenly applied white noise. The mean square response is a useful measure of response, but for many applications it is more important to have an estimate of the peak or envelope response.

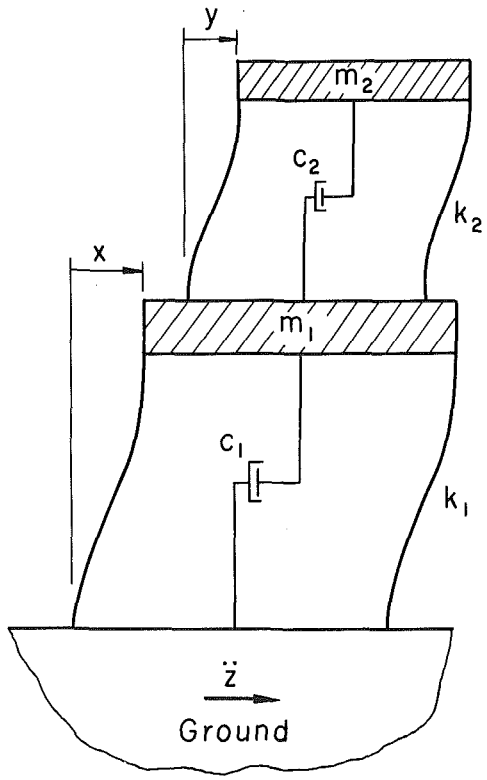


Fig. 1 **Single-degree-of-freedom secondary system attached to single-degree-of-freedom primary system**

This paper presents an approach to the secondary system problem which leads to a general expression for the time-

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 18, 1986; final revision January 24, 1987.

varying statistics of the envelope of the response of a secondary system to excitation represented as a general stochastic process. The primary system and secondary system are each represented as single-degree-of-freedom systems. However, the results derived here may be combined to approximate the behavior of a more complicated primary system (Penzien and Chopra, 1965).

Primary System Response

There are many approaches to analyzing the response of a single-degree-of-freedom linear oscillator to stochastic excitation. The approach chosen herein is analogous to that used subsequently for the secondary system. Only a few results which will be useful later will be presented by way of review.

Let the system be defined as in Fig. 1. Then, neglecting interaction, the relative displacement $x(t)$ of the primary system is described by the equation

$$\ddot{x}+2\zeta_1\omega_1\dot{x}+\omega_1^2 x=-\ddot{z}(t) \tag{1}$$

The response $x(t)$ to any excitation $z(t)$ may, for zero initial conditions, be expressed as

$$x(t)=\int_0^t \frac{1}{\omega_{d1}} e^{-\zeta_1\omega_1(t-\tau)} \sin \omega_{d1}(t-\tau)[-\ddot{z}(\tau)]d\tau \tag{2}$$

where $\omega_{d1} = \omega_1\sqrt{1-\zeta_1^2}$. This integral may be decomposed as

$$x(t)=x_1(t)\cos\omega_{d1}t+x_2(t)\sin\omega_{d1}t \tag{3}$$

where

$$\begin{aligned} x_1(t) &= -\int_0^t \frac{1}{\omega_{d1}} e^{-\zeta_1\omega_1(t-\tau)} \sin\omega_{d1}\tau[-\ddot{z}(\tau)]d\tau \\ x_2(t) &= \int_0^t \frac{1}{\omega_{d1}} e^{-\zeta_1\omega_1(t-\tau)} \cos\omega_{d1}\tau[-\ddot{z}(\tau)]d\tau \end{aligned} \tag{4}$$

For class of problems of interest, it will be assumed that

$$x(t)=a_1(t)\cos[\omega_{d1}t+\phi_1(t)] \tag{5}$$

where $a_1(t)$ and $\phi_1(t)$ are nearly constant over any one period $2\pi/\omega_{d1}$. Equation (5) may be decomposed in the form of equation (3) to give

$$\begin{aligned} x_1(t) &= a_1(t)\cos\phi_1(t) \\ x_2(t) &= -a_1(t)\sin\phi_1(t) \end{aligned} \tag{6}$$

The envelope of the response $a_1(t)$ may be expressed as

$$\begin{aligned} a_1^2(t) &= x_1^2(t)+x_2^2(t) \\ &= \int_0^t\int_0^t \frac{1}{\omega_{d1}^2} e^{-\zeta_1\omega_1(2t-\tau_1-\tau_2)} \\ &\quad \cos\omega_{d1}(\tau_1-\tau_2)[\ddot{z}(\tau_1)][\ddot{z}(\tau_2)]d\tau_1 d\tau_2 \end{aligned} \tag{7}$$

If $z(t)$ is a stochastic process, $a_1(t)$ will also be a stochastic process. Taking the ensemble average of $a_1^2(t)$ gives

$$\begin{aligned} E[a_1^2(t)] &= \frac{1}{\omega_{d1}^2}\int_0^t\int_0^t e^{-\zeta_1\omega_1(2t-\tau_1-\tau_2)} \\ &\quad \cos\omega_{d1}(\tau_1-\tau_2)E[\ddot{z}(\tau_1)\ddot{z}(\tau_2)]d\tau_1 d\tau_2 \end{aligned} \tag{8}$$

For a broad-band stochastic excitation, it may be shown (Smith, 1985) that

$$E[\ddot{z}(\tau_1)\ddot{z}(\tau_2)]\approx\int_{-\infty}^{\infty} S(\omega,\tau_1)e^{i\omega(\tau_2-\tau_1)}d\omega \tag{9}$$

where $S(\omega, t)$ is the evolutionary power spectral density (Priestley, 1965) of the process $\ddot{z}(t)$. Substituting equation (9) into equation (8) and using symmetry in τ_1 and τ_2 gives

$$\begin{aligned} E[a_1^2(t)] &\approx \frac{1}{\omega_{d1}^2}\int_0^t\int_{-\infty}^{\infty} e^{-2\zeta_1\omega_1(t-\tau_1)} \\ &\quad S(\omega,\tau_1)\xi(\omega,\tau_1)d\omega d\tau_1 \end{aligned} \tag{10}$$

where

$$\xi(\omega,t)=2\int_0^t e^{-\zeta_1\omega_1\tau}\cos\omega_{d1}\tau\, e^{i\omega\tau}d\tau \tag{11}$$

It may be shown that the function $\xi(\omega, t)$ is sharply peaked at $\omega = \omega_{d1}$, and that

$$\int_{-\infty}^{\infty}\xi(\omega,t)d\omega=2\pi$$

Taking $S(\omega,t)$ as nearly constant over the region of maximum ξ, thus yields

$$E[a_1^2(t)]\approx\frac{2\pi}{\omega_{d1}^2}\int_0^t e^{-2\zeta_1\omega_1(t-\tau_1)}S(\omega_{d1},\tau_1)d\tau_1 \tag{12}$$

This result is identical to that obtained by Spanos and Lutes (1980), except that ω_1 is replaced here by ω_{d1}.

Secondary System Response

The relative displacement $y(t)$ of a noninteracting secondary system is governed by the equation

$$\ddot{y}+2\zeta_2\omega_2\dot{y}+\omega_2^2 y=-\ddot{z}(t)-\ddot{x}(t) \tag{13}$$

Using equation (1) for the primary system, the right-hand side of equation (13) may be expressed more simply as

$$\begin{aligned} -\ddot{z}-\ddot{x} &= 2\zeta_1\omega_1\dot{x}+\omega_1^2 x \\ &\approx \omega_1^2 a_1(t)\cos[\omega_1 t+\phi_1(t)] \end{aligned} \tag{14}$$

where equation (5) has been used together with the assumption that $\zeta_1 << 1$. $y(t)$ may now be expressed in terms of the Duhammel integral representation

$$\begin{aligned} y(t) &= \int_0^t \frac{1}{\omega_2} e^{-\zeta_2\omega_2(t-\tau)}\sin\omega_2(t-\tau)\omega_1^2 a_1(\tau) \\ &\quad \cos[\omega_1\tau+\phi_1(\tau)]d\tau \end{aligned} \tag{15}$$

Note that $\omega_{d2} = \omega_2\sqrt{1-\zeta_2^2}$ has been replaced by ω_2 in equation (15), since it is assumed that $\zeta_2 << 1$.

It will be assumed that $y(t)$ has the appearance of a harmonic oscillation with slowly varying amplitude and phase, and possibly a slowly varying frequency. That is,

$$y(t)=a_2(t)\cos[\omega'(t)t+\phi_2(t)] \tag{16}$$

where $a_2(t)$, $\omega'(t)$, and $\phi_2(t)$ are slowly varying random functions of time.

Equation (15) exhibits two characteristic frequencies, ω_1 and ω_2. However, one of these frequencies may dominate the other. When the natural frequencies of the primary and secondary systems are well separated, the response of the secondary system is generally dominated by the lowest natural frequency, although a small amount of the higher frequency is still present. Equation (15) will first be decomposed in harmonics of ω_1. This procedure may be expected to yield good results for $\omega_1 \leq \omega_2$, since the frequency ω_1 will certainly be present in $y(t)$ in this case.

Let

$$y(t)=y_1(t)\cos\omega_1 t+y_2(t)\sin\omega_1 t \tag{17}$$

where

$$\begin{aligned} y_1(t) &= \int_0^t \frac{1}{\omega_2} e^{-\zeta_2\omega_2(t-\tau)}\sin\omega_2(t-\tau)\omega_1^2 a_1(\tau) \\ &\quad \cos[\omega_1(t-\tau)-\phi_1(\tau)]d\tau \\ y_2(t) &= \int_0^t \frac{1}{\omega_2} e^{-\zeta_2\omega_2(t-\tau)}\sin\omega_2(t-\tau)\omega_1^2 a_1(\tau) \\ &\quad \sin[\omega_1(t-\tau)-\phi_1(\tau)]d\tau \end{aligned} \tag{18}$$

If equation (16) is decomposed in a similar manner, it may be seen that

$$y_1(t) = a_2(t)\cos[\omega'(t)t - \omega_1 t + \phi_2(t)]$$
$$y_2(t) = -a_2(t)\sin[\omega'(t)t - \omega_1 t + \phi_2(t)] \tag{19}$$

Thus,

$$a_2^2(t) = y_1^2(t) + y_2^2(t)$$
$$= \int_0^t \int_0^t e^{-\zeta_2\omega_2(2t-\tau_1-\tau_2)}\sin\omega_2(t-\tau_1)\sin\omega_2(t-\tau_2)$$
$$\omega_1^4\chi(\tau_1,\tau_2)d\tau_1 d\tau_2 \tag{20}$$

where

$$\chi(\tau_1,\tau_2) = a_1(\tau_1)a_1(\tau_2)\cos[\omega_1(\tau_1-\tau_2)+\phi_1(\tau_1)-\phi_1(\tau_2)] \tag{21}$$

The function $\chi(\tau_1, \tau_2)$ is a random function which depends only on the primary response $x(t)$. $\chi(\tau_1, \tau_2)$ can, in turn, be written in terms of the excitation $z(t)$ using the random functions $x_1(t)$ and $x_1(t)$ so that (Smith, 1985)

$$\chi(t_1,t_2) = \int_0^{t_1}\int_0^{t_2}\frac{1}{\omega_1^2}e^{-\zeta_1\omega_1(t_1+t_2-\tau_1-\tau_2)}$$
$$\cos\omega_1(t_1-t_2-\tau_1+\tau_2)[\ddot{z}(t_1)][\ddot{z}(\tau_2)]d\tau_1 d\tau_2 \tag{22}$$

Equations (20) and (22) express $a_2^2(t)$ as a four-fold integral involving the random function $\ddot{z}(\tau_1)\ddot{z}(\tau_2)$.

A similar calculation may be made in the case that $\omega_2 < \omega_1$. In this case, it will be appropriate to decompose equations (15) and (16) in harmonics of ω_2. This leads to the expression

$$a_2^2(t) = \int_0^t\int_0^t\frac{\omega_1^2}{\omega_2^2}e^{-\zeta_2\omega_2(2t-\tau_1-\tau_2)}$$
$$\cos\omega_2(\tau_1-\tau_2)\hat{x}(\tau_1)\hat{x}(\tau_2)d\tau_1 d\tau_2 \tag{23}$$

where

$$\hat{x}(t) = \int_0^t e^{-\zeta_1\omega_1(t-\tau)}\left[\sin\omega_1(t-\tau) + \frac{2\zeta_1}{\omega_1}\cos\omega_1(t-\tau)\right][-\ddot{z}(\tau)]d\tau \tag{24}$$

Closed Form Solution

The equations derived above provide a description of the stochastic process $a_2(t)$ in terms of the excitation process $\ddot{z}(t)$. Let

$$\omega_\lambda = \min(\omega_1,\omega_2)$$
$$\omega_u = \max(\omega_1,\omega_2) \tag{25}$$

Then, taking ensemble averages in equations (20) and (22), making use of the sharply peaked nature of the kernel function as above and assuming that $S(\omega, \tau_3)$ is a slowly varying function of ω, it may be shown that (Smith, 1985)

$$E[a_2^2(t)] = \int_0^t \mu(t-s)S(\omega_\lambda,s)ds \tag{26}$$

where

$$\mu(t) = \frac{4\pi\omega_1^2}{\omega_2^2}e^{-2\zeta_\lambda\omega_\lambda t}\int_0^t\int_0^{\tau_2}e^{(\zeta_\lambda\omega_\lambda-\zeta_u\omega_u)(\tau_1+\tau_2)}$$
$$\cos\omega_\lambda(\tau_1-\tau_2)\sin\omega_u\tau_1\sin\omega_u\tau_2\,d\tau_1 d\tau_2 \tag{27}$$

Let

$$\nu = \zeta_u\omega_u - \zeta_\lambda\omega_\lambda$$
$$\Delta\omega = \sqrt{(\omega_u-\omega_\lambda)^2+\nu^2} \tag{28}$$

Then, when $\Delta\omega \neq 0$, a closed form expression for $\mu(t)$ may be obtained as

$$\mu(t) = \frac{\frac{\pi\omega_1^2}{\omega_2^2}}{(\Delta\omega)^2[(\omega_\lambda+\omega_u)^2+\nu^2]}\Big\{2\omega_u^2 e^{-2\zeta_\lambda\omega_\lambda t}$$
$$+e^{-2\zeta_u\omega_u t}[2\omega_u^2+(\omega_\lambda^2-\omega_u^2+\nu)^2)(1-\cos 2\omega_u t)$$
$$+2\nu\omega_u\sin 2\omega_u t]$$
$$+2\omega_u e^{-(\zeta_\lambda\omega_\lambda+\zeta_u\omega_u)t}[(\omega_\lambda-\omega_u)\cos(\omega_\lambda+\omega_u)t-\nu\sin(\omega_\lambda+\omega_u)t$$
$$-(\omega_\lambda+\omega_u)\cos(\omega_\lambda-\omega_u)t+\nu\sin(\omega_\lambda-\omega_u)t]\Big\} \tag{29}$$

For $\Delta\omega = 0$, the corresponding result is

$$\mu(t) = \pi e^{-2\zeta_1\omega_1 t}\left[\tfrac{1}{2}t^2 - \frac{1}{2\omega_1}t\sin 2\omega_1 t + \frac{1}{4\omega_1^2}(1-\cos 2\omega_1 t)\right] \tag{30}$$

It may be shown that $\mu(t) \geq 0$.

In summary, the second moment $E[a_2^2(t)]$ is seen to be expressible as a convolution of the evolutionary power spectral density $S(\omega, t)$ of the excitation process $\ddot{z}(t)$ with the nonnegative function $\mu(t)$. In order to obtain the probability density function of $a_2(t)$, moments of all orders must be computed. However, it is reasonable to suppose that the secondary envelope $a_2(t)$ is nearly Rayleigh distributed if the excitation is Gaussian and the response is narrow-banded. Since the Rayleigh distribution has only a single parameter, the probability density is determined by the second moment.

Response to Finite Duration Stationary Excitation

One application of the above results is to the case where the excitation is derived from a stationary process which is modulated by a rectangular function in time. The evolutionary power spectral density of the power acceleration will be taken as

$$S(\omega,t) = \begin{cases} S_0(\omega) & 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases} \tag{31}$$

where $S_0(\omega)$ is the power spectral density of the underlying stationary process, and T is the duration of shaking. Neglecting terms of higher than first order in ζ_1 and ζ_2, it may be shown that

$$E[a_2^2(t)] = S_0(\omega_\lambda)\rho(t) \tag{32}$$

where

$$\rho(t) = \begin{cases} \rho^*(t) & 0 \leq t \leq T \\ \rho^*(t) - \rho^*(t-T) & t > T \end{cases} \tag{33}$$

$\rho^*(t)$ describes the buildup of the secondary response from rest to stationarity, when T is made arbitrarily large.

$$\rho^*(t) \approx \frac{\pi\omega_1^2}{(\Delta\omega)^2\omega_2^2(\omega_\lambda+\omega_u)^2}\Big\{\frac{\omega_u^2}{\zeta_\lambda\omega_\lambda}(1-e^{-2\zeta_\lambda\omega_\lambda t})$$
$$+\frac{\omega_\lambda^2+\omega_u^2}{2\zeta_u\omega_u}(1-e^{-2\zeta_u\omega_u t})$$
$$-\frac{2\omega_u}{(\omega_\lambda-\omega_u)^2+\omega_\lambda\omega_u(\zeta_\lambda+\zeta_u)^2}$$
$$[2(\zeta_\lambda\omega_\lambda^2+\zeta_u\omega_u^2)(1-e^{-(\zeta_\lambda\omega_\lambda+\zeta_u\omega_u)t}\cos(\omega_\lambda-\omega_u)t)$$
$$+(\omega_\lambda^2-\omega_u^2)e^{-(\zeta_\lambda\omega_\lambda+\zeta_u\omega_u)t}\sin(\omega_\lambda-\omega_u)t]\Big\} \tag{34}$$

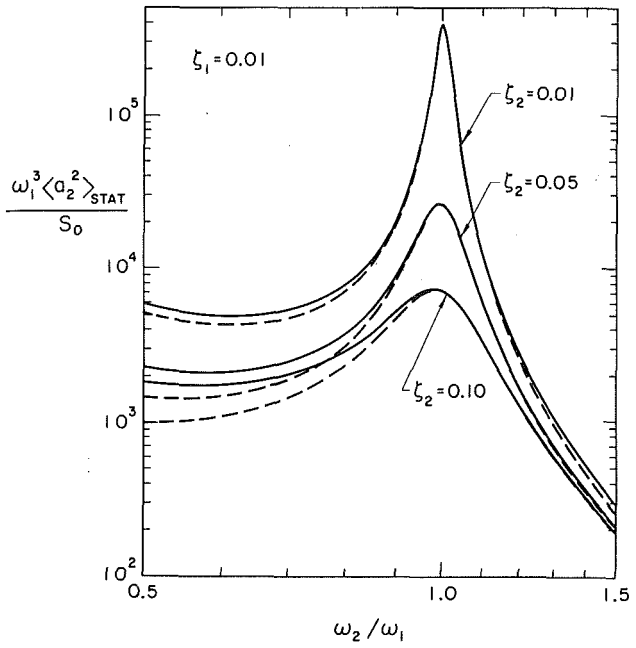


Fig. 2 Comparison of stationary mean square secondary envelope (solid) with narrow-band prediction (dashed): $\zeta_1 = 0.01$; $\zeta_2 = 0.01, 0.05, 0.10$

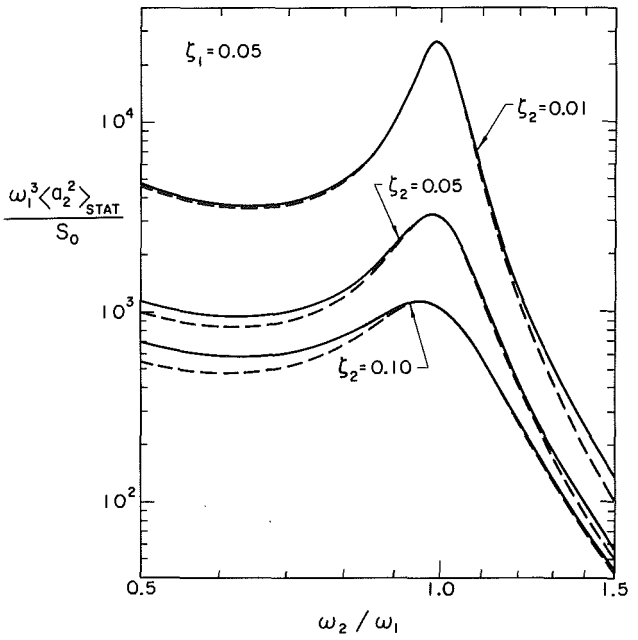


Fig. 3 Comparison of stationary mean square secondary envelope (solid) with narrow-band prediction (dashed): $\zeta_1 = 0.05$; $\zeta_2 = 0.01, 0.05, 0.10$

If $\Delta\omega = 0$, equation (34) is interminate. The corresponding result for this case is

$$\rho^*(t) \approx \frac{\pi}{8\zeta_1^3\omega_1^3}\left[1-(1+2\zeta_1\omega_1 t+2\zeta_1^2\omega_1^2 t^2)e^{-2\zeta_1\omega_1 t}\right] \quad (35)$$

If T is large, the system will eventually achieve stationarity. At stationarity

$$E[a_2^2]_{stat} = \rho^*(\infty)\, S_0(\omega_\lambda) \quad (36)$$

If the secondary response is narrow-banded, then

$$\langle y^2 \rangle_{stat} \approx \tfrac{1}{2} \langle a_2^2 \rangle_{stat} \quad (37)$$

The stationary variance of the displacement of a secondary system driven by white noise may be shown to be (Crandall and Mark, 1963)

$$E[y^2]_{stat} = \frac{\pi S_0}{2\zeta_2\omega_2^3}\,\frac{\omega_1^4+\omega_1\omega_2^3(\zeta_2/\zeta_1)+A}{(\omega_1^2-\omega_2^2)^2+A} \quad (38)$$

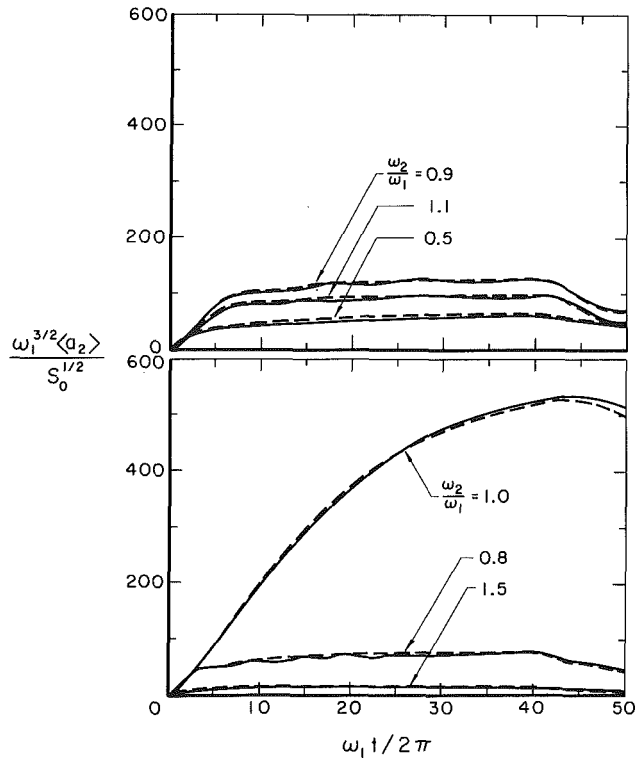


Fig. 4 Comparison of mean secondary envelope (dashed) with simulation (solid): $\zeta_1 = \zeta_2 = 0.01$; $\omega_1 T = 80\,\pi$

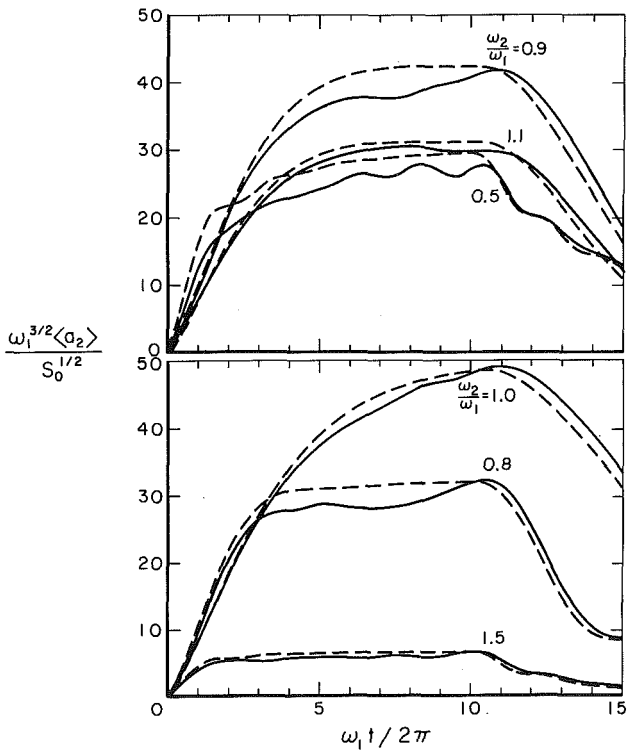


Fig. 5 Comparison of mean secondary envelope (dashed) with simulation (solid): $\zeta_1 = \zeta_2 = 0.05$; $\omega_1 T = 20\,\pi$

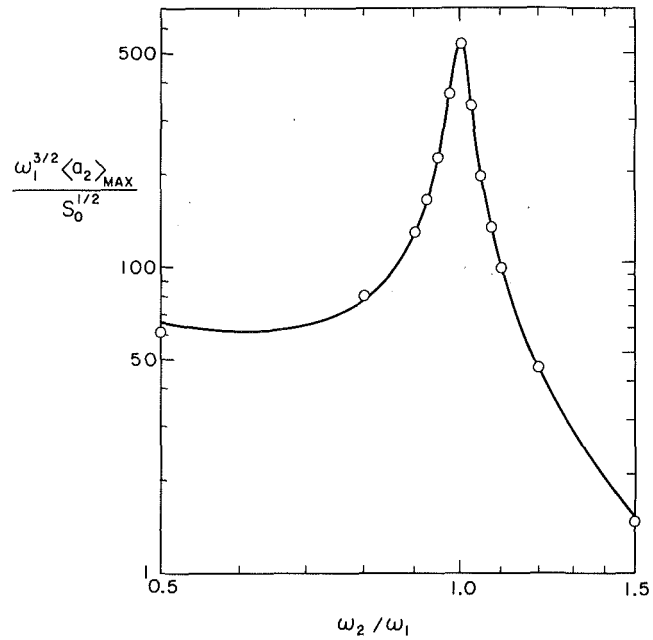


Fig. 6 Comparison of maximum mean secondary envelope with simulation (circles): $\zeta_1 = \zeta_2 = 0.01$; $\omega_1 T = 80\pi$

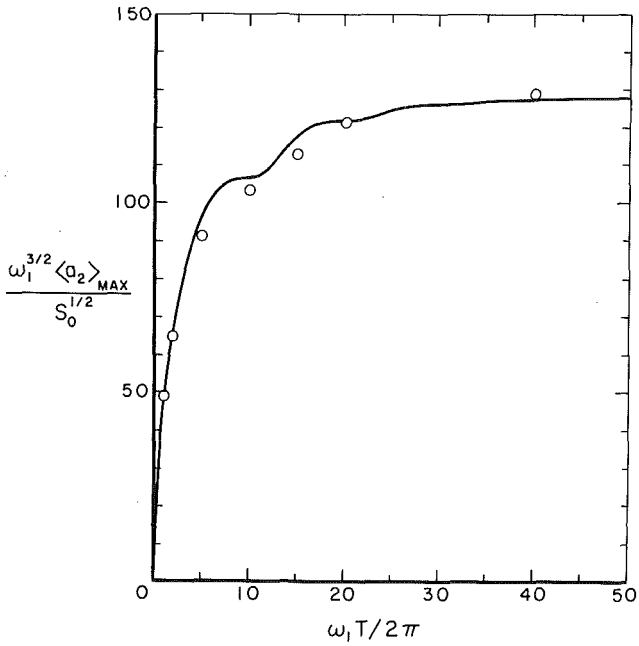


Fig. 8 Comparison of maximum mean secondary envelope with simulation (circles): $\zeta_1 = \zeta_2 = 0.01$; $\omega_1 T = 80\pi$

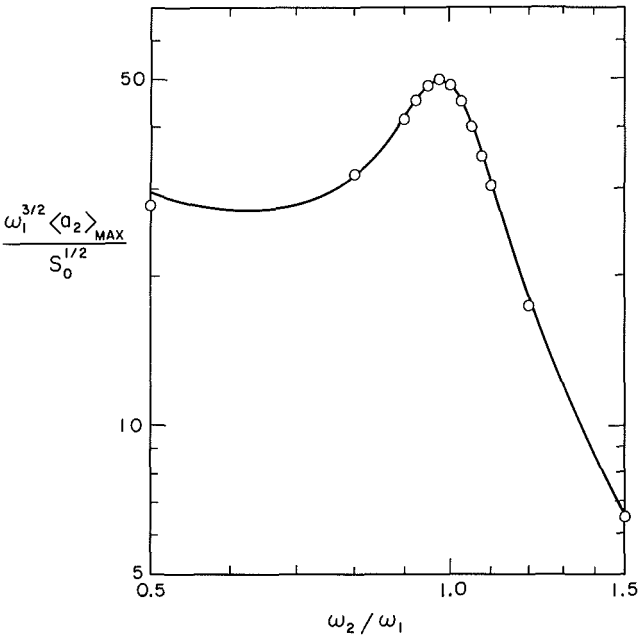


Fig. 7 Comparison of maximum mean secondary envelope with simulation (circles): $\zeta_1 = \zeta_2 = 0.05$; $\omega_1 T = 20\pi$

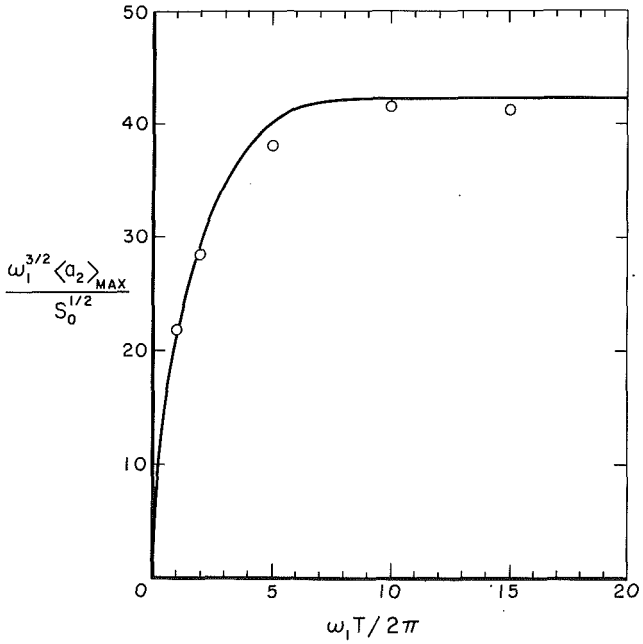


Fig. 9 Comparison of maximum mean secondary envelope with simulation (circles): $\zeta_1 = \zeta_2 = 0.05$; $\omega_1 T = 20\pi$

where

$$A = 4\omega_1\omega_2[\zeta_1\zeta_2(\omega_1^2+\omega_2^2)+\omega_1\omega_2(\zeta_1^2+\zeta_2^2)] \tag{39}$$

A comparison of equations (37) and (38) shows asymptotic agreement as $\omega_2/\omega_1 \to 0$, as $\omega_2/\omega_1 \to 1$, and as $\omega_2/\omega_1 \to \infty$. In these cases, the assumption of narrow-bandedness is valid. For intermediate values of ω_2/ω_1, however, $y(t)$ has a bimodal behavior.

Figures 2 and 3 show a comparison of the actual stationary response and that predicted by the narrow-bandedness assumption. It may be seen that the narrow--bandedness assumption generally leads to an underestimation of the mean square envelope response.

Comparison With Simulation Results

In order to verify the analytical results a series of numerical simulations were performed. The average value of $a_2(t)$ for 250 samples is shown in Figs. 4 and 5 for various values of system parameters. The theoretical mean envelope value is shown for comparison. In general, the theory gives conservative results.The difference is most pronounced at higher levels of damping and away from resonance, when the bimodality of $y(t)$ is greatest.

The intensity of response may be characterized by the maximum achieved by the mean value of the envelope. Figures 6 and 7 show a comparison of simulated and theoretical values for $\max_{0\le t<\infty} E[a_2(t)]$, as a function of ω_2/ω_1 for fixed duration. Figures 8 and 9 show the same comparison as a function of

duration for fixed ω_2/ω_1. In all cases, the theoretical values correspond closely to those obtained by simulation.

Interaction Effects

So far, it has been assumed that the mass of the secondary system is negligible in comparison with the mass of the primary system, so that the response of the primary system is unaffected by the presence of the secondary system. This assumption simplifies the analysis considerably by reducing the combined system to two chained single-degree-of-freedom systems.

As the mass ratio increases, the noninteraction approximation becomes progressively worse, particularly if $\omega_1 \approx \omega_2$, Ignoring interaction effects generally leads to conservative estimates of system response, since the secondary system will actually absorb energy from the primary system. However, in some cases (especially near resonance), the assumption of noninteraction leads to gross overestimates of system response.

The equations of motion for the two-degree-of-freedom primary/secondary system, including interaction effects, are:

$$\ddot{x}+2\zeta_1\omega_1\dot{x}+\omega_1^2 x-\epsilon(2\zeta_2\omega_2\dot{y}+\omega_2^2 y)=-\ddot{z}$$
$$\ddot{y}+2\zeta_2\omega_2\dot{y}+\omega_2^2 y=-\ddot{z}-\ddot{x} \tag{40}$$

where $\epsilon = m_2/m_1$ is the mass ratio. If $\zeta_1\omega_2 = \zeta_2\omega_1$, it is possible to uncouple these equations as

$$\ddot{x}_e+2\zeta_1'\omega_1'\dot{x}_e+(\omega_1')^2 x_e=-\ddot{z}$$
$$\ddot{y}_e+2\zeta_2'\omega_2'\dot{y}_e+(\omega_2')^2 y_e=-\ddot{z}-\ddot{x}_e \tag{41}$$

where

$$\zeta_1'=\sqrt{1+\theta}\,\zeta_1$$
$$\omega_1'=\sqrt{1+\theta}\,\omega_1$$
$$\zeta_2'=\sqrt{\frac{1}{1+\theta}}\,\zeta_2 \tag{42}$$
$$\omega_2'=\sqrt{\frac{1}{1+\theta}}\,\omega_2$$

and

$$x(t)=x_e(t)+\frac{\theta}{1+\theta}\,y_e(t)$$
$$y(t)=\frac{1}{(1+\theta)}\,y_e(t) \tag{43}$$

with

$$\theta=\tfrac{1}{2}\left[(1+\epsilon)\left(\frac{\omega_2}{\omega_1}\right)^2-1\right]+\operatorname{sgn}(\omega_1^2-\omega_2^2)$$
$$\left\{\tfrac{1}{4}\left[(1+\epsilon)\left(\frac{\omega_2}{\omega_1}\right)^2-1\right]^2+\epsilon\left(\frac{\omega_2}{\omega_1}\right)^2\right\}^{1/2} \tag{44}$$

The actual (interacting) secondary envelope is therefore $1/(1+\theta)$ times the secondary envelope of the equivalent noninteracting system with system parameters modified as in equations (42). Thus, the results already obtained for the noninteracting system can be carried over to the interacting system simply by modifying the system parameters and scaling the response.

Figure 10 shows θ as a function of the mass ratio ϵ, for several values of ω_2/ω_1. It may be seen that θ is positive when $\omega_2 < \omega_1$, and negative when $\omega_2 > \omega_1$. It may be shown that the natural frequencies of the equivalent noninteracting system in equation (42) are more widely spaced than those of the original system. This effect is most prominent near resonance, where a slight "detuning" may have a large influence on the response. By comparison, the changes in damping and the scaling of $y(t)$ are much less significant.

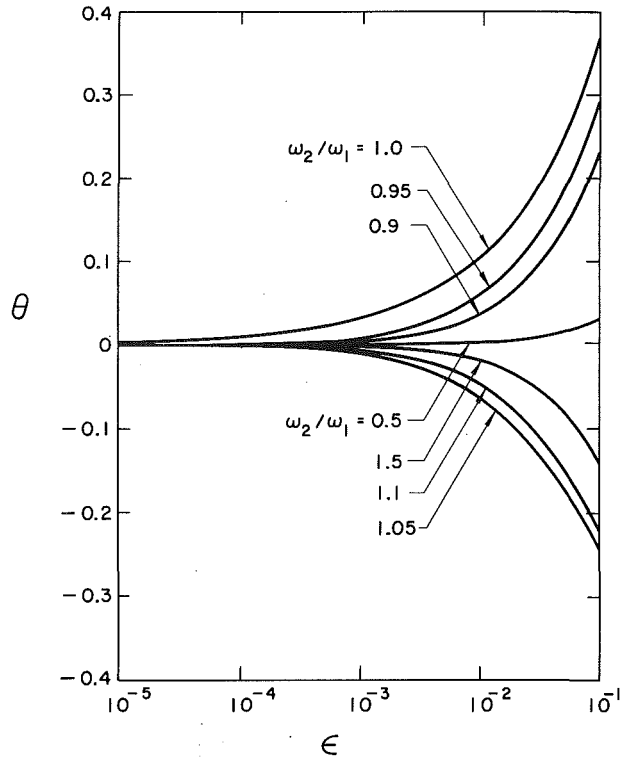


Fig. 10 Interaction parameter θ versus mass ratio ϵ for several values of ω_2/ω_1

Near resonance, for small θ, interaction effects may be neglected if $|\theta| << \zeta_1$. At perfect resonance interaction may be neglected if $\epsilon << \zeta_1^2$. Note that the restriction to classical damping requires $\zeta_1 \approx \zeta_2$ when $\omega_1 \approx \omega_2$.

Away from resonance, the detuning is less significant, so $\epsilon << \zeta_1^2$ remains a conservative criterion. If $\zeta_1 = \zeta_2 = 0.05$, this requires that the secondary system's mass be substantially less than 1/400 of the primary system's mass for interaction effects to be ignored. If the two-degree-of-freedom system is not classically damped, the above analysis does not apply. In this case, there seems to be no simplification of the interacting system, and the complete system must be considered. The stationary analysis (Igusa and Der Kiureghian, 1985) indicates that the effect of interaction may be neglected if

$$\epsilon << \zeta_1\zeta_2 \tag{45}$$

In any event, the assumption of noninteraction gives conservative results when ϵ is small.

Summary and Conclusions

An approximate closed form solution has been presented for the mean square envelope response of a secondary system which is attached to a noninteracting primary system subjected to nonstationary random excitation. The predictions of the closed form solution agree well with results of numerical simulation. The analytical solution is useful in the analysis of composite systems subjected to transient random excitation, such as earthquake loading.

Under certain restrictions, the effects of primary-secondary system interaction have been shown to be expressible primarily in terms of a frequency detuning of the two system subelements. This simple result provides useful qualitative insight into the nature of interaction effects.

Acknowledgments

This investigation was sponsored by Grant No. CEE81-19962 from the National Science Foundation. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Chakravorty, M. K., and Vanmarcke, E. H., 1973, "Probabilistic Seismic Analysis of Light Equipment Within Buildings," *Proceedings of the 5th World Conference on Earthquake Engineering*, Rome, Italy.

Crandall, S., Mark, W., 1963, *Random Vibration in Mechanical Systems*, Academic Press, New York, Chapter 2.

Igusa, T., and Der Kiureghian, A., 1983, "Dynamic Analysis of Multiply Tuned and Arbitrarily Supported Secondary Systems," Report No. UCB/EERC-83/07, Earthquke Engineering Research Center, University of California, Berkeley, CA.

Igusa, T., and Der Kiureghian, 1985, "Generation of Floor Response Spectra Including Oscillator-Structure Interaction," *Earthquake Engineering and Structural Dynamics*, Vol. 13, pp. 661-676.

Penzien, J., and Chopra, A. K., 1965, "Earthquake Response of Appendage on a Multi-story Building," *Proceedings of the 3rd World Conference on Earthquake Engineering*, New Zealand.

Priestley, M. B., 1965, "Evolutionary Spectra and Non-Stationary Processes," *Journal of Royal Statistical Society*, Vol. 27, pp. 204-228.

Singh, M. P., 1975, "Generation of Seismic Flow Spectra," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 101, No. EM5, pp. 563-607.

Smith, Kenneth Scott, 1985, "Stochastic Analysis of the Seismic Response of Secondary Systems," Ph.D. Thesis, California Institute of Technology, EERL Report 85-01, Pasadena, CA.

Spanos, P-T. D., and Lutes, L. D., 1980, "Probability of Response to Evolutionary Process," *Joural of the Engineering Mechanics Division*, ASCE, Vol. 106, No. EM2, pp. 213-224.

Characteristics of Numerical Simulations of Chaotic Systems[1]

B. H. Tongue
School of Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332
Assoc. Mem. ASME

A generalized form of Duffing's Equation is examined in order to gain insight into the characteristics and properties of chaotic motion. It is shown that variations in the forcing function parameters as well as variations in the system's initial conditions can lead to a chaotic response. The incidence of chaos is presented in the form of chaos maps and the structure of these maps is discussed. The influence of linear spring force on these maps is also examined. Finally, it is shown that an improper choice of time step can cause spurious results with regard to the existence of chaotic motion.

Introduction

It has been realized for over two hundred years that although most meaningful problems are nonlinear in nature, the difficulty associated with solving such problems makes an analytical study of them impractical. This lack of tractability has serious implications due to the fact that nonlinear equations often do not behave as intuition might indicate. The variety of behaviors that nonlinear systems can generate is very rich and is only hinted at by a linear analysis.

One partial solution to this problem has been the development of various approximate techniques (Bogoliubov and Mitropolsky, 1961; Hagedorn, 1981; Minorsky, 1983; Nayfeh and Mook, 1979). Such techniques usually assume that the system nonlinearities are small and often are based upon an assumed form of the actual solution. These techniques (harmonic balance, multiple scales, etc.) have allowed investigators to uncover some essential details regarding the response of nonlinear systems. However, these methods are limited and only apply to a restricted range of parameter values. None of the techniques even suggest the existence of chaotic motions (Holmes, 1979; Ueda and Akamatsu, 1981; Moon and Shaw, 1983; Tongue, 1986), the subject of this paper.

One of the primary reasons for examining chaotic systems is that they are capable of producing regular, well behaved, periodic outputs which are easily handled by approximate techniques as well as outputs that appear random and which are not well approximated by any analytical techniques. As more and more researchers start investigations of chaos, it becomes important to make certain that the problems associated with numerical analyses of chaotic systems are clearly understood. Unfortunately, most papers presume a prior knowledge both of chaos from a theoretical framework as well as from a computational one, making it difficult for the beginning researcher to be aware of the different problems that can occur. The following section addresses these problems.

Computational Difficulties

There exists an essential difference between exact analytical solutions of differential equations and numerical integrations of the equations, namely that analytical solutions are continuous in time whereas numerical integrations involve finite time steps. Because of this, numerical integrations are best regarded as the solutions of difference equations rather than differential equations. For small time steps the difference between the differential and difference equation is presumed to vanish. With regular, well behaved, linear equations, the question of an appropriate time step for the integration is easily handled. One simply determines what is the fastest time constant of interest (highest natural frequency for a modal analysis) and chooses a time step small enough to accurately track this. Thus the time step might be chosen so that thirty time increments equals the period of the highest frequency.

When dealing with chaotic systems this question is not so easily handled. Since the overall response contains a continuous spectra of frequencies, it is not clear how small a time increment is needed. By choosing incorrectly, qualitative errors can result. Figure 1 illustrates this problem. This figure shows two Poincaré maps for the equation:

$$\ddot{x}+0.2\dot{x}-x+x^3=0.3\cos(1.2t) \tag{1}$$

The small and large dots were generated with time steps equal to 0.2618 and 0.0654, respectively. The time discretization equal to 0.2618 has clearly given rise to a strange attractor and thus one might conclude that the system is behaving chaotically. However, that is not what occurs when the time scale is refined to 0.0654 (large dots). Now the Poincaré map

[1]This work was supported by the National Science Foundation, Grant No. MSM-8451186.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13–18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 31, 1985; final revision, January 16, 1987.

Paper No. 87-WA/APM-4.

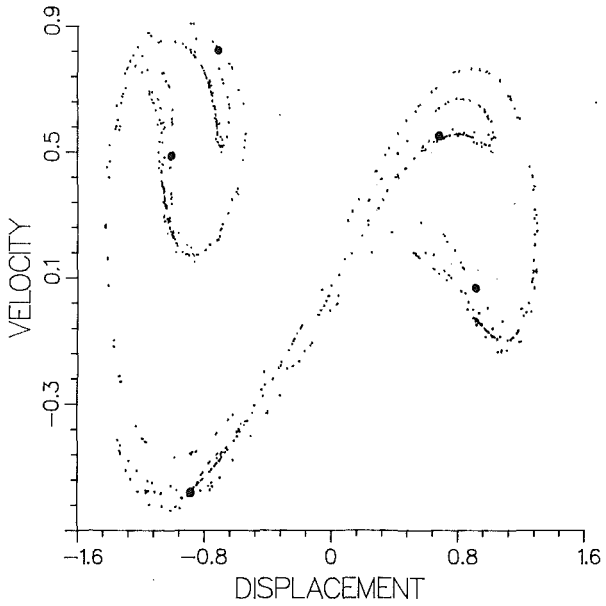


Fig. 1 Poincare map, $\Delta t = 0.2618$ and 0.0654

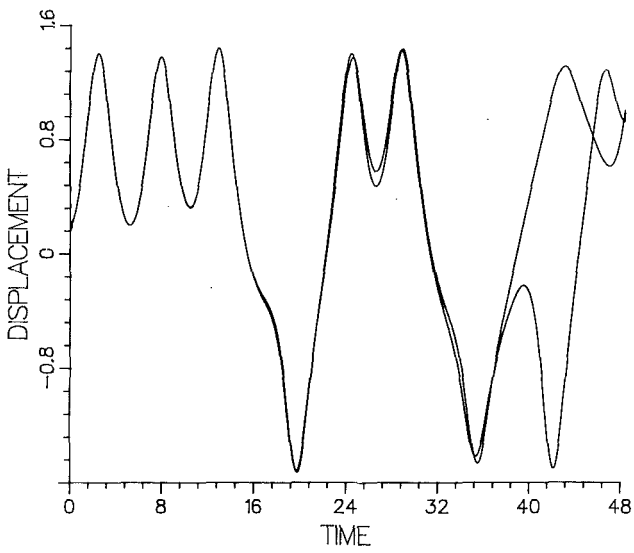


Fig. 2 Exponential divergence of trajectories

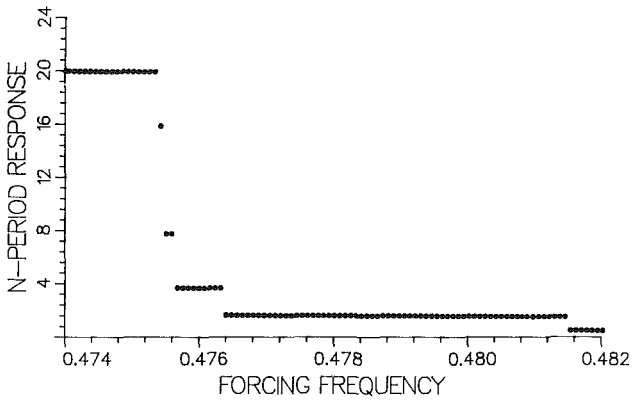


Fig. 3 Period bifurcations: $\Delta t = 0.0393/\omega$

shows that the response is actually a period five motion. Additional reductions of the integration time step cause no further changes in the output.

Two qualitatively different solutions have been generated from a single differential equation. The only parameter that distinguishes between the two cases is the length of the integration time step. The conclusion to be drawn is that chaotic responses must be viewed with care. The strange attractor of Fig. 1 is actually from spurious chaos. The true response of the governing differential equation is periodic while the response of the numerical approximation to the actual differential equation, for a time step of 0.2618, is chaotic. But it must be stressed that this means that the difference equation that has replaced the differential equation supports chaos, not the differential equation itself and presumably not the physical system that gave rise to the differential equation.

It is well known that for a chaotic system, two solutions starting arbitrarily close together will quickly diverge from each other. This observation has been used as the basis for an analytical measure of chaotic systems (Ueda, 1979; Wolf, 1985). The rate of divergence of two closely spaced points is found to be exponential in time (separation $\approx e^{L_i t}$). The values L_i are called Lyapunov exponents and they govern the rate of separation with respect to specific directions in phase space. Chaos is associated with a positive exponent, which implies that two trajectories starting infinitesimally close to each other will diverge exponentially. This is illustrated in Fig. 2 for the equation

$$\ddot{x} + 0.2\dot{x} - x + x^3 = 0.33\cos(1.2t) \tag{2}$$

An exactly analogous situation will occur for any change in the time discretization. If two trajectories, starting from identical initial conditions, are generated using slightly different time discretizations ($\Delta t = 0.0524$ and 0.0518), the same kind of time histories as seen in Fig. 2 will be generated. Note that only a one percent change in the time step is used.

At first this seems odd, since very small changes in time discretization usually lead to very small changes in the integrated output of well-behaved systems. However, the system being considered is chaotic. Although the two initial conditions were identical, the next calculated state for each trajectory will be different due to the different duration of their associated time step. Thus one should expect divergence to occur, as it did for identical discretizations and differing initial conditions. It should be noted that this same behavior can be observed even for solutions that are ultimately periodic. In this case the time history can go through a transiently chaotic stage before settling into periodicity (Yorke and Yorke, 1979).

Another interesting point to consider is how a system progresses into chaos. Work has been done that predicts a regular route to chaos (Feigenbaum, 1978, 1980). As a critical parameter in the generating equation changes, the response changes from period one motion to period two, period four, and so on. Theoretically, a limiting value of the critical parameter exists for which the period of the motion goes to infinity, at which point the response becomes aperiodic. For parameter values above this critical point, chaos can occur. This kind of response is easily generated by simple recursive systems such as the logistic equation.

A question can now be raised as to whether this simple route to chaos can occur for a nonconservative dynamical system such as the one examined in the previous part of this paper. Simulations were therefore run to obtain the response of

$$\ddot{x} + 0.1\dot{x} - x + x^3 = 3.2\cos(\omega t) \tag{3}$$

for values of frequency near the chaotic regime. The results are shown in Figs. 3–5.

The figures illustrate the kind of steady state motion that is obtained for the given time discretization at the given driving frequencies. They record the periodicity of the output (period one, period two, etc.) as a function of the driving frequency. The results shown in Fig. 3 demonstrate that simple period doubling can be observed in a dynamical system and that the

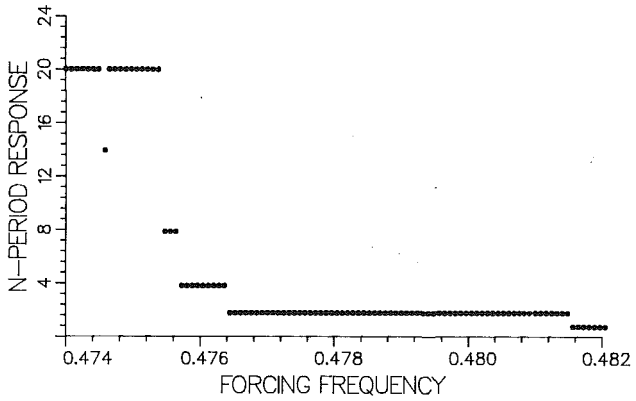


Fig. 4 Period bifurcations: $\Delta t = 0.0785/\omega$

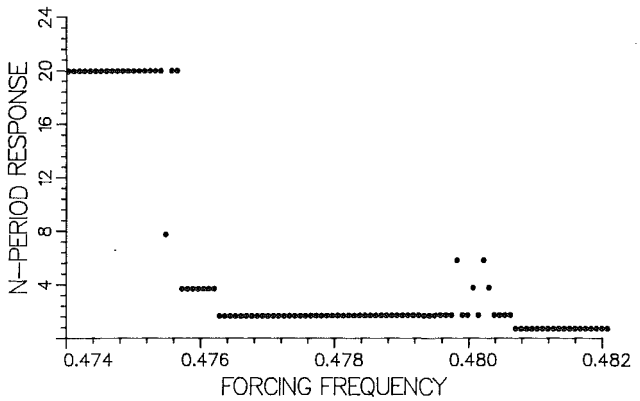


Fig. 5 Period bifurcations: $\Delta t = 0.1571/\omega$

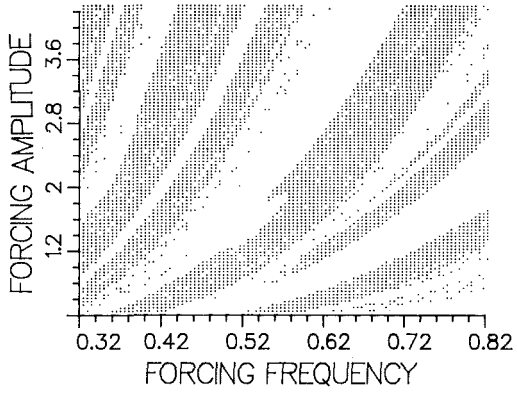


Fig. 6 Regions of chaos in parameter space: $\epsilon - 1.0$; $\Delta G = 0.04$; $\Delta\omega = 0.005$

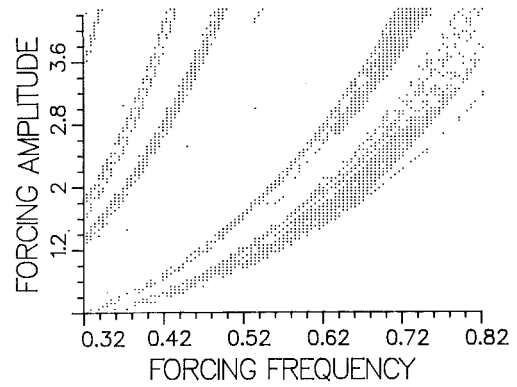


Fig. 7 Regions of chaos in parameter space: $\epsilon = 0.0$; $\Delta G = 0.04$; $\Delta\omega = 0.005$

time discretization again affects the response. As the driving frequency is decreased from $\omega = 0.482$, with a time step equal to $0.0393/\omega$, the steady state response progresses from a period one motion, to period two, four, eight, sixteen, and finally into chaos (represented by period 20 motion). Clearly, frequency doubling is occurring. If the size of the time step is increased to $0.0785/\omega$ the responses alter slightly (Fig. 4). The period sixteen motion of Fig. 3 disappears and a period fourteen motion appears in what was previously a chaotic regime. Coarsening the time step yet further (Fig. 5), to $0.1571/\omega$, causes a large change in the response and greatly obscures the period doubling that was previously so evident. Thus the observation of period doubling can depend on a proper choice of time increment. It is important to note that these results do not imply that period doubling is the only route to chaos, merely that it is one possible route.

Results

It is of interest to determine when chaos can occur for different parameter combinations. One reason to do this is to decide if chaos is a rare occurrence or a common one. The shape of the chaotic regions in parameter space may also be of help in yielding an overall understanding of the system.

Figure 6 shows the regions of chaos and the regions of periodic responses associated with different values of forcing amplitude, G, and forcing frequency, ω, for the equation

$$\ddot{x} + 0.1\dot{x} + \epsilon x + x^3 = G\cos(\omega t) \tag{4}$$

for $\epsilon = -1$. This kind of plot will be called a chaos map. The region of parameter space was divided by a mesh of spacing $\Delta G = 0.04$ and $\Delta\omega = 0.005$. One hundred different values for G and ω were used, yielding ten thousand different combinations of amplitude and frequency. For each of these combinations, equation (4) was integrated and the resultant output analyzed. If the output was chaotic, then a dot was placed on the corresponding grid point. If the response was periodic, then the grid intersection was left blank. Only the indicated ten thousand points of parameter space were investigated; one cannot say anything with absolute certainty about the infinitude of combinations that occur at untested values.

Overall qualitative trends are quite apparent. First, it seems that chaos is not a rare occurrence but occurs at a roughly equal frequency to that of periodic responses. Second, even at the present scale of mesh spacing it can be seen that periodic responses occur within the mainly chaotic regions. Also, the mainly chaotic regions are structured and all fall along curving paths from the origin. Figures 7 and 8 show how this chaos map varies for the linear spring coefficient, ϵ, equal to 0 and 0.3, respectively.

It should be recognized that these plots are closely related to the concept of basins of attraction (McDonald, 1985). Numerous researchers have tried to determine what the steady state position of a nonlinear system is as a function of some system parameters. For example, the unforced case of equation (4) will come to rest in one of two equilibrium conditions, either $x = 1.$ or $x = -1.$ One can then ask which regions of initial condition space will send the system to $x = 1.$ and which will send it to $x = -1.$ These regions are called basins of attraction. The boundary that separates the regions is the basin boundary. Grebogi et al. (1984) defines an attractor as a compact set for which almost all points in a neighborhood of the set tend toward the set as time goes to infinity. Thus in the present example we have the very simple situation in which $x = \pm 1.$ acts as the attractor set. One can further generalize the problem by introducing a forcing team. If the forcing is small,

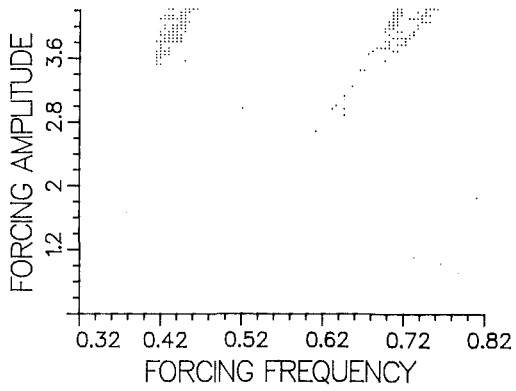


Fig. 8 Regions of chaos in parameter space: $\epsilon = 0.3$; $\Delta G = 0.04$; $\Delta\omega = 0.005$

the system trajectories will oscillate about $x = 1.$ or $x = -1.$ Therefore, basins of attraction can again be constructed on the basis of whether the final oscillations are about $x = 1.$ or $-1.$

Figures 6–8 do not show in which region of the phase space the trajectory ends but rather indicate the existence of equalitatively different global responses, namely periodic and chaotic. They give an immediate sense of how likely it is that chaos will occur, as a function of the driving parameters. Clearly chaos is more common for $\epsilon = -1.$ than for $\epsilon = 0.3$.

There is a qualitative difference between the system's static equilibria for ϵ greater than zero and ϵ less than zero. In the former case only one equilibrium exists while the latter case supports three distinct equilibria. The existence of three distinct equilibria (two stable separated by one unstable) has led some investigators (Moon, 1980) to postulate that chaos can occur when the energy flow of the system is such that the oscillation amplitude of the particle approaches the separation distance between a stable and unstable equilibrium.

The author (Tongue, 1986) has shown that this is not a necessary condition for chaos and Figs. 6–8 illustrate this quite clearly. As ϵ is increased to zero and then to positive values, the regions of chaos alter in shape and draw away from the $G = 0.$ boundary. This supports the notion that, although multiple static equilibria encourage the existence of chaos, they are not necessary. For large values of G, the mass particle is forced to be far from the origin for a great deal of the time. Thus, as far as the mass is concerned, the conditions of ϵ greater than or less than zero are essentially the same. The fine details of equilibria at $x = \pm\sqrt{\epsilon}$, 0 or three roots at $x = 0.$ are lost. However for small G (i.e., oscillation amplitudes on the order of the equilibrium spacing), the existence of separate equilibria becomes important and the interaction of the potential walls associated with them dominates the response. In this case multiple equilibria are quite important.

It should be noted that a system of the given type can sustain multiple dynamic equilibria (Hagedorn, 1981). Thus, even though the system does not have multiple static equilibria, it does have the possibility of multiple dynamic equilibria whose interaction may lead to the observed chaotic response. Finally, in addition to requiring higher levels of forcing, the overall frequency of occurrence of chaos is seen to be reduced as ϵ is increased. This implies that it is less probable that chaos will occur when multiple equilibria are absent for arbitrary choices of forcing amplitude and frequency.

As can be seen from Fig. 6, the chaotic boundaries appear to be reasonably smooth. McDonald et al. (1985) have investigated several examples of systems having basin boundaries and have observed that these boundaries are often fractal (Mandelbrot, 1977) in nature. The question then arises as to whether these chaos maps exhibit fractal behavior. Figure 9 is a chaos map for the restricted region of Fig. 6 in which $0.52 \le \omega \le 0.56$ and $0.8 \le G \le 1.2$. The scale of resolution has

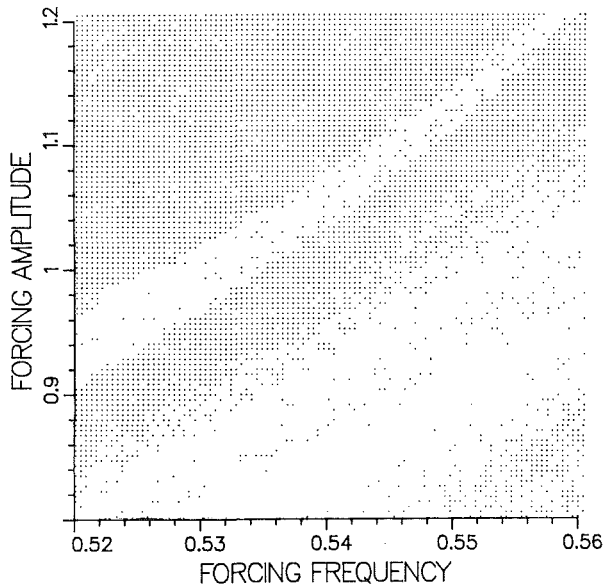


Fig. 9 Regions of chaos in parameter space: $\epsilon = -1.0$; $\Delta G = 0.004$; $\Delta\omega = 0.0004$

been increased by a factor of 125. While the two upper boundaries appear relatively smooth, the others seem quite ragged. An increase in surface detail with increasing resolution is one characteristic of fractal structures. Thus the lower boundaries seem fractal-like whereas the upper ones are relatively smooth; if they are fractal their dimension would appear to be close to unity. A comprehensive answer to the question of the dimension of the boundaries will require the use of fractal interpolation (Barnsley) or of covering sets (Mandelbrot, 1977) and will be accomplished in a later paper.

As a final observation it should be noted that, although the system under examination in this paper appears to be reasonably simple, there are a total of eight parameters that must be taken into account. The general equation is seen to be:

$$\ddot{x} + \sigma\dot{x} + \epsilon x + \alpha x^3 = G\cos(\omega t + \theta) \tag{5}$$

$$\text{I.C.: } x(0) \text{ and } \dot{x}(0)$$

The eight parameters are σ, ϵ, α, G, ω, θ, $x(0)$, and $\dot{x}(0)$.

Of course, it is possible to nondimensionalize in order to reduce the total parameter set. If one defines $\tau = \sqrt{\epsilon}\ t$ and $x = (\epsilon/\alpha)^{1/2}\ y$ then equation (5) becomes

$$y'' + \bar{\sigma}y' + y + y^3 = \bar{G}\cos(\bar{\omega}\tau + \theta)$$

where

$$(\ )' \equiv \frac{d(\ )}{d\tau},\quad \bar{\sigma} = \frac{\sigma}{\sqrt{\epsilon}},\quad \bar{G} = \frac{G\alpha^{1/2}}{\epsilon^{3/2}} \text{ and } \bar{\omega} = \frac{\omega\tau}{\sqrt{\epsilon}}.$$

The difficulty with such a reduction is that it restricts the analysis to ϵ and α values of the same sign and won't permit $\epsilon = 0$. Since both of these cases are of interest, the anlysis will treat the form of equation (5).

It is possible to express this problem as a coupled second order and first order equation:

$$\ddot{x} + \sigma\dot{x} + \epsilon x + \alpha x^3 - G\cos(\eta) = 0 \quad \text{I.C.} = x(0), \dot{x}(0)$$

$$\dot{\eta} = \omega \quad \text{I.C.} = \eta(0)$$

However, instead of expressing the system as two coupled equations, an equivalent fourth order equation can be found[2] in which the system's dependent variable's initial conditions are clearly related to $x(0)$, $\dot{x}(0)$, G and ω. Since these are the

[2]This alternate formulation, used in a different context, was pointed out to the author by Dr. Gary Anderson of the Army Research Office during a conference on nonlinear systems, held at the Georgia Institute of Technology.

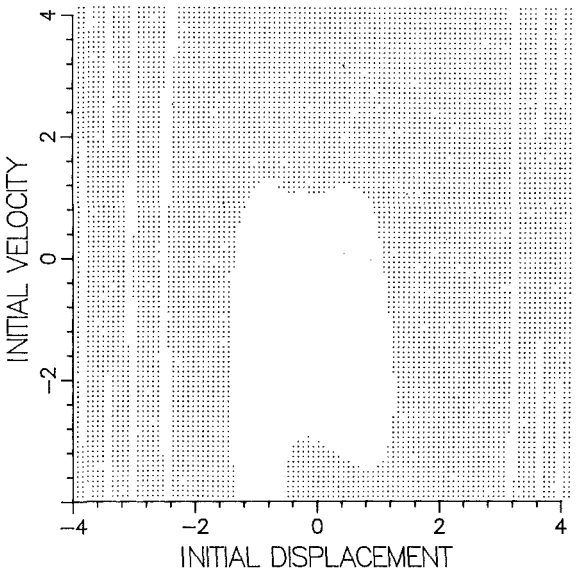


Fig. 10 Regions of chaos in initial condition space: $\Delta x = 0.08$, $\Delta \dot{x} = 0.08$

parameters of main interest in this paper, this final formulation shall be presented. To start, express equation (5) as:

$$\ddot{x} + \sigma\dot{x} + \epsilon x + \alpha x^3 = q; \quad \text{I.C.:} x(0), \dot{x}(0) \tag{6}$$

$$\ddot{q} + \omega^2 q = 0; \quad \text{I.C.:} q(0), \dot{q}(0) \tag{7}$$

Clearly, a proper choice of $q(0)$ and $\dot{q}(0)$ will yield a solution for $\theta(t)$ that will match $G\cos(\omega t + \theta)$ for any G, θ combination.

Differentiating equation (6) twice yields an expression for $\ddot{q}$ which when substituted into equation (7) along with the expression for q will yield the following fourth order homogeneous equation:

$$x^{IV} + \sigma\dddot{x} + \ddot{x}(\epsilon + \omega^2) + 6\alpha\dot{x}^2 x + 3\alpha\ddot{x}x^2 + \sigma\omega^2\dot{x} + \epsilon\omega^2 x + \omega^2\alpha x^3 = 0 \tag{8}$$

$$\text{I.C.: } x(0), \dot{x}(0), \ddot{x}(0), \dddot{x}(0)$$

The appropriate initial conditions for $\ddot{x}$ and x are seen to be:

$$\ddot{x}(0) = q(0) - \alpha x^3(0) - \epsilon x(0) - \sigma\dot{x}(0) \tag{9}$$

$$\dddot{x}(0) = \dot{q}(0) - 3\alpha x^2(0)\dot{x}(0) - \epsilon\dot{x}(0) - \sigma\ddot{x}(0) \tag{10}$$

The choice of a homogeneous or nonhomogeneous form is thus seen to be arbitrary and can be based on convenience in regard to physical interpretation. It is thus interesting to note that the forcing parameters (G and ω) can be considered to be serving the same purpose as the initial conditions ($x(0)$ and $\dot{x}(0)$).

It has been shown that variations in G and ω can give rise to a chaotic response. Since variations in G and ω will cause variations in $\ddot{x}(0)$ and $\dddot{x}(0)$ for the fourth order formulation of equation (5) it should not be surprising that variations in $x(0)$ and $\dot{x}(0)$ will also produce chaos (McDonald, 1985). This is illustrated in Fig. 10, which examines

$$\ddot{x} + 0.1\dot{x} + x + 3.8x^3 = 12\cos(1.25t)$$

for various initial conditions $x(0)$ and $\dot{x}(0)$.

Figure 10 shows an initial condition chaos map for a fairly wide range of initial conditions. As can be seen, there are again predominantly chaotic regions and predominantly periodic regions. This map does not have the same ordered form as that exhibited by the G,ω chaos maps.

It is clear that only periodic responses result for a wide range of initial conditions ($-1.5 \le x(0) \le 1.$ and $-3. \le \dot{x}(0) \le 1.$). However, beyond this region chaos begins to predominate. This supports the supposition that the existence of chaos depends not only on the governing equation's parameters but also on the system's initial conditions. Thus one cannot simply look at a G,ω chaos map when trying to determine the probability of chaos but must also take the system initial conditions into account. Of course this type of behavior is in keeping with classical nonlinear analysis, for which initial conditions play a key role in determining a system's steady state response.

Conclusions

1 Classical techniques are not adequate to predict or describe chaotic motion.

2 Chaos does not always occur but depends on an appropriate combination of system parameters.

3 Numerical integration can give spurious results with regard to the existence of chaos due to insufficiently small time steps.

4 Chaos maps occur for both positive and negative linear stiffness coefficients. Chaotic behavior is more probable for the case of negative linear stiffness.

References

Barnsley, M. F., "Fractal Functions and Interpolation," *Journal of Constructive Approximation*, to appear.

Bogoliubov, N. N., and Mitropolsky, Y. A., 1961, *Asymptotic Methods in the Theory of Nonlinear Oscillations*, Hindustan Publishing Co.

Feigenbaum, M. J., 1980, "Universal Behavior in Nonlinear Systems," *Los Alamos Science*, Summer, pp. 2–27.

Feigenbaum, M. J., 1978, "Quantitative Universality for a Class of Nonlinear Transformations," *Journal of Statistical Physics*, Vol. 19, pp. 25–52.

Grebogi, C., et al., 1984, "Strange Attractors That Are Not Chaotic," *Physica*, Vol. 13D, pp. 261–268.

Hagedorn, P., 1981, *Nonlinear Oscillations*, Clarendon Press.

Holmes, P. J., 1979, "A Nonlinear Oscillator with a Strange Attractor," *Philosophical Transactions of the Royal Society*, Vol. 292, No. A 394, pp. 419–448.

Kailath, T., 1980, *Linear Systems*, Prentice-Hall, Inc.

Mandelbrot, B. B., 1977, *Fractals: Form, Change and Dimension*, W. H. Freeman, San Francisco, CA.

McDonald, S., et al., 1985, "Fractal Basin Boundaries," *Physica*, Vol. 17D, pp. 125–153.

Minorsky, N., 1983, *Nonlinear Oscillations*, Krieger Publishing Co.

Moon, F. C., and Shaw, S., 1983, "Chaotic Vibrations of a Beam with Non-Linear Boundary Conditions," *Int. J. of Non-Linear Mechanics*, Vol. 18, No. 6, pp. 465–477.

Nayfeh, A. H., and Mook, D. T., *Nonlinear Oscillations*, Wiley Interscience, New York.

Tongue, B. H., 1986, "Existence of Chaos in a One-Degree-of-Freedom System," *Journal of Sound and Vibration*, Vol. 110, No. 1., pp. 69–78.

Ueda, Y., 1979, "Randomly Transitional Phenomena in the System Governed by Duffing's Equation," *Journal of Statistical Physics*, Vol. 2, No. 2, pp. 181–197.

Ueda, Y., and Akamatsu, N., 1981, "Chaotically Transitional Phenomena in the Forced Negative Resistance Oscillator," *IEEE Transactions on Circuits and Systems*, Vol. CAS-28, No. 23, pp. 217–223.

Wolf, A., "Quantifying Chaos with Lyapunov Exponents," to appear in *Nonlinear Science: Theory and Applications*, Manchester University Press.

Wolf, A., et al., 1985, "Determining Lyapunov Exponents from a Time Series," *Physica*, Vol. 16D, pp. 285–317.

Yorke, J. A., and Yorke, E. D., 1979, "Metastable Chaos: The Transition to Sustained Chaotic Behavior in the Lorenz Model," *Journal of Statistical Physics*, Vol. 21, No. 3, pp. 263–277.

A New Method for Finding Symmetric Form of Asymmetric Finite-Dimensional Dynamic Systems

Mehdi Ahmadian
Assistant Professor,
Assoc. Mem. ASME

Shui-Hang Chou
Graduate Research Assistant.

Department of Mechanical Engineering,
Clemson University,
Clemson, South Carolina 29634-0921

A subclass of general lumped-parameter dynamic systems which can be transformed into an equivalent symmetric form is considered here. For the purpose of the present study, these systems are divided into two categories: those without velocity dependent forces (pseudo-conservative systems) and those with velocity dependent forces (pseudo-symmetric systems). For each category, the results on symmetrizability of matrices are used to develop an effective, systematic technique for computing the coordinate system in which the system is symmetric. The primary advantages of the technique presented in this study are twofold. First, it is computationally efficient and stable. Second, it can effectively handle systems with many degrees-of-freedom, unlike the trial and error approach suggested in previous studies.

Introduction

Lumped-parameter dynamic systems with symmetric matrix coefficients have been studied for many decades. The development of results which analyze the overall dynamic behavior of such systems via studying the properties of the coefficient matrices are abundant and well documented, for example those by Lord Rayleigh (1945), Foss (1958), Caughey (1960), Caughey and O'Kelly (1965), Moran (1970), Walker and Schmitendorf (1973), Meirovitch (1980), and Inman and Andry (1980). However, systems that are under the influence of general types of forces, which possess asymmetric matrix coefficients, are often encountered, such as those described by the vector differential equation

$$M\ddot{x}+(C+G)\dot{x}+(K+E)x=0. \tag{1}$$

Here, M is a nonsingular matrix indicating the inertial forces, C and K are symmetric matrices representing the dissipative and conservative forces, and G and E are skew-symmetric matrices denoting the gyroscopic and follower forces. Usually, equation (1) is rewritten in a simpler form as

$$\ddot{x}+A\dot{x}+Bx=0, \tag{2}$$

where $A = M^{-1}(C+G)$ and $B = M^{-1}(K+E)$ are real asymmetric matrices. Systems of this nature arise in many areas of engineering and have received increased attention in recent years with the advent of modern control techniques for dynamic systems.

One way to analyze various dynamic aspects of a system described by equation (2) is to transform it into a state space form and solve for the eigenvalues of the state matrix. Although many eigensolution techniques are available to handle this task effectively, it is usually desirable to analyze the dynamic behavior of asymmetric systems through studying the properties of the coefficient matrices, just as symmetric systems. The advantages of such an approach are twofold. First, it involves computations in n-dimensional space as opposed to $2n$-dimensional space needed for the state-space approach. Second, a physical sense for the problem is preserved, since the results are direct in terms of the coefficient matrices. Consequently, several studies have focused on finding such results; to name a few one can cite those by Mingori (1970), and Ahmadian and Inman (1984, 1986). However, one point is evident in all of these studies: the asymmetric nature of the coefficient matrices in equation (2) does not allow for results that are as sharp and clear as those existing for symmetric systems.

To eliminate the problem of dealing with asymmetric matrices, different ideas have been proposed. One idea suggests ignoring the asymmetries and simply approximating the system as a symmetric system, so that the results developed for symmetric systems can be applied. Although this approach is simple in nature, it suffers from a major defect in the sense that the approximation could lead to intolerable errors, which would make the analysis useless. Another idea suggests reducing the system to a symmetric form via a similarity transformation. Since this approach is exact, it is more attractive and has been the focus of several studies such as those by Huseyin

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Boston, MA, December 13-18, 1987, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May 16, 1986; final revision, October 15, 1986.

Paper No. 87-WA/APM-1.

and Leipholz (1973), Huseyin (1978), Inman (1983), and Ahmadian and Inman (1985).

Huseyin (1978) presents conditions under which a system without velocity-dependent forces (i.e., $A=0$ in equation (2)) can be transformed into a symmetric form; he classifies this class of systems as *pseudo-conservative* systems. In another study (Inman, 1983), Inman considers systems with velocity-dependent forces, and he takes advantage of the results presented by Taussky (1968) and Zassenhaus (1969) to show that if the coefficient matrices A and B are both symmetrizable and they have at least one symmetric factor in common, then the system can be transferred to a symmetric form via a similarity transformation.

The advantages of such a transformation are obvious. Now, the well-known results of symmetric systems can be used to analyze various dynamic aspects of an asymmetric system based on the properties of its coefficient matrices. However, the major difficulty which still remains is the computation of the coordinate system in which the system is symmetric. The problem is in determining the symmetric factors of the asymmetric coefficient matrices. Huseyin (1978) recognizes this problem and illustrates a method for systems with 2 degrees-of-freedom. For systems of higher order, he proposes a trial and error approach based on a special form of the symmetric factors of the coefficient matrix B in equation (2). Inman (1983) also realizes this problem, and he also suggests a trial and error approach for determining the symmetric form of the system (2).

However, a trial and error strategy is hardly sufficient for most systems, since it is inefficient and can result in serious computational difficulties even for low-order systems. In addition, there is no assurance that a trial and error approach can yield a solution. Consequently, a more efficient, systematic technique is needed.

The study presented here discusses a successful computational technique for determining a transformation which yields the symmetric form of an asymmetric system. Two examples are used to illustrate the effectiveness of the results. It is shown that the developed method is capable of handling systems with many degrees-of-freedom, unlike the trial and error approach suggested previously.

Background

It has been shown in the past (Tausskey, 1968) that a real asymmetric matrix is symmetrizable if and only if it satisfies any of the following conditions:

1. It has real eigenvalues and a complete set of real eigenvectors;
2. it is similar to a symmetric matrix;
3. it is similar to its transpose via a symmetric, positive definite transformation (i.e., $P = Q^{-1}P^TQ$); and
4. it becomes symmetric when multiplied by a suitable positive definite matrix.

A symmetrizable matrix can always be expressed as the product of two symmetric matrices, one of which is positive definite, i.e.,

$$P=P_1P_2;\ P_1=P_1^T>0,\ P_2=P_2^T,$$

where P_1^T indicates the transpose of the matrix P_1 and the notation > 0 denotes positive definiteness; so P_1 is symmetric and positive definite and P_2 is symmetric.

Using the above concept, it has been demonstrated (Inman, 1983) that for dynamic systems of the form

$$\ddot{x}+A\dot{x}+Bx=0,$$

the coefficient matrices A and B are simultaneously symmetrizable if and only if they are separately symmetrizable, i.e.,

$$A=S_1S_2;\ S_1=S_1^T>0,\ S_2=S_2^T,$$

$$B=T_1T_2;\ T_1=T_1^T>0,\ T_2=T_2^T,$$

and have at least one factorization in common, i.e., $S_1 = T_1$. Using this, Inman shows that the aforementioned dynamic systems can be transformed into a symmetric form as follows:

$$\ddot{q}+\tilde{A}\,\dot{q}+\tilde{B}\,q=0,$$

where

$$\tilde{A}=\tilde{A}^T=S_1^{1/2}S_2S_1^{1/2}$$

$$\tilde{B}=\tilde{B}^T=S_1^{1/2}T_2S_1^{1/2}$$

and

$$q=S_1^{-1/2}x.$$

Here, $S_1^{1/2}$ indicates the square root of the positive definite matrix S_1, defined according to

$$S_1^{1/2}=\omega\Delta_{S_1}^{1/2}\omega^T,$$

where ω is the modal matrix of S_1 and Δ_{S_1} is a diagonal matrix of the eigenvalues of S_1.

Results

Before discussing the derivation of the results let us introduce the following two definitions:

Definition 1. The coordinate system in which the equation of motion is symmetric is referred to as *symmetric coordinates*.

Definition 2. An asymmetric system which possesses symmetric coordinates is called a *pseudo-symmetric* system.

As mentioned earlier, for the purpose of this study, asymmetric systems are divided into two categories: those without velocity-dependent forces and those with velocity-dependent forces. Now, the results for each category are presented.

Pseudo-Conservative Systems

Dynamic systems with nonconservative, follower forces, as described by the equation

$$M\ddot{x}+(K+E)x=0 \tag{3}$$

$$|M|\neq 0$$

$$K=K^T$$

$$E=-E^T$$

or

$$\ddot{x}+Bx=0 \tag{4}$$

where

$$B=M^{-1}(K+E) \tag{5}$$

are considered first. For instance Pflüger's column (Peflüger, 1964) which consists of a simply-supported elastic rod subjected to a uniformly-distributed tangential force along its length constituties a typical example of this class of systems. If the matrix B is symmetrizable, i.e.,

$$B=T_1T_2;\ T_1=T_1^T>0,\ T_2=T_2^T \tag{6}$$

then the asymmetric system shown in equation (4) can be transformed into a symmetric form via a similarity transformation; such systems are commonly referred to as pseudo-conservative systems. The significance of pseudo-conservative systems is that they can be reduced to a form where they resemble conservative systems, even though they are asymmetric in their physical coordinates.

To provide a systematic scheme for computing the matrices T_1 and T_2 in equations (6), one can write them as

$$T_1=\Phi\ \gamma\ \Phi^T \tag{7a}$$

$$\epsilon = \gamma^{-1}\Delta_B \tag{7b}$$

and

$$T_2 = \Phi^{-T}\epsilon\,\Phi^{-1} \tag{7c}$$

where γ is an arbitrary positive definite diagonal matrix (i.e., $\gamma_{ii} > 0, i = 1, 2, \ldots, n$), Φ is the modal matrix of B, and Δ_B is a diagonal matrix which contains the eigenvalues of B, i.e.,

$$\Delta_B = \Phi^{-1}B\Phi. \tag{8}$$

Since T_1 is positive definite, according to Cholesky decomposition it can be expressed as

$$T_1 = L\,L^T, \tag{9}$$

where L is a lower triangular matrix whose entries are

$$\ell_{ii} = \left[(T_1)_{ii} - \sum_{j=1}^{i=1} \ell_{ij}^2\right]^{1/2}; \quad i = 1,2, \ldots, n \tag{10a}$$

$$\ell_{ki} = \left[(T_1)_{ik} - \sum_{j=1}^{i=1} \ell_{ij}\ell_{kj}\right]/\ell_{ii};$$

$$k = i+1, i+2, \ldots, n \tag{10b}$$

Since matrix L is now known it can serve as the similarity transformation which reduces B to a symmetric form. To illustrate this more clearly, rewrite equation (4) as

$$\ddot{x} + LL^T\,T_2 x = 0, \tag{11}$$

let

$$x = Lq, \tag{12}$$

and premultiply (11) by L^{-1}; this yields

$$\ddot{q} + \tilde{B}q = 0, \tag{13}$$

where

$$\tilde{B} = \tilde{B}^T = L^T T_2 L \tag{14}$$

For computational purposes, the above process can be outlined as follows:

1. Compute the modal matrix and eigenvalues of the matrix B (i.e., form Φ and Δ_B);
2. choose the arbitrary positive definite matrix γ (the simplest choice is the identity matrix) and calculate the matrix ϵ according to equation (7b);
3. compute the matrices T_1 and T_2 according to equations (7a) and (7c); and
4. determine the matrix L according to equations (10), and find the symmetric coordinates q and the symmetric matrix $\tilde{B}$ as presented in equations (12) and (14), respectively.

Example 1. As an example, consider a system with 7 degrees-of-freedom where

$M =$ diag. [3.00 2.00 3.00 2.00 2.00 3.00 1.00]

$$K = \begin{bmatrix} 24.91 & -4.54 & -4.47 & -4.23 & -1.94 & -2.85 & -1.24 \\ -4.54 & 20.43 & -1.53 & -1.43 & -1.36 & -2.14 & -3.34 \\ -4.47 & -1.53 & 30.55 & -5.09 & -4.37 & -2.91 & -2.14 \\ -4.23 & -1.43 & -5.09 & 31.51 & -3.47 & -5.00 & -4.23 \\ -1.94 & -1.36 & -4.37 & -3.47 & 26.33 & -2.09 & -1.89 \\ -2.85 & -2.14 & -2.91 & -5.00 & -2.09 & 32.67 & -1.98 \\ -1.24 & -3.34 & -2.14 & -4.23 & -1.89 & -1.98 & 21.10 \end{bmatrix}$$

$$E = \begin{bmatrix} 0.00 & -0.51 & 0.07 & -0.89 & -0.32 & 0.52 & -0.16 \\ 0.51 & 0.00 & -0.12 & -0.20 & 0.13 & 0.55 & -1.14 \\ -0.07 & 0.12 & 0.00 & -1.28 & -0.35 & 0.53 & -1.24 \\ 0.89 & 0.20 & 1.28 & 0.00 & 0.61 & 1.87 & -1.34 \\ 0.32 & -0.13 & 0.35 & -0.61 & 0.00 & 0.72 & -0.57 \\ -0.52 & -0.55 & -0.53 & -1.87 & -0.72 & 0.00 & -0.92 \\ 0.16 & 1.14 & 1.24 & 1.34 & 0.57 & 0.92 & 0.00 \end{bmatrix}$$

which gives

$$B = \begin{bmatrix} 8.30 & -1.68 & -1.47 & -1.71 & -0.75 & -0.77 & -0.46 \\ -2.01 & 10.21 & -0.82 & -0.82 & -0.61 & -0.80 & -2.24 \\ -1.51 & -0.47 & 10.18 & -2.12 & -1.57 & -0.79 & -1.13 \\ -1.67 & -0.61 & -1.91 & 15.75 & -1.43 & -1.57 & -2.78 \\ -0.81 & -0.74 & -2.01 & -2.04 & 13.16 & -0.69 & -1.23 \\ -1.12 & -0.89 & -1.14 & -2.29 & -0.94 & 10.89 & -0.97 \\ -1.08 & -2.21 & -0.90 & -2.89 & -1.31 & -1.06 & 21.09 \end{bmatrix}$$

Solving for eigenvalues and eigenvectors of the matrix B and choosing the arbitrary matrix γ as

$\gamma =$ diag. [5.81 .48 .043 .12 .47 4.68 1.20]

yields

$$T_1 = \begin{bmatrix} 1.2192 & -0.1676 & -0.1387 & -0.1449 & -0.1524 & -0.1844 & 0.0000 \\ -0.1676 & 1.5240 & -0.1297 & 0.0000 & 0.0000 & 0.0000 & -0.1524 \\ -0.1387 & -0.1297 & 1.3716 & -0.1676 & -0.1419 & 0.0000 & -0.1524 \\ -0.1449 & 0.0000 & -0.1676 & 1.6764 & -0.1524 & -0.1387 & -0.3353 \\ -0.1524 & 0.0000 & -0.1419 & -0.1524 & 1.9812 & -0.1524 & 0.0000 \\ -0.1844 & 0.0000 & 0.0000 & -0.1387 & -0.1524 & 1.8288 & 0.0000 \\ 0.0000 & -0.1524 & -0.1524 & -0.3353 & 0.0000 & 0.0000 & 2.2860 \end{bmatrix}$$

$$T_2 = \begin{bmatrix} 6.5616 & -0.7218 & -0.6561 & -0.6240 & 0.0000 & -0.0000 & -0.6562 \\ -0.7218 & 6.5616 & 0.0000 & -0.6561 & -0.5315 & -0.6562 & -0.6240 \\ -0.6561 & 0.0000 & 7.2177 & -0.5971 & -0.6562 & -0.7940 & 0.0000 \\ -0.6240 & -0.6561 & -0.5971 & 9.1862 & -0.4659 & -0.6562 & 0.0000 \\ 0.0000 & -0.5315 & -0.6562 & -0.4659 & 6.5616 & 0.0000 & -0.7218 \\ 0.0000 & -0.6562 & -0.7940 & -0.6562 & 0.0000 & 5.9054 & -0.6562 \\ -0.6562 & -0.6240 & 0.0000 & 0.0000 & -0.7218 & -0.6562 & 9.1862 \end{bmatrix}$$

$$L = \begin{bmatrix} 1.1042 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\ -0.15179 & 1.2251 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\ -0.12561 & -0.12143 & 1.1580 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\ -0.13123 & -.016259 & -0.16067 & 1.2779 & 0.00000 & 0.00000 & 0.00000 \\ -0.13802 & -.017100 & -0.13930 & -0.15116 & 1.3855 & 0.00000 & 0.00000 \\ -0.16700 & -.020691 & -.020284 & -0.12850 & -0.14295 & 1.3278 & 0.00000 \\ 0.00000 & -0.12439 & 0.14464 & -0.28215 & -.046860 & -.036497 & 1.4719 \end{bmatrix}$$

according to the process outlined earlier. Now, the original system can be transformed into a symmetric form, as shown in equation (13). For this case the symmetric matrix $\tilde{B}$ is

$$\tilde{B} = \begin{bmatrix} 9.11 & -1.60 & -1.12 & -1.66 & -.82 & -.91 & -.62 \\ -1.60 & 10.36 & -.47 & -.38 & -.62 & -.91 & -2.77 \\ -1.12 & -.47 & 10.65 & -1.99 & -1.89 & -1.07 & -1.79 \\ -1.66 & -.38 & -1.99 & 16.27 & -1.60 & -1.79 & -3.53 \\ -.82 & -.62 & -1.89 & -1.60 & 12.82 & -1.03 & -1.97 \\ -.91 & -.91 & -1.07 & -1.79 & -1.03 & 10.49 & -1.78 \\ -.62 & -2.77 & -1.79 & -3.53 & -1.97 & -1.78 & 19.90 \end{bmatrix}$$

Pseudo-Symmetric Systems

In the presence of dissipative and gyroscopic forces the equation of motion described in equation (3) becomes

$$M\ddot{x} + (C+G)\dot{x} + (K+E)x = 0, \tag{15}$$

where

$$C = C^T,$$

$$G = -G^T,$$

and the other matrices are as defined earlier. Equation (15) can be rewritten as

$$\ddot{x} + A\dot{x} + Bx = 0, \tag{16}$$

where

$$A = M^{-1}(C+G) \tag{17a}$$

and

$$B = M^{-1}(K+E) \tag{17b}$$

are general asymmetric matrices.

For this class of systems the following theorem is used to demonstrate a systematic method for calculating the symmetric coordinates and the symmetric form of the system (16).

Theorem 1[1]. An asymmetric system possesses symmetric coordinates if and ony if the coefficient matrices A and B are simultaneously symmetrizable.

The above theorem implies that for a pseudo-symmetric system matrices A and B in equation (16) can be expressed as

$$A = S_1 S_2;\ S_1 = S_1^T > 0,\ S_2 = S_2^T \tag{18a}$$

$$B = T_1 T_2;\ T_1 = T_1^T > 0,\ T_2 = T_2^T \tag{18b}$$

$$S_1 = T_1 \tag{18c}$$

Rewriting the above symmetric factors, in a manner similar to that presented earlier for pseudo-conservative systems, yields

[1] Proof of all theorems are included in Appendix A.

$$T_1 = \Phi\ \gamma\ \Phi^T \tag{19a}$$

$$\epsilon = \gamma^{-1}\Delta_B \tag{19b}$$

$$T_2 = \Phi^{-T}\epsilon\ \Phi^{-1} \tag{19c}$$

and

$$S_1 = \theta\ \alpha\ \theta^T \tag{20a}$$

$$\beta = \alpha^{-1}\Delta_A \tag{20b}$$

$$S_2 = \theta^{-T}\beta\ \theta^{-1}, \tag{20c}$$

where γ, Φ, and Δ_B are as defined in equations (7a–c), α is an arbitrary diagonal matrix with positive entries (.i.e., $\alpha_{ii} > 0$, $i = 1, 2, \ldots, n$), θ is the modal matrix of A, and Δ_A is a diagonal matrix which contains the eigenvalues of the matrix A, i.e.,

$$\Delta_A = \theta^{-1}A\ \theta \tag{21}$$

Now, the following theorem is in order.

Theorem 2. An asymmetric system described by equation (16) is pseudo-symmetric (i.e., possesses symmetric coordinates) if and only if the arbitrary matrices α and γ are related according to

$$\alpha = \theta^{-1}\Phi\ \gamma\ \Phi^T\theta^{-T} \tag{22}$$

or

$$\alpha = \Psi\ \gamma\ \Psi^T, \tag{23}$$

where

$$\Psi = \theta^{-1}\Phi \tag{24}$$

Expanding the matrix α defined in equation (23) gives

$$\alpha_{ij} = \sum_{k=1}^{n} \gamma_{kk}\Psi_{ik}\Psi_{jk} \qquad \text{for} \qquad i \neq j \tag{25a}$$

and

$$\alpha_{ii} = \sum_{k=1}^{n} \gamma_{kk}\Psi_{ik}^2 \qquad \text{otherwise} \tag{25b}$$

$$i = 1,2, \ldots, n$$

$$j = 1,2, \ldots, n$$

Since the matrix α is chosen to be diagonal and positive definite, the right-hand-side of equations (25a) and (25b) must be zero and positive, respectively . However, the term expressed on the right-hand-side of equation (25b) is always positive, since γ_{kk} is selected to be positive; therefore, the only condition which needs to be satisfied is

$$\sum_{k=1}^{n} \gamma_{kk}\psi_{ik}\psi_{jk} = 0 \quad \begin{matrix} i = 1,2, \ldots, n \\ j = 1,2, \ldots, n \\ i \neq j \end{matrix} \tag{26}$$

The above equation can be written in a matrix form as

$$Pz = 0 \tag{27}$$

where the matrix P is a $n^2 - n/2 \times n$ matrix and the unknown vector z is a $n \times 1$ vector given according to

$$z_i = \gamma_{ii} \quad i = 1,2, \ldots, n \tag{28}$$

The elements of the matrix ψ constitutive the entries of the matrix P; they are found according to

$$\psi_{ik}\psi_{jk} = P_{\ell k}\ ; \quad k = 1,2, \ldots, n \tag{29}$$

$$i = 2, \ldots, n;\ j = 1,2, \ldots, i-1,$$

$$\ell = 1 \qquad \text{if } i = 2$$

$$\ell = j + (i-1)(i-2)/2 \qquad \text{if } i > 2$$

For a nontrivial solution of the simultaneous equations shown in equation (27), the number of independent equations must be less than the number of unknowns, which indicates the rank of P must be less than n. This means the system shown in equation (16) is not pseudo-symmetric if rank $(P) = n$. Now, equations (22) through (29) can be used to develop a systematic computational technique for calculating the matrices S_1, S_2, and T_2. This technique is outlined below:

Step 1. Compute the eigenvalues and eigenvectors of A and B, and form the matrices θ, Δ_A, Φ, and Δ_B.

Step 2. Given that all eigenvalues of A and B are real (i.e., A and B are at least separately symmetrizable), form the $n^2 - n/2 \times n$ matrix P according to equation (29).

Step 3. Transfer the matrix P to its reduced form, as defined by O'Neil (1983) and explained in Appendix B. Call the reduced matrix P_R.

Step 4. Using the matrix P_R, check for m, the rank of P; if $m = n$, then the system is not pseudo-symmetric; otherwise (i.e., $m < n$), the system can be pseudo-symmetric and a symmetric coordinate system may be found.

Step 5. Using P_R, choose the vector z such that it has all positive entries and satisfies the equation

$$P_R z = 0,$$

where, now, P_R is a $m \times n$ matrix and z is a $n \times 1$ vector. This is often a very simple task, due to the special form of the matrix P_R.

Step 6. Upon determining the vector z the symmetric factors can be calculated according to equations (19), (20).

Once the matrices S_1, S_2, and T_2 are calculated, the system shown in equation (16) can be expressed as

$$\ddot{x} + S_1S_2\dot{x} + S_1T_2x = 0 \tag{30}$$

where

$$S_1 = LL^T, \tag{31}$$

according to Cholesky decomposition (Meirovitch, 1980). Letting

$$x = Lq \tag{32}$$

and premultiplying equation (30) by L^{-1} yields the symmetric form of the system

$$\ddot{q} + L^TS_2L\dot{q} + L^TT_2Lq = 0 \tag{33}$$

where

$$L^TS_2L = (L^TS_2L)^T \tag{34a}$$

$$L^TT_2L = (L^TT_2L)^T, \tag{34b}$$

and q is the symmetric coordinates.

Example 2. Consider a system described by equation (15), where

$$M = \text{diag. } [1.00 \quad 1.00 \quad 1.00 \quad 1.00 \quad 1.00]$$

$$C = \begin{bmatrix} 34.00 & -1.50 & -11.00 & -13.50 & -7.00 \\ -1.50 & 34.00 & -6.00 & -19.50 & -4.50 \\ -11.00 & -6.00 & 48.00 & -13.50 & -15.00 \\ -13.50 & -19.50 & -13.50 & 89.00 & -17.50 \\ -7.00 & -4.50 & -15.00 & -17.50 & 102.00 \end{bmatrix}$$

$$G = \begin{bmatrix} 0.00 & 0.50 & 0.00 & -0.50 & -1.00 \\ -0.50 & 0.00 & -2.00 & 0.50 & 2.50 \\ 0.00 & 2.00 & 0.00 & 0.50 & 2.00 \\ 0.50 & -0.50 & -0.50 & 0.00 & 2.50 \\ 1.00 & -2.50 & -2.00 & -2.50 & 0.00 \end{bmatrix}$$

$$K = \begin{bmatrix} 42.00 & -1.50 & -14.00 & -14.50 & -10.00 \\ -1.50 & 38.00 & -14.00 & -14.00 & -4.00 \\ -14.00 & -14.00 & 75.00 & -15.50 & -27.50 \\ -14.50 & -14.00 & -15.50 & 88.00 & -19.50 \\ -10.00 & -4.00 & -27.50 & -19.50 & 153.00 \end{bmatrix}$$

$$E = \begin{bmatrix} 0.00 & 0.50 & -1.00 & 0.50 & -2.00 \\ -0.50 & 0.00 & -3.00 & -1.00 & 3.00 \\ 1.00 & 3.00 & 0.00 & 1.50 & 4.50 \\ -0.50 & 1.00 & -1.50 & 0.00 & 0.50 \\ 2.00 & -3.00 & -4.50 & -0.50 & 0.00 \end{bmatrix}$$

After reducing the system to the form shown in equation (16), steps 1–6, as described earlier, are followed to find the symmetric factors of the system. The result is as follows:

$$S_1 = T_1 = \begin{bmatrix} 0.54 & 0.00 & -0.13 & -0.13 & -0.13 \\ 0.00 & 0.54 & -0.13 & -0.13 & -0.00 \\ -0.13 & -0.13 & 0.67 & -0.13 & -0.13 \\ -0.13 & -0.13 & -0.13 & 0.95 & -0.13 \\ -0.13 & -0.00 & -0.13 & -0.13 & 1.35 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} 59.08 & -7.39 & -7.38 & -7.38 & 0.00 \\ -7.39 & 59.08 & 0.00 & -14.77 & -7.38 \\ -7.38 & 0.00 & 66.47 & -7.39 & -7.38 \\ -7.38 & -14.77 & -7.39 & 88.62 & -7.38 \\ 0.00 & -7.38 & -7.38 & -7.38 & 73.85 \end{bmatrix}$$

and

$$T_2 = \begin{bmatrix} 73.85 & -7.39 & -7.38 & -7.38 & 0.00 \\ -7.39 & 66.49 & -7.38 & -7.38 & -7.38 \\ -7.38 & -7.38 & 103.39 & -7.38 & -14.77 \\ -7.38 & -7.38 & -7.38 & 88.62 & -7.38 \\ 0.00 & -7.38 & -14.77 & -7.38 & 110.78 \end{bmatrix}$$

Next, calculating the transformation matrix L and the symmetric form of the system yields

$$L = \begin{bmatrix} 0.74 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.74 & 0.00 & 0.00 & 0.00 \\ -0.18 & -0.18 & 0.78 & 0.00 & 0.00 \\ -0.18 & -0.18 & -0.26 & 0.90 & 0.00 \\ -0.18 & 0.00 & -0.22 & -0.25 & 1.10 \end{bmatrix}$$

$$\tilde{B} = L^T S_2 L = \begin{bmatrix} 42.25 & 5.25 & -4.36 & -14.42 & -11.96 \\ 5.25 & 40.75 & -1.18 & -22.59 & -2.99 \\ -4.36 & -1.18 & 54.64 & -19.57 & -21.85 \\ 14.42 & -22.59 & -19.57 & 79.96 & -27.67 \\ 11.96 & -2.99 & -21.85 & -27.67 & 89.40 \end{bmatrix}$$

$$\tilde{C} = L^T T_2 L = \begin{bmatrix} 52.25 & 6.25 & -7.42 & -13.06 & -17.94 \\ 6.25 & 46.00 & -12.43 & -18.03 & -1.50 \\ -7.42 & -12.43 & 81.38 & -16.12 & -37.00 \\ -13.06 & -18.03 & -16.12 & 82.28 & -37.84 \\ -17.94 & -1.50 & -37.00 & -37.84 & 134.10 \end{bmatrix}$$

Summary

Lumped-parameter dynamic systems under the influence of general types of forces such as dissipative, gyroscopic, conservative, and follower forces were considered. The main emphasis of this study was placed on developing an effective method for reducing asymmetric dynamic equations to a symmetric form via a similarity transformation. To this end, systems that are transferable to a symmetric form were divided into two groups: those without velocity dependent forces, referred to as pseudo-conservative systems (Huseyin 1978), and those with velocity dependent forces, named pseudo-symmetric systems.

For pseudo-conservative systems, the symmetrizability conditions for an asymmetric matrix as stated by Taussky were used to find a systematic approach for calculating the symmetric coordinates (i.e., the coordinate system in which the system is symmetric).

For pseudo-symmetric systems, a similar approach was followed to allow for simultaneous symmetrizabilty of the coefficient matrices, therefore allowing the system to be reduced to a symmetric form. The scheme presented here computes the symmetric factors of the coefficient matrices and uses Cholesky decomposition to find the symmetric coordinates.

The primary advantages of the technique developed here are twofold. First, it is computationally efficient and stable. Second, it can effectively handle systems with many degrees-of-freedom, unlike the trial and error approach suggested in previous studies. Two examples involving systems with 5 and 7 degrees-of-freedom were used to illustrate the utility of the technique.

References

Ahmadian, M., and Inman, D. J., 1984, "Classical Normal Modes in Asymmetric Non-Conservative Dynamic Systems," *AIAA Journal*, Vol. 22, No. 7, pp. 1012–1015.

Ahmadian, M., and Inman, D. J., 1985, "On the Stability of General Dynamic Systems Using a Liapunov's Direct Method Approach," *Journal of Computers and Structures*, Vol. 20, No. 1-3, pp. 287–292.

Ahmadian, M., and Inman, D. J., 1986, "Some Stability Results for General Linear Lumped-Parameter Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 10–14.

Caughey, T. K., 1960, "Classical Normal Modes in Damped Linear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 27, pp. 269–271.

Caughey, T. K., and O'Kelly, M. E. J., 1965, "Classical Normal Modes in Damped Linear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 583–587.

Foss, K. A., 1958, "Co-Ordinates which Uncouples the Equation of Damped Linear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 25, pp. 361–364.

Huseyin, K., and Leipholz, H. H. E., 1973, "Divergence Instability of Multiple-Parameter Circulatory Systems," *Quarterly of Applied Math*, Vol. 31, No. 2, pp. 25–40.

Huseyin, K., 1978, *Vibrations and Stability of Multiparameter Systems*, Noordhoff International Publishing, The Netherlands.

Inman, D. J., and Andry, A. N., Jr., 1980, "Some Results on the Nature of Eigenvalues of Discrete Damped Linear Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 927–930.

Inman, D. J., 1983, "Dynamics of Asymmetric Non-Conservative Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 199–203.

Meirovitch, L., 1980, *Computational Methods in Structural Dynamics*, Sijthoff and Noordhoff, The Netherlands.

Mingori, D. L., 1970, "A Stability Theorem for Mechanical Systems with Constraint Damping," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 253–258.

Moran, T. J., 1970, "A Simple Alternative to the Routh-Hurwitz Criterion for Symmetric Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 1168–1170.

O'Neil, P. V., 1983, *Advanced Engineering Mathematics*, Wadsworth Publishing Co., 1st Ed., pp. 470–477.

Peflüger, A., 1964, *Stabilitatsprobleme der Elastostatik*, Springer-Verlag.

Lord Rayleigh, 1945, *The Theory of Sound*, Vol. 1, Dover, New York.

Taussky, O., 1968, "Positive Definite Matrices and Their Role in the Study of the Characteristic Roots of General Matrices," *Advances in Mathemtatics*, Vol. 2, pp. 175–186

Taussky, O., and Zassenhaus, H., 1969, "On the Similarity Transformation Between a Matrix and its Transpose," *Pacific Journal of Mathematics*, Vol. 9, pp. 839–896.

Walker, J. A., and Schmitendorf, W. E., 1973, "A Simple Test for Asymptotic Stability in Partially Dissipative Symmetric Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, No. 4, pp. 1120–1121.

APPENDIX A

Theorem 1. The proof of this theorem is as follows. If the coefficient matrices are simultaneously symmetrizable, then according to Inman (1983) equation (16) can be rewritten as

$$\ddot{x} + S_1 S_2 \dot{x} + S_1 T_2 x = 0 \qquad (A1)$$

$$S_1 = S_1^T > 0$$

$$S_2 = S_2^T$$

$$T_2 = T_2^T$$

or

$$\ddot{x} + LL^T S_2 \dot{x} + LL^T T_2 x = 0, \qquad (A2)$$

based on Cholesky decomposition. Letting $x = Lq$ and premultiplying equation (A2) by L^{-1} yields

$$\ddot{q} + (L^T S_2 L)\dot{q} + (L^T T_2 L) q = 0 \qquad (A3)$$

which is symmetric. Therefore, the system shown in equation (16) possesses symmetric coordinates. Conversely, if the system (16) has symmetric coordinates z, then it can be written as

$$\ddot{z}+L_1^{-1}AL_1\dot{z}+L_1^{-1}BL_1z=0,$$

where the matrices $(L_1^{-1}AL_1)$ and $(L_1^{-1}BL_1)$ are symmetric, i.e.,

$$L_1^{-1}AL_1=(L_1^{-1}AL_1)^T \tag{A4}$$

and

$$(L_1^{-1}BL_1)=(L_1^{-1}BL_1)^T \tag{A5}$$

Equation (A4) can be rewritten as

$$A^TL_1^{-T}L_1^{-1}=(A^TL_1^{-T}L_1^{-1})^T \tag{A6}$$

which implies that the matrix $A^TL_1^{-T}L_1^{-1}$ is symmetric. Furthermore, equation (A4) implies that

$$A=(L_1L_1^T)(A^TL_1^{-T}L_1^{-1}) \tag{A7}$$

This indicates that the matrix A is symmetrizable, since it can be written as the product of two symmetric matrices one of which is positive definite. Similarly it is possible to use (A5) to show that B is symmetrizable and can be written as

$$B=(L_1L_1^T)(B^TL_1^{-T}L_1^{-1}). \tag{A8}$$

Therefore, according to (A7) and (A8), the matrices A and B are simultaneously symmetrizable.

Theorem 2. The proof of this theorem is based on Theorem 1. If the system is pseudo-symmetric, then

$$A=S_1S_2,\ S_1=S_1^T>0,\ S_2=S_2^T \tag{A9}$$

$$B=T_1T_2,\ T_1=T_1^T>0,\ T_2=T_2^T \tag{A10}$$

and

$$S_1=T_1. \tag{A11}$$

based on equations (19*a*) and (20*a*), equation (A11) can be rewritten as

$$\theta\,\alpha\,\theta^T=\Phi\,\gamma\,\Phi^T \tag{A12}$$

or

$$\alpha=\theta^{-1}\Phi\,\gamma\,\Phi^T\theta^{-T} \tag{A13}$$

Conversely, if the arbitrary matrices α and γ are related to each other according to (A13), then one gets

$$\theta\,\alpha\,\theta^T=\Phi\,\gamma\,\Phi^T \tag{A14}$$

or

$$S_1=T_1, \tag{A15}$$

using equations (19*a*) and (20*a*). Therefore, the system is pseudo-symmetric if and only if the arbitrary matrices α and γ are related according to the equation (A13).

APPENDIX B

A reduced matrix is a matrix which satisfies the following conditions:

I. The first nonzero entry in each row is 1; this is called the leading entry.
II. If row r has its leading entry in column c, then all other entries of column c are zero.
III. Each row having all zero entries, if there is such a row, lies below any row having a nonzero entry.
IV. If the first nonzero entry in row r lies in column c_1 and the first nonzero entry of row r_2 is in column c_2, and if $r_1 < r_2$, then $c_1 < c_2$.

A complete discussion of reduced matrices can be found in Section 10.5 of O'Neil (1983) where it is proved that every matrix has exactly one reduced matrix row-equivalent to it.

ERRATA

Errata on "Asymmetric Wave Propagation in an Elastic Half-Space by a Method of Potentials," by R. Y. S. Pak and published in the March 1987 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 121–126.

On page 124, equation (30) should read:

$$X_m = \tilde{P}_m^{m-1}(\xi) - i\tilde{Q}_m^{m-1}(\xi);$$

Entrainment of Self-Sustained Flow Oscillations: Phaselocking or Asynchronous Quenching?

T. Staubli
Hydraulic Research and Development,
Sulzer Escher Wyss Ltd.,
Zürich, Switzerland
Mem. ASME

Asynchronous quenching and phaselocking are two different mechanisms leading to the onset of synchronization of flow instabilities with externally excited oscillations. Experimental evidence for asynchronous quenching as well as for phaselocking is given from response measurements of representative pressures, velocities, or of forces for the following types of flow - forcing interactions: an oscillating circular cylinder in crossflow; interaction of an unstable, planar jet with an oscillating leading edge; a forced mixing layer between parallel streams; and a thermally forced cavity shear-layer.

Introduction

If self-sustained flow oscillations are subjected to external forcing, they usually synchronize with the excitation over a certain range of excitation frequencies. Such synchronization effects, commonly known as "lock-in" or "locking-on," are observed for many types of flow instabilities, e.g., for the Kármán vortex street forming behind cylinders or blunt trailing edges, for free shear-layer instabilities, for oscillations of impinging flows such as flow past cavities, and for jet - edge (edgetone) oscillations. The external forcing also encompasses many different types such as mechanical forcing, acoustic forcing by a loudspeaker, or thermal forcing.

In this context of flow systems containing fluctuating vorticity, our purpose will be to demonstrate that external excitation of such systems exhibits the typical characteristics of quasiperiodicities and phase-locking in forced dynamical systems. The contents of this paper focus on the onset of synchronization of self-sustained oscillations, arising from the flow instability, with the external forcing. From experimental evidence, it will be shown, that there are at least two means of attaining this synchronization: *phaselocking* and *asynchronous quenching*. This terminology is chosen in accordance with that of Dewan (1972), who discusses the entrainment of van der Pol oscillations.

For *phaselocking*, the external forcing causes a systematic phase retardation or advancement of the self-sustained oscillations in such a manner that the frequency of the self-sustained oscillation becomes identical with the excitation frequency. Retardation occurs for excitation frequencies below the natural frequency and advancement occurs for higher excitation frequencies. As synchronization is approached, this phase shifting leads to coalescence of the self-sustained frequency and the excitation frequency. For phaselocking, therefore, both components coexist undistinguishably at the same frequency, contribute to the oscillations, and lead to resonance. Phaselocking usually is observed for small excitation levels.

In the case of *quenching* the self-sustained oscillation of the flow is attenuated and disappears when the excitation frequency approaches that of the self-sustained oscillation. We call it asynchronous quenching if this suppression of the self-sustained oscillations occurs at a frequency different from the excitation frequency. Quenching usually requires, in contrast to phaselocking, higher excitation levels.

Synchronization of flow oscillations and body motion can also be observed for coupled systems involving a self-excited, fluid-dynamic oscillator and a passive mechanical oscillator, e.g., an elastically mounted cylinder in crossflow. Although the frequency of the mechanical system may, for such cases, be affected considerably by the flow effects, the flow oscillations themselves will be entrained by the body motion and synchronize with it over certain ranges of the parameters chosen in the experiment, e.g., flow velocity. Typical for such coupled systems is that there are ranges where two stable states of oscillation can be found at the same frequency, which explains experimentally observed hysteresis effects and jumps in amplitude and phase (e.g., Feng, 1968). This multi-amplitude response arises from the coupling of the nonlinear fluid-dynamic system with a linear mechanical oscillator (Staubli, 1983b). Unfortunately, this coupling makes it difficult to discuss the response of the fluid-dynamic part alone since the mechanical system typically has a narrow-band resonant response due to weak damping, thereby veiling the details in the onset of synchronization of the self-sustained flow oscillations.

This paper discusses the onset of synchronization for the following types of interaction of external forcing with self-sustained flow oscillations: an oscillating circular cylinder in crossflow (Staubli, 1983a); an impinging jet interacting with an oscillating edge (Staubli and Rockwell, 1987); a forced mix-

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 23, 1986.

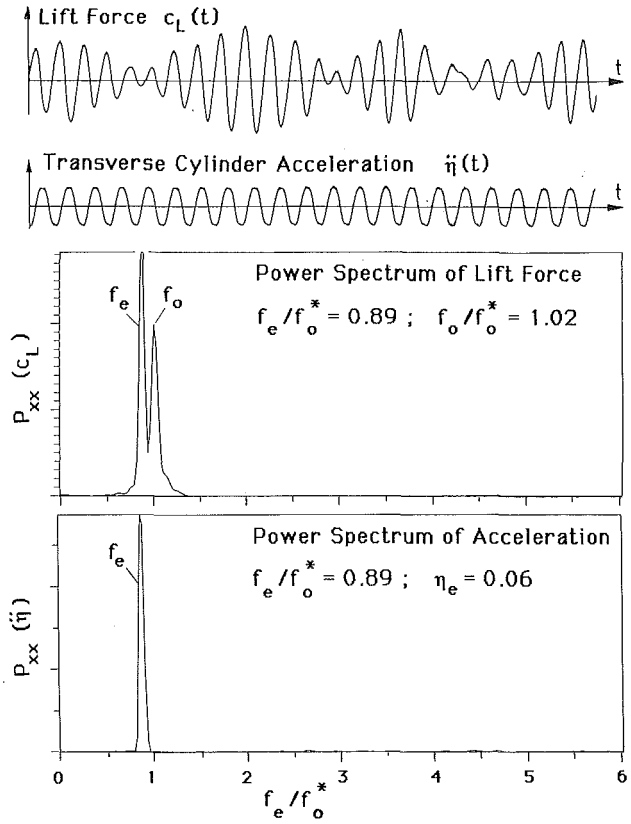


Fig. 1 Time records and power spectra of lift force and acceleration for an oscillating circular cylinder in crossflow in case of not synchronized vortex shedding (Staubli, 1983a)

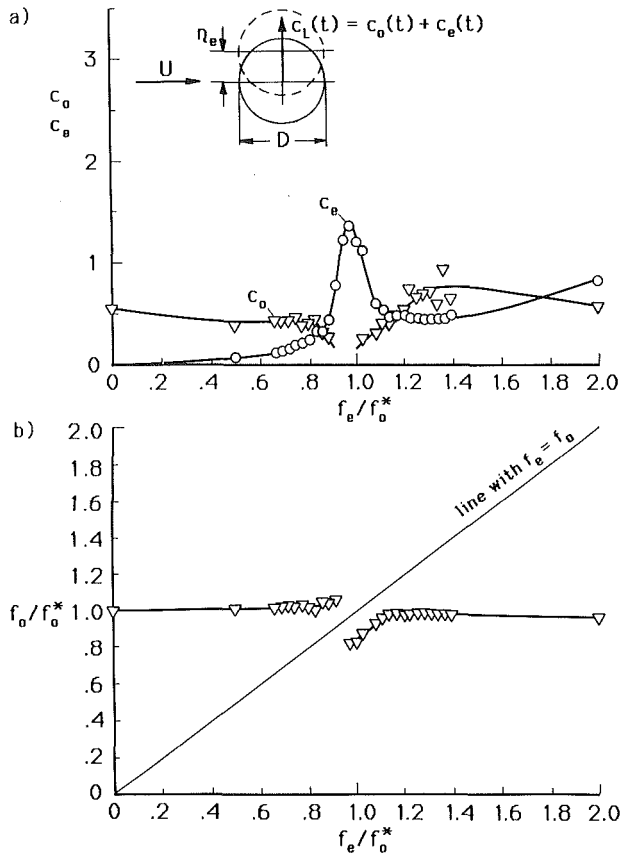


Fig. 2 Response characteristics of the lift force of an oscillating circular cylinder ($\eta_e = 0.11$) in crossflow showing: (a) the lift coefficients c_o (self-excited component) and c_e (externally excited component); (b) the normalized frequency f_o/f_o^* of the self-sustained lift component c_o as a function of the normalized excitation frequency f_e/f_o^* (Staubli, 1983a)

ing layer between parallel streams (Oster and Wygnanski, 1983); and a thermally forced, oscillating cavity shear-layer (Gharib, 1983).

The main parameters in all these studies are the frequency and amplitude of the forced excitation. Because of the nonlinearity of the fluid-dynamic system (self-sustained flow oscillations) only harmonic forcing can be considered. Typically, the response of the fluid-dynamic system is determined by power spectra of global, or representative local, flow properties such as force, pressure, or velocity.

With the frequency and amplitude of forcing as parameters, there are two ways of attaining synchronization. Either the amplitude may be increased at a constant excitation frequency, or the amplitude may be maintained constant while the frequency is varied. The latter way is preferable to show the resonance effects that accompany synchronization. If we wish to ascertain the type of entrainment, both methods are equivalent and complement each other.

Knowledge of the type of entrainment of the self-sustained flow oscillations by the external excitation may provide insight into the physcial events leading to synchronization and, further, may lead to an appropriate model in the form of a nonlinear differential equation modelling the global properties of the fluid-dynamic system or flow oscillator.

Loading on an Oscillating Circular Cylinder in Crossflow

Fluid forces acting on an externally-driven cylinder were measured in a water tank using a towing technique. The experimental system and the associated measurements are described by Staubli (1983a). Selected measurements of fluctuating lift will be presented here to demonstrate the transition to synchronization in a case where the self-excited oscillation in the flow is of the Kármán vortex street type.

The power spectra of the measured lift forces showed that only two discrete frequency components have to be taken into account for an accurate description of the response of the fluid forces in case of forced sinusoidal oscillation of the cylinder. This discrete two-frequency response was observed under the experimental restriction of small amplitudes of cylinder oscillations ($<0.8D$). For such small amplitudes the fluid-dynamic response, in terms of lift forces, is typically deterministic. Concerning the dependency on Reynolds number, we find this characteristical two-frequency response for the whole range of technically important Reynolds numbers with the exception of the transitional range between $Re = 2 \cdot 10^5$ and 10^6; for supercritical Reynolds numbers the response to sinusoidal excitation was investigated by Szechenyi and Loiseau (1975).

The correctness of the superposition of two cosine components for approximation of the lift force is confirmed by the measured time records which show modulations with the periodicity $1/|f_e - f_o|$ (f_e = excitation frequency; f_o = frequency of self-sustained vortex shedding). An example of time records and power spectral density distributions is given in Fig. 1. While the lift signal shows a typical beating wave form in the time domain, the measured acceleration signal of the cylinder oscillations shows to a good approximation a sinusoidal time record.

Accordingly, the forcing function which is given by the sinusoidal displacement signal $\eta(t)$ (η_e = Amplitude/Diameter) and the resulting lift forces $c_L(t)$ can be written as follows:

displacement

$$\eta(t) = \eta_e \cos(2\pi f_e t) \tag{1}$$

lift coefficient

$$c_L(t) = c_o \cos(2\pi f_o t) + c_e \cos(2\pi f_e t + \phi). \tag{2}$$

In the absence of cylinder oscillations the self-sustained vortex shedding in the wake gives rise to lift fluctuations of the form

$$c_L(t) = c_o \cos(2\pi f_o^* t), \tag{3}$$

which is a special case of equation (2), where $c_e = 0$ and $f_o \rightarrow$

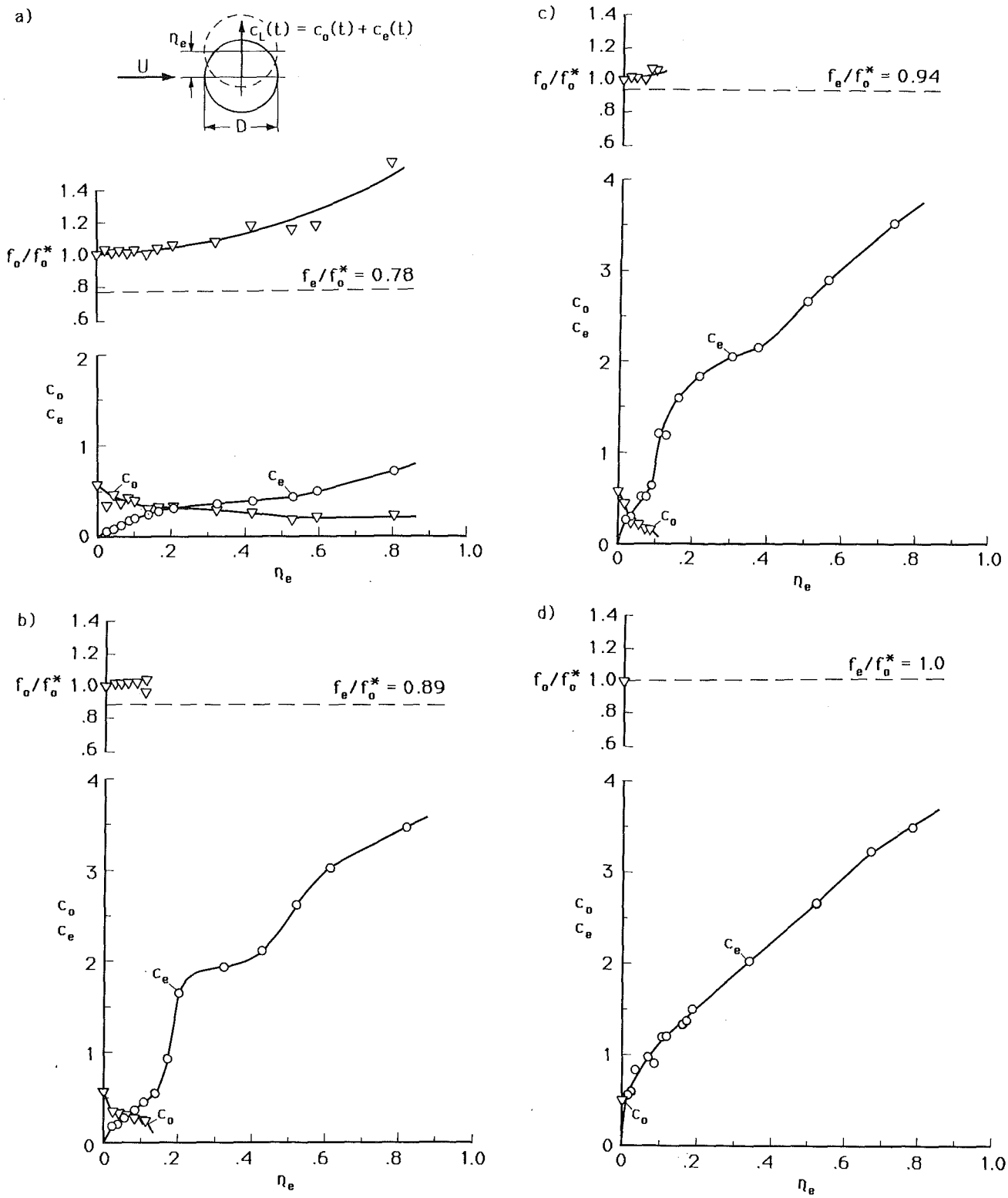


Fig. 3 The lift coefficients c_o, c_e and the frequency f_o of the self-sustained lift component c_o for varying displacement amplitudes η_e and excitation at a frequency ratio of: (a) $f_e/f_o^* = 0.78$; (b) $f_e/f_o^* = 0.89$; (c) $f_e/f_o^* = 0.94$; (d) $f_e/f_o^* = 1.0$ (Staubli, 1983a)

f_o^*. The frequency f_o^* denotes the frequency of the "natural" Kármán vortex shedding in absence of cylinder oscillations. Especially near synchronization, the frequencies f_o and f_o^* are not identical. The measured Strouhal number was $S = f_o^* D/U = 0.180$ for the Reynolds number of $\mathrm{Re} = 6 \cdot 10^4$.

The spectral components c_e at the excitation frequency f_e and c_o at the self-excited frequency f_o were determined from spectral density distributions after taking window effects (Hanning windows) into account and by assuming minimal modulations of f_e and f_o. It should be noted here, that the self-sustained component was not completely free of modulations; however, we can conclude from the shape of the peaks in the spectra that these modulations were minor.

Figure 2(*a*) displays the lift coefficients c_o (amplitude of self-sustained component) and c_e (amplitude of forced component) as a function of the normalized excitation frequency f_e/f_o^* for a constant amplitude of displacement $\eta_e = 0.11$. There is a synchronization range near $f_e/f_o^* = 1$ where the cylinder oscillation entrains the flow oscillation. The coefficient c_e shows a distinct resonance behavior within that range; the maximum lies at f_e/f_o^* slightly less than one. For high excitation frequencies, the lift coefficient c_e increases due to added mass effects and is proportional to f_e^2. For frequencies above synchronization, the self-sustained flow oscillations recover, and the lift component c_o shows amplitudes comparable to those of the nonoscillating cylinder, i.e., $f_e/f_o^* = 0$.

The onset of synchronization clearly indicates an attenuation of the self-sustained lift coefficient c_o on both sides of the synchronization range. Extrapolation of the measured data points indicates a complete quenching of the self-sustained flow oscillation near synchronization. Here, the resolution of the smallest values of the coefficients c_o which still can be determined accurately in the measurements is limited by the signal-to-noise ratio. Although we note, e.g., in Fig. 1, a noise level due to stochastic or other signal contributions, which is negligible in comparison to the deterministic peaks in the spectra, this noise makes the determination of the self-sustained component c_o more difficult when the latter is small relative to the externally excited component c_e. As can be seen in Fig. 2(*a*), the coefficient c_o becomes small relative to c_e for the onset of synchronization; it is in this range that noise contributes significantly to the peak of c_o in the power spectrum and adds a positive systematic error to the true values of c_o. Thus, without any noise effects, a stronger drop of c_o could be expected in Fig. 2(*a*) for the onset of synchronization. In an absolute sense, there is no experimental evidence of an increased noise level for the onset of synchronization.

Even if the noise significantly influences the amplitude of the lift coefficient c_o of the self-sustained oscillation, the associated frequency f_o still can be determined accurately. This frequency f_o normalized with the frequency f_o^* of the "natural" vortex shedding, is plotted in Fig. 2(*b*) as a function of the normalized excitation frequency f_e/f_o^*. Since the measured data points do not approach the line $f_e = f_o$, the self-excited and the excitation frequencies do not coalesce. Especially for $f_e/f_o^* > 1$, the measured frequency f_o shifts away from its original value f_o^* towards lower frequencies. This result confirms measurements of Stansby (1976) for an oscillating cylinder having the same relative displacement amplitude but a different Reynolds number. These findings concerning the amplitude of the lift component c_o and the shift of the frequency f_o near synchronization represent an example of entrainment of self-sustained flow oscillations by asynchronous quenching.

Measurements for constant excitation frequency and increasing displacement amplitude are displayed in the Figs. 3(*a*) through 3(*d*). In none of these cases do the frequencies f_o and f_e coalesce. Further, the self-excited component of the lift force, the coefficient c_o, is always attenuated by increasing the amplitude of excitation; in fact, Figs. 3(*b*) and 3(*c*) show a complete suppression of c_o. This suppression of the self-excited component for larger amplitudes of displacement demonstrates the well-known effect of broadening of the synchronization range at larger excitation amplitudes.

Somewhat different is the case shown Figs. 3(*d*), where the excitation frequency is identical to the natural Kármán shedding frequency, that is $f_e/f_o^* = 1$. For this case, and also for $f_e/f_o^* \approx 1$, we should closely examine the limit of very low excitation amplitude. By definition, for $\eta_e \to 0$ the coefficient c_e must approach zero, while c_o shows the value of the nonoscillating cylinder. With the available equipment it was not possible to obtain adequate resolution for displacement amplitudes smaller than $\eta_e \approx 0.02$. For such a displacement the measured lift coefficient at excitation frequency f_e is already larger than the coefficient c_o of the nonoscillating cylinder.

However, since c_e grows from a zero value, there must be a range of excitation at very small amplitudes where we will find a superposition of two independent components, the self-excited and the externally forced ones. They both contribute in case of $f_e/f_o^* = 1$ to the lift force at the same frequency. Of course, it is not possible to separate contributions at the same frequency by measurement, but we can conclude from this transition to zero excitation amplitude that here, in contrast to larger amplitudes, quenching of the self-sustained flow oscillations cannot be the mechanism leading to synchronization. We rather expect synchronization by phaselocking which requires the coexistence of the two contributions at the same frequency as a necessary condition. In general, we can conclude from the above measurements that for the oscillating circular cylinder, oscillating at amplitudes of practical interest, synchronization is achieved by asynchronous quenching.

Similar experimental observations of attenuation in the onset of synchronization are reported by Graham and Maull (1971), who investigated the interaction of vortex shedding behind a plate with an oscillating trailing edge. They also observed an attenuation of the self-excited components when the excitation frequency approached the frequency of the natural vortex shedding. Their measurements of velocity fluctuations show in the spectra, somehow overemphasized by the logarithmic scale, modulations of the self-excited frequency components. This effect is due to overall modulations of the self-sustained vortex shedding, but also due to the fact that the single point measurement of velocity represents only a local property in a turbulent wake.

Impinging Jet Interacting With an Oscillating Edge

The case of an unstable, planar jet interacting with an oscillating leading edge has been investigated by Staubli and Rockwell (1987). Some of the measurements will be presented here to provide insight into the type of synchronization occurring when the controlled frequency f_e of the edge oscillations approaches the frequency f_o^* of the self-sustained jet oscillation.

In these experiments, pressure fluctuations $p(t)$ were measured at a representative location on the leading edge. This localized pressure loading was used to characterize – as lift forces were in case of the oscillating cylinder – the response of the jet oscillations to mechanically forced oscillations of the edge. The geometry and the Reynolds number were chosen small enough so that the response of the jet-edge interaction could be described to a good approximation, with only two discrete frequency components (larger impingement length and higher Reynolds numbers exhibit multiple frequency components of the self-sustained jet oscillation; Lucas and Rockwell, 1984). Thus, the pressure fluctuation $p(t)$ can be written in analogy to equation (2) as a superposition of the following two cosine components with the frequencies f_o and f_e.

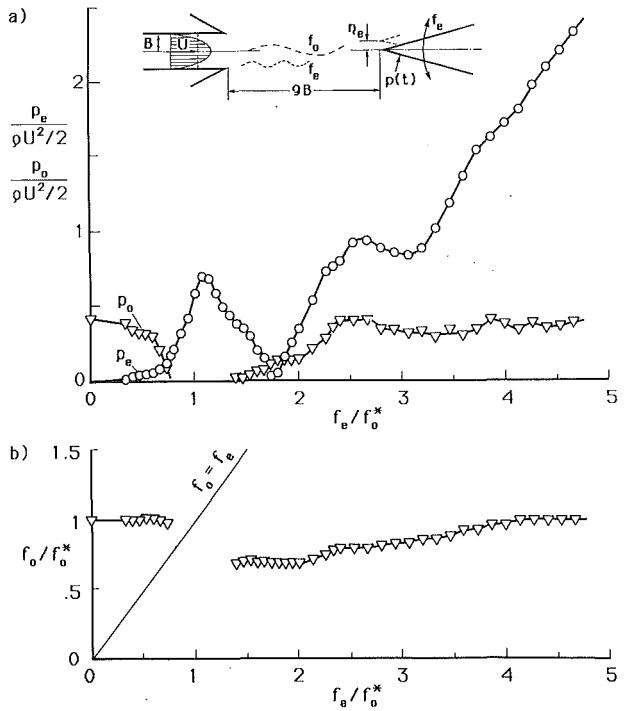


Fig. 4 Response characteristics of the pressure loading due to jet-edge interaction (tip displacement η_e = 0.16) showing: (*a*) self-excited pressure component p_o and externally excited pressure component p_e; (*b*) normalized frequency f_o/f_o^* of the self-sustained jet oscillation as a function of the normalized excitation frequency f_e/f_o^* (Staubli and Rockwell, 1987)

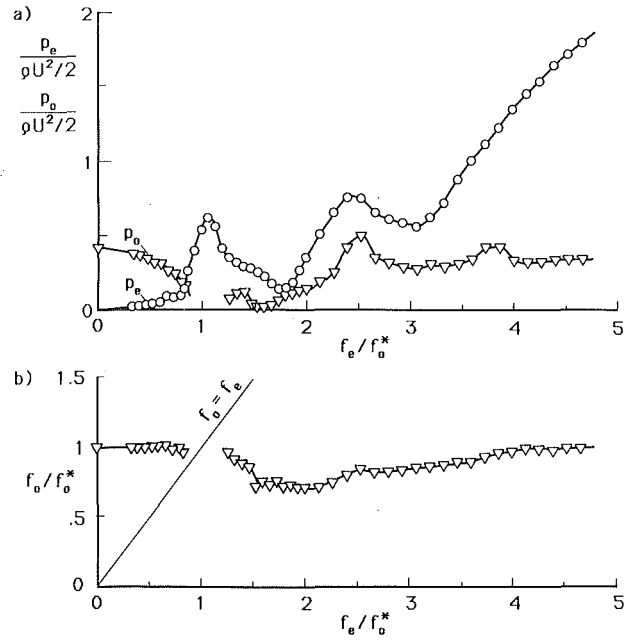


Fig. 5 Response characteristics of the pressure loading due to jet-edge interaction (tip displacement η_e = 0.11) showing: (*a*) self-excited pressure component p_o and externally excited pressure component p_e; (*b*) the normalized frequency f_o/f_o^* of the self-sustained jet oscillation (Staubli and Rockwell, 1987)

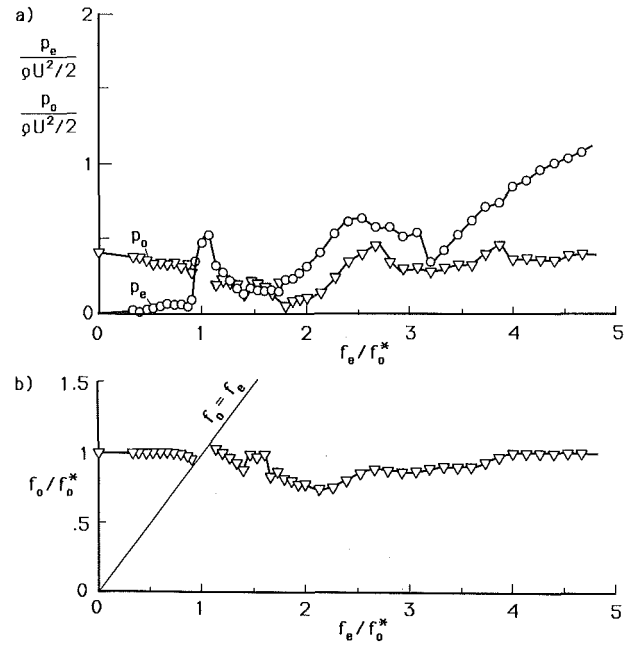


Fig. 6 Response characteristics of the pressure loading due to jet-edge interaction (tip displacement η_e = 0.07) showing: (*a*) self-excited pressure component p_o and externally excited pressure component p_e; (*b*) the normalized frequency f_o/f_o^* of the self-sustained jet oscillation (Staubli and Rockwell, 1987)

$$p(t) = p_o \cos(2\pi f_o t) + p_e \cos(2\pi f_e + \phi) \quad (4)$$

The phase angle ϕ is here the phase between the edge displacement and the pressure fluctuations. In absence of edge oscillations, equation (4) degenerates to

$$p(t) = p_o \cos(2\pi f_o^* t) \quad (5)$$

The frequency f_o^* is the frequency of the "natural," self-sustained jet oscillation which corresponds to the most amplified jet instability. The forced excitation is described in terms of the tip displacement of the leading edge (nondimensional with nozzle width $2B$)

$$\eta(t) = \eta_e \cos(2\pi f_e t) \quad (6)$$

The response characteristics of the self-excited component p_o and of the externally excited component p_e for the constant amplitude of displacement $\eta_e = 0.16$ are displayed in Fig. 4(a). There is a striking similarity with the measurements of the loading on an oscillating circular cylinder in crossflow. The forced pressure component p_e indicates a pronounced resonance for $f_e/f_o^* \approx 1$, comparable to that of the lift coefficient c_o in Fig. 2(a). For higher excitation frequencies the pressure fluctuation p_e increases due to added mass effects and is proportional to f_e^2. For the high frequencies of excitation the self-sustained jet oscillation completely recover; the pressure amplitudes p_o at $f_e/f_o^* > 2.5$ reach values comparable to those of the stationary edge with $f_e/f_o^* = 0$. For the onset of synchronization the self-sustained oscillations are attenuated down to a level where they disappear within an overall noise level. This can be observed on both sides of the synchronization range.

Figure 4(b) displays the frequency of the self-sustained jet oscillation f_o as a function of the excitation frequency f_e, both normalized with the frequency f_o^* of the "natural" jet oscillations. For the low frequency ratios f_e/f_o^* the external excitation hardly effects the self-excited frequency f_o. For excitation frequencies above synchronization, the ratio f_o/f_o^* becomes considerably smaller than one and recovers to a value of one for increasing excitation frequency when the interaction of the two components decreases.

Figures 5 and 6 show two cases with smaller amplitudes of displacement, that is $\eta_e = 0.11$ and $\eta_e = 0.07$. The response characteristics of the forced component p_e are qualitatively the same for all three investigated amplitudes of displacement.

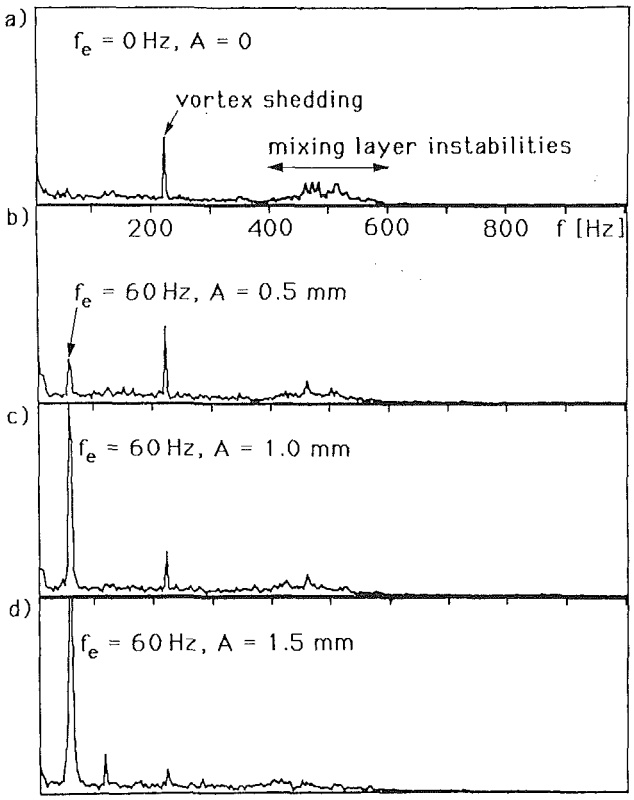


Fig. 7 Power spectra of the streamwise velocity fluctuations u' for a forced mixing layer between parallel streams (Oster and Wygnanski, 1983). ($r = U_1/U_2 = 0.6$; $x = 100$ mm; $y = 0$; A = amplitude of flap oscillation.)

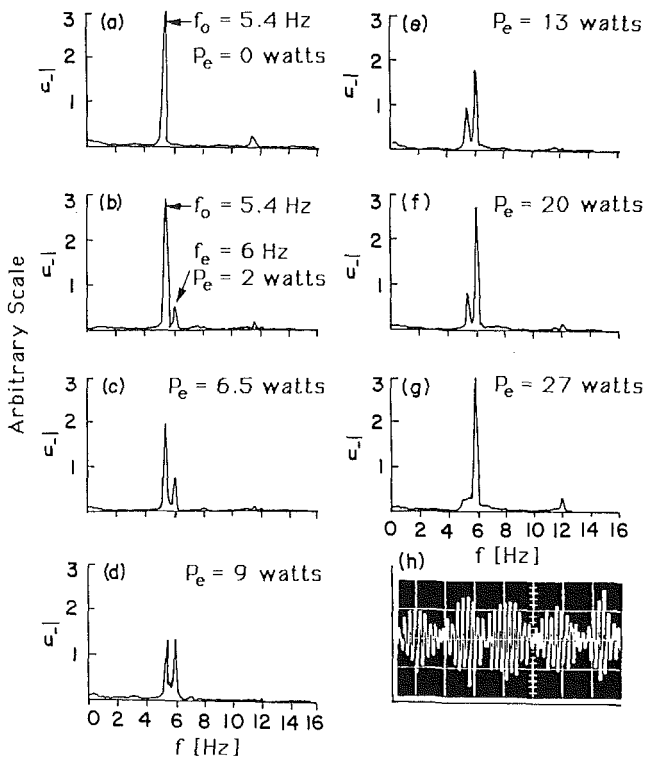


Fig. 8 Velocity fluctuation spectra, presenting interaction of the natural and forced frequencies in a cavity shear layer for different power levels (P_e); (*h*) scope trace of case *(d)* (Gharib, 1983)

At resonance, there is a minor reduction of the maximum amplitudes; and, for high frequencies f_e, the smaller added mass effect proportionally reduces the pressure fluctuation p_e. The self-excited pressure p_o shows a stronger effect of the reduction in displacement amplitude, especially near synchronization. As expected for smaller excitation levels, the attenuation of p_o is reduced and the actual synchronization range becomes smaller. At least for $\eta_e = 0.07$, shown in Fig. 6(*a*), there is no more indication of a complete suppression of p_o. Thus, this experiment shows that there is an amplitude limit below which the external excitation cannot suppress the self-sustained flow oscillations, a necessary condition for quenching. Further indication that we have phaselocking in this case of very low displacement amplitude comes from Fig. 6(*b*) showing the frequency f_o of the self-excited component versus the excitation frequency f_e. For this graph, and also already in Fig. 5(*b*), we observe that the measured points tend to approach the line $f_e = f_o$. This coalescence of the self-excited frequency with the externally excited frequency is the second necessary condition for phaselocking.

Thus, we can conclude that the response of the flow oscillation of the impinging jet on forced edge oscillations is, in general, comparable to the response of the Kármán vortex shedding behind an oscillating cylinder in terms of force and pressure measurements; for the large displacement amplitudes the measurements clearly indicate an entrainment of the flow oscillations of the asynchronous quenching type. From the measurements at very small amplitudes in the jet-edge case, as well as from deduction of the limit of excitation amplitudes tending to zero, we can conclude that synchronization is caused there by phaselocking.

The Forced Mixing Layer Between Parallel Streams

Oster and Wygnanski (1983) present velocity measurements in a turbulent mixing layer. Controlled oscillations of a flap (amplitude A) were applied at the initiation of mixing between two parallel streams. The frequency of forcing was at least one order of magnitude lower than the initial instability frequency. In contrast to the above shown examples, the self-excited frequencies are in this case not restricted to one discrete frequency but instead represent frequency components in a broad band between 400 and 600 Hz, as can be seen from Fig. 7(*a*). The peak in the power spectrum at $f = 230$ Hz is a subharmonic of the shedding frequency in the immediate neighborhood of the trailing edge. Oster and Wygnanski report the following observation: "With increasing amplitude of surging (Figs. 7(*c*), 7(*d*)) one may note a marked increase in the energy content at the forcing frequency ($f_e = 60$ Hz) and a relative reduction at high frequencies, so that the integral of all spectral components of u remains approximately constant. Further investigation is required to determine how the fluctuating energy shifts to the forcing frequency,. . . ."

Thus, these measurements indicate another case of attenuation of the self-excited components, while the level of the external excitation is increased. We can, therefore, classify this onset of synchronization with the excitation frequency as asynchronous quenching. Of course, this classification does not explain how the energy is transferred from one frequency component to another; but one expects, from similarity to the foregoing cases of attenuation of the self-sustained oscillation in the onset of synchronization, that there might be similar physical principles leading to this attenuation in all cases of asynchronous quenching.

In contrast to the cases of the oscillating cylinder or of the jet-edge interaction, which showed only one discrete frequency component resulting from the self-excited flow oscillations, the self-excited contribution consists in this mixing layer example of a broad-band of effects in the frequency spectrum. Under these conditions, the response of the fluid cannot be approximated anymore with two-frequency components, which

was done in equations (2) and (4). For a description of the response we have to consider in this case the integral over a frequency range associated with the self-excited flow effects. Nevertheless, this example shows that the classification of entrainment by quenching is not necessarily restricted to a group of flow situations where only discrete frequency components are observed in the flow oscillations.

Effect of Forcing on a Naturally Oscillating Cavity Flow

Gharib (1983) describes oscillations arising from flow past a rectangular cavity and a means of controlling them. His experiments have been performed in a water tunnel using an axisymmetric cavity model. The forced oscillations were introduced by a sinusoidally heated thin-film strip, which excited the Tollmien-Schlichting waves in the boundary layer upstream of the cavity.

Figure 8 shows measurements taken at a constant frequency f_e and increasing power P_e supplied to the heated strip. Gharib mentions: "Spectra reveal that as the amplitude of the forcing increases, the amplitude of the natural oscillation decreases and eventually disappears." Furthermore, "strong modulation of the signal indicates that both frequencies are present simultaneously." From these comments and from Fig. 8 we can conclude that this example describes still another case of well-defined asynchronous quenching.

Conclusion

In the foregoing, we have addressed four different examples of self-sustained, fluid-dynamic oscillations subjected to external forcing. In spite of differences between these various flow configurations, there are qualitative similarities in the response of the respective flow instabilities to external forcing, including the onset of synchronization.

Of the two means of attaining synchronization, asynchronous quenching is detected more often. In case of jet-oscillating edge interaction, measurements indicate phaselocking for very small amplitudes of excitation. This also can be concluded for all other cases from the fact that there is a lower limit in amplitude of external forcing for which self-sustained oscillations can be suppressed.

The similarity of the measured responses for these flow oscillations suggests that, in all cases, similar nonlinear differential equations might be formulated to describe the measured fluid-dynamic response. These differential equations for global flow properties might be second or higher-order differential equations with small nonlinearities for damping (causing self-excitation) and also with a small nonlinear spring.

That a nonlinear equation for the flow instability, e.g., the van der Pol equation, coupled with a linear oscillator for the mechanical system can successfully describe the vortex induced oscillation of an elastically mounted cylinder in crossflow was first demonstrated by Hartlen and Currie (1970). The great advantage of formulating differential equations describing the fluid-dynamic response to external excitation would be that once these equations are known, any mechanical oscillation can be calculated from coupling these equations with differential equations describing the mechanical system. Of course, such amplitudes of oscillation of mechanical systems can also be obtained directly from measured, fluid-dynamic data, a means which involves extended computations, as shown by Staubli (1983b).

Acknowledgment

I gratefully thank Professor Donald Rockwell of Lehigh University for valuable discussions and helpful comments. Furthermore, I acknowledge the Swiss National Science Foundation who funded a two-year stay of myself at Lehigh University.

References

Dewan, E. M., 1972, "Harmonic Entrainment of van der Pol Oscillations: Phaselocking and Asynchronous Quenching," *IEEE Transactions on Automatic Control*, Vol. AC-17, No. 5, pp. 655-663.

Feng, C. C., 1968, "The Measurement of Vortex-Induced Effects in Flow Past Stationary and Oscillating Circular and D-Section Cylinders," MASc thesis, University of British Columbia.

Gharib, M., 1983, "The Effect of Flow Oscillations on Cavity Drag, and a Technique for Their Control," PhD thesis, Graduate Aeronautical Laboratories California Institute of Technology.

Graham, J. M. R., and Maull, D. J., 1971, "The Effects of an Oscillating Flap and an Acoustic Resonance on Vortex Shedding," *Journal of Sound and Vibration*, Vol. 18, Part 3, pp. 371-380.

Hartlen, R. T., and Currie, I. G., 1970, "Lift Oscillator Model of Vortex-Induced Vibration," *Proceedings of the ASCE, Journal of the Engineering Mechanics Division*, EM5, pp. 577-591.

Lucas, M., and Rockwell, D., 1984, "Self-Excited Jet: Upstream Modulation and Multiple Frequencies," *Journal of Fluid Mechanics*, Vol. 147, pp. 333-352.

Oster, D., and Wygnanski, I., 1983, "The Forced Mixing Layer between Parallel Streams," *Journal of Fluid Mechanics*, Vol. 123, pp. 91-130.

Stansby, P. K., 1976, "The Locking-On of Vortex Shedding to the Cross-Stream Vibration of Circular Cylinders in Uniform and Shear Flows," *Journal of Fluid Mechanics*, Vol. 74, Part 4, pp. 641-665.

Staubli, T., 1983a, "Untersuchung der Oszillierenden Kräft am Querangeströmten, Schwingenden Kreiszylinder," Disseration, Swiss Federal Institute of Technology, Zürich, Diss. ETH 7322.

Staubli, T., 1983b, "Calculation of the Vibration of an Elastically Mounted Cylinder Using Experimental Data From Forced Oscillation," ASME *Journal of Fluids Engineering*, Vol. 105, pp. 225-229.

Staubli, T., and Rockwell, D., 1987, "Interaction of an Unstable, Planar Jet with an Oscillating Leading Edge," *Journal of Fluid Mechanics*, Vol. 176, pp. 135-167.

Szechenyi, E., and Loiseau, H., 1975, "Portance Instationnaires sur un Cylindre Vibrant dans un écoulement supercritique," *Recherche Aerospatiale, No. 1, pp. 45-57.*

A New Method for Predicting the Critical Taylor Number in Rotating Cylindrical Flows

J. O. Cruickshank
Avco Research Laboratory, Inc.,
Everett, MA 02149

A method for determining the boundaries of dynamic stability of a fluid system, as distinct from the prediction of the subsequent motion, is presented. The method is based on well-known approaches to the problem of instability in elastic systems. The extension of these methods to fluid systems, specifically, to the stability of flow between concentric cylinders, confirms that it may be possible in some cases to determine the boundaries of stability of fluid systems without recourse to an Orr-Sommerfeld type treatment. The results also suggest that the concept of apparent (virtual) viscosity may have implications for fluid stability outside the current realm of turbulence modelling. Finally, it is also shown that flow instability may be preceded by the onset of a critical stress condition in analogy with elastic systems.

1 Introduction

The study of the stability of fluid systems has been a part of the ongoing research into the behavior of fluids going back as far as the famous dye experiments of Osborne Reynolds (1883) and perhaps earlier. In general, there are two problems of interest to those who work in the field: the prediction of the conditions under which a given fluid system would become unstable, and the description of the subsequent motion.

Currently, most techniques for the study of unstable flows are based on models that arise from the perturbation of time-dependent Navier-Stokes equations. Typical of the resulting equations are the Orr-Sommerfeld equations, the derivation of which may be found in any good text on Fluid Mechanics (Schlichting, 1968; White, 1974).

The stability analysis of elastic systems provide a clue as to a possible alternate approach of this problem. Simple models of elastic systems, based on linear assumptions and displacements, have been used extensively for the prediction of the boundaries of stability of such systems. The results, such as the critical buckling load for columns, have found wide use in engineering, and the fact that these simple models stop short of predicting the subsequent (very often nonlinear) motion has for most situations been irrelevant. It is the proposition here that some of these techniques from solid mechanics are applicable to some problems in fluid mechanics. In this paper, we use the problem of instability of rotating flow between two concentric cylinders to show that flow instability can be analyzed using concepts analogous to those used for studying instability in elastic systems and thus we open up the possibility of an integration of these two currently disparate fields.

The general methods for the determination of the boundaries of stability for elastic systems are fully described in Bolotin (1964a). In the following modelling, the philosophical approach is based almost entirely on that work, due regard being paid when necessary to the differences that naturally exist between solids and fluids.

It is perhaps worth recalling that the equations that generally govern the behavior of solids and fluids, the equations of continuum mechanics, are the same until appropriate constitutive relations for stress are invoked. Thus, clearly, as long as the stress terms are retained and proper attention is paid to the other subtle differences that exist between fluids and solids, it is not surprising that some of the equations from solid mechanics can be used without major modifications.

The notion that some aspects of fluid stability are analogous to the behavior of solids has its origins in the concept of fluid buckling. Experimental and theoretical work in this area by Cruickshank (1980), Blake and Bejan (1984), Bejan (1981), Suleiman and Munson (1981), and Cruickshank and Munson (1981) suggests that the investigation of this link can shed valuable light on flow stability problems ranging from those at low Reynolds number to high Reynolds number turbulent situations. Previous work in this area have generally dealt with unique, isolated instances of the fluid buckling phenomenon. In this paper, the Euler buckling analogy inherent in the concept of fluid buckling is extended to one of the more classical problems of fluid mechanics.

2 The Governing Equations

A. The Differential Equations of Dynamic Stability. Consider a three-dimensional element of a continuum performing a regular three-dimensional motion with transport velocities (u, v, w) on which a three-dimensional oscillation is then imposed. We restrict ourselves to small, though finite

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March 17, 1986.

displacements resulting from the oscillations. During the oscillatory displacement, the following assumptions as stated by Bolotin (1964b) are assumed to hold, and they refer to the effects of the oscillation on the medium:

(*a*) Lengths of linear elements, the areas and volumes change; however, the influence of these changes on the stability is usually negligibly small and can be neglected for small deformation relative to unity.

(*b*) Angles of rotation and the general nondimensional oscillatory displacements in comparison to unity may not be neglected. Thus, changes in cross section and length are not considered; however, the additional forces which arise from rotation of the cross section due to the oscillation are taken into account.

So far, it is clear that even though the original Bolotin work was developed for solids, its applicability to any medium that satisfies the concept of the continuum is preserved by not invoking the constitutive relation for stress.

The finite displacements, using the terminology of Bolotin, may be described in a system of rectangular Cartesian or curvilinear coordinates x_1, x_2, x_3 referred to the undeformed medium (in the case of a fluid system, this would be the coordinate system applicable to the base, unperturbed flow). If the corresponding displacements of points in the medium are u_1, u_2, u_3 then the new coordinates will be

$$\xi_1 = x_1 + u_1(x_1, x_2, x_3) \tag{1}$$

$$\xi_2 = x_2 + u_2(x_1, x_2, x_3) \tag{2}$$

$$\xi_3 = x_3 + u_3(x_1, x_2, x_3) \tag{3}$$

Such a formulation would be a Lagrangian as distinct from an Eulerian formulation. In the Eulerian system, the coordinates of points of the deformed medium (ξ_1, ξ_2, ξ_3) are taken as independent variables, hence according to Bolotin, the equations that govern the displacement (or oscillation) are given by:

$$\sum_{k=1}^{3} \frac{\partial \sigma_{ik}}{\partial \xi_k} + X_i = \rho \frac{D^2 u_i}{Dt^2} \quad i = 1,2,3 \tag{4}$$

In equation (4), σ_{ik} are the components of the stress, X_i are the components of the body force per unit volume in the deformed state, ρ is the density of the deformed medium, and D/Dt is the substantial derivative. This equation in its simpler manifestations is part of the standard engineering literature. For example, the vibrating string equation is obtained by setting $u_1 = 0$, $u_3 = 0$, and $x_1 = x$; then for this problem $x_2 = 0$, the component of stress σ_{12} is $T\,\partial u_2/\partial x$, and neglecting the body force, we have

$$T \frac{\partial^2 u_2}{\partial x^2} = \rho \frac{\partial^2 u_2}{\partial t^2}$$

where T is the tension in the string and ρ is its mass per unit length. If the string is also moving with velocity v in the x direction, the resulting equation would be the so-called moving threadline equation (Swope and Ames, 1963).

B. The Two-Dimensional Problem. We now consider the application of equation (4) to an element with base (unperturbed) motion in a horizontal (x, z) plane, and subjected to oscillatory displacements perpendicular to this plane. This simplification is carried out with the application to the Taylor problem specifically in mind. Other stability problems may require slight variations from this approach. The oscillations are thus given by

$$u_1 = 0, u_3 = 0, u_2 \neq 0$$

Thus,

$$\xi_1 = x_1 = x$$

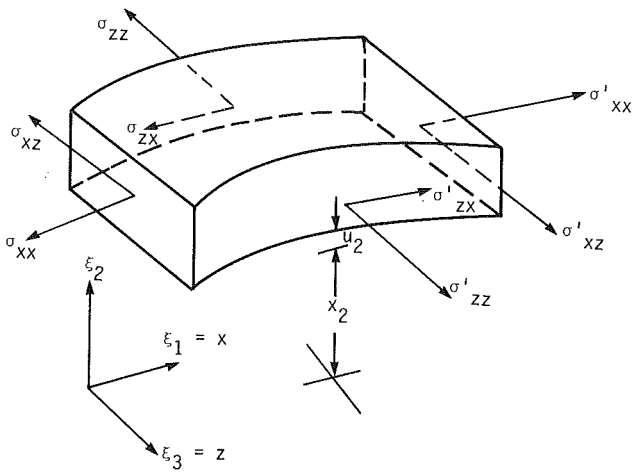


Fig. 1 An element displaced u_2 above plane of motion (x, z)

and

$$\xi_3 = x_3 = z$$

Furthermore, if the stress components in the ξ_2 direction are independent of that direction, then, upon neglect of body forces, equation (4) becomes:

$$\frac{\partial \sigma_{11}}{\partial \xi_1} + \frac{\partial \sigma_{13}}{\partial \xi_3} = 0 \tag{4a}$$

$$\frac{\partial \sigma_{21}}{\partial \xi_1} + \frac{\partial \sigma_{23}}{\partial \xi_3} = \rho \frac{D^2 u_2}{Dt^2} \tag{4b}$$

$$\frac{\partial \sigma_{31}}{\partial \xi_1} + \frac{\partial \sigma_{33}}{\partial \xi_3} = 0 \tag{4c}$$

Equations (4*a*) and (4*c*) simply imply that for small displacements, there is no appreciable change in the stress components in the corresponding directions. This is equivalent to the equation

$$T\cos\Theta - T\cos(\Theta + d\Theta) = 0$$

where T is the tension and Θ is its angle with the horizontal, obtained for a horizontal resolution of forces during the derivation of the vibrating string equation. For small $d\Theta$ we obtain the expression $T\cos\Theta = T\cos\Theta$ which is true but irrelevant to the further analysis of the problem. Similarly, we drop equations (4*a*) and (4*c*) from further consideration. Hence, if we choose a medium that is infinitely deep in the ξ_2 direction such that variations with respect to ξ_2 may be neglected, then the behavior of an element at any (ξ_1, ξ_3) plane will be the same and the component of stress in the ξ_2 direction will be independent of ξ_2. The component of the transport velocity in that direction, v, may also be assumed to be zero.
This results in the following equation:

$$\frac{\partial \sigma_{21}}{\partial \xi_1} + \frac{\partial \sigma_{23}}{\partial \xi_3} = \rho \frac{D^2 u_2}{Dt^2} \tag{4d}$$

Note that σ_{21} and σ_{23} have a meaning that is slightly different from the obvious usage, as demonstrated earlier and in equations (5*a*) and (5*b*) below.

Thus, we consider the motion of an element with steady (u, w) transport velocity components in a horizontal (x, z) plane with the superimposed oscillatory displacements u_2 occurring perpendicular to this plane. A much more general analysis would include the variation of the stress components in the ξ_2 direction and all three transport velocities would be retained. For the specified problem, however, this is not necessary. Figure 1 shows a typical such element after the deformation described.

It can be shown that the components of stress in the u_2 direction after the deformation, are

$$\sigma_{21} = \sigma_{xx} \frac{\partial u_2}{\partial x} + \sigma_{zx} \frac{\partial u_2}{\partial z} \tag{5a}$$

where σ_{2l} is the sum of all stresses on the x face resolved in the u_2 direction (Fig. 1). $\partial u_2/\partial x$ and $\partial u_2/\partial z$ account for angular change in the orientation of the faces of the element due to the rotation arising from the displacement u_2.

$$\sigma_{23} = \sigma_{zz} \frac{\partial u_2}{\partial z} + \sigma_{xz} \frac{\partial u_2}{\partial x} \tag{5b}$$

Using

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + w \frac{\partial}{\partial z}$$

the corresponding equation obtainable from equation (4*d*) is

$$\frac{\partial}{\partial x}\left(\sigma_{xx} \frac{\partial u_2}{\partial x}\right) + \frac{\partial}{\partial z}\left(\sigma_{zz} \frac{\partial u_2}{\partial z}\right) + \frac{\partial}{\partial z}\left(\sigma_{xz} \frac{\partial u_2}{\partial x}\right)$$

$$+ \frac{\partial}{\partial x}\left(\sigma_{zx} \frac{\partial u_2}{\partial z}\right)$$

$$= \rho\left[\frac{\partial^2 u_2}{\partial t^2} + 2\left(u \frac{\partial^2 u_2}{\partial x \partial t} + w \frac{\partial^2 u_2}{\partial z \partial t}\right) + 2uw \frac{\partial^2 u_2}{\partial x \partial z}\right.$$

$$+ u^2 \frac{\partial^2 u_2}{\partial x^2} + w^2 \frac{\partial^2 u_2}{\partial z^2} + \frac{\partial u_2}{\partial x}\left(u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial z}\right)$$

$$\left. + \frac{\partial u_2}{\partial z}\left(u \frac{\partial w}{\partial x} + w \frac{\partial w}{\partial z}\right)\right] \tag{6}$$

From the momentum equations of continuum mechanics

$$\rho\left(u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial z}\right) = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \tag{7}$$

and

$$\rho\left(u \frac{\partial w}{\partial x} + w \frac{\partial w}{\partial z}\right) = \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \tag{8}$$

Substituting equations (7) and (8) into equation (6), we obtain

$$\left(\frac{\sigma_{xx}}{\rho} - u^2\right)\frac{\partial^2 u_2}{\partial x^2} + 2\left(\frac{\sigma_{xz}}{\rho} - uw\right)\frac{\partial^2 u_2}{\partial x \partial z} + \left(\frac{\sigma_{zz}}{\rho} - w^2\right)\frac{\partial^2 u_2}{\partial z^2}$$

$$- 2u \frac{\partial^2 u_2}{\partial x \partial t} - 2w \frac{\partial^2 u_2}{\partial z \partial t} - \frac{\partial^2 u_2}{\partial t^2} = 0 \tag{9}$$

Equation (9) would be applicable to both solids and fluids as we still have not invoked the constitutive relations for stress.

3 Application to the Taylor Stability Problem

Equation (9) will now be transformed to cylindrical coordinates for application to the problem of flow between two concentric cylinders. The outer cylinder is stationary and the inner one is rotating. It is assumed that the steady base flow in the (x, z) plane is subjected to an oscillatory disturbance such that a particle is displaced a distance u_2 above its base (undeformed) plane, consistent with the earlier theory. We will show that the simple form of disturbance assumed in this model is sufficient to determine the critical conditions under which this flow will become unstable. More complicated forms are, therefore, unnecessary.

When end effects are neglected, the rotating cylinder problem fits the "infinitely deep" assumption of the formulation leading to equation (9). Clearly, the component of stress in the direction of the axis of the cylinders is independent of that direction when end effects are neglected, and the disturbance itself is imposed such that this independence is preserved.

Continuing this development, we propose that we can, for the fluid motion considered here, replace the steady stress components of the previous sections arising from the steady nature of the transport velocities with a combination of this original steady-state stress and a very small time dependent component. It is the assumption that so long as this time-dependent component is very small, the previous equations will remain essentially correct. Thus, for example, we will replace the stress σ_{xx}, with a new one $\sigma_{xx}{}^{\circ} + \sigma_{xx}(t)$ where $\sigma_{xx}{}^{\circ}$ is the original steady-state stress and $\sigma_{xx}(t)$ is a small time-dependent stress the form of which will be determined for the particular problem at hand from the appropriate constitutive relations. For a fluid, this additional stress component can only be induced by a time-dependent velocity or pressure. Since we retain the steady-state pressure and velocity, we need to look for an additional plausible time-dependent velocity component that would provide the time-dependence in the stress terms, but which at the same time is small enough to not affect in a major way the basic steady-state flow velocities and pressures derived from the Navier-Stokes equations.

A. The Equations for Taylor Stability Analysis. It can be shown that for the case under consideration, equation (9) reduces to:

$$\frac{\sigma_{rr}}{\rho} \frac{\partial^2 u_2}{\partial r^2} + \left[\frac{\sigma_{\Theta\Theta}}{\rho} - V_\Theta^2\right]\frac{1}{r} \frac{\partial u_2}{\partial r} = \frac{\partial^2 u_2}{\partial t^2} \tag{10}$$

where

$$V_\Theta(r) = \frac{\omega r_1^2}{r_2^2 - r_1^2}\left[\frac{r_2^2}{r} - r\right],$$

ω is the angular velocity of the inner cylinder, and r_1, r_2 are the radii of the inner and outer cylinders, respectively. V_Θ is the steady state flow velocity.

We allow the original steady stresses to become time-dependent in the manner indicated earlier, and these are substituted into equation (10). The steady-state stresses are obtained from the steady-state solution. We are now left to determine the small time-dependent stress components.

The constitutive relation for the normal stresses in equation (10) are

$$\sigma_{rr} = -p + 2\mu \frac{\partial V_r}{\partial r} \tag{11}$$

and

$$\sigma_{\Theta\Theta} = -p + 2\mu\left(\frac{1}{r} \frac{\partial V_\Theta}{\partial \Theta} + \frac{V_r}{r}\right) \tag{12}$$

For the classical steady-state problem, $V_\Theta \neq f(\Theta)$ and $V_r = 0$. We retain, in this model, the Θ-independence but allow the existence of a small V_r only as it affects these stresses. The continuity equation is then used to obtain a time-dependent value of V_r and this becomes the basis for the small time-dependent components in the stresses alluded to earlier.

From the continuity equation

$$\frac{\partial V_r}{\partial r} + \frac{V_r}{r} + \frac{1}{r} \frac{\partial V_\Theta}{\partial \Theta} = 0 \tag{13}$$

Hence,

$$\frac{\partial}{\partial r}(r\, V_r) = 0$$

The continuity equation allows us to obtain an expression for V_r that is at least consistent with continuity.

Let $r\, V_r = f(t)$ where $f(t)$ is an arbitrary function of time, then the continuity equation is satisfied. From Fourier analysis, we may write

$$f(t) = \epsilon \sin\Omega t + \ldots\ldots$$

where Ω is some frequency and ϵ is sufficiently small,

$$\text{then } V_r = \epsilon / r \sin\Omega t \tag{14}$$

and the normal stresses become

$$\sigma_{rr} = -p - 2\mu\, \epsilon / r^2 \sin\Omega t \tag{15}$$

$$\sigma_{\Theta\Theta} = -p + 2\mu\, \epsilon / r^2 \sin\Omega t \tag{16}$$

As stated earlier, V_r is allowed to affect only the stress terms of this model. V_Θ and p are retained from their steady-state form unchanged.

Thus equation (10) becomes

$$\left(-\frac{p}{\rho} - 2\nu \frac{\epsilon}{r^2} \sin\Omega t\right) \frac{\partial^2 u_2}{\partial r^2} + \left(-\frac{p}{\rho} + 2\nu \frac{\epsilon}{r^2} \sin\Omega t - V_\Theta{}^2\right) \frac{1}{r} \frac{\partial u_2}{\partial r} = \frac{\partial^2 u_2}{\partial t^2} \tag{17}$$

This equation, along with appropriate boundary and initial conditions will be solved numerically to determine the value of ϵ at which u_2 becomes unstable.

To summarize then, we intend to study the stability of the Taylor flow problem by studying the physical displacement of the fluid particles in a direction perpendicular to the (r, Θ) plane of the flow. We do this by interpreting the stress terms in the Bolotin equations to include a small time-dependent component obtained by solving the continuity equation for V_r and substituting this into the constitutive relations. V_r is set small enough so that the values of $p(r)$ and $V_\Theta(r)$ are unchanged, by its presence, from their steady-state values. Its only significance is in modifying the stress terms that are used for the prediction of the oscillatory displacements. In short, V_r is small enough not to change V_Θ or p, but the theory here is that it causes the stress terms to acquire a small time dependent component as shown in equations (15) and (16).

B. Apparent Viscosity and ϵ. It is easily shown that the parameter ϵ in equations (15) and (16) has dimensions of kinematic viscosity. Clearly, since it arises specifically as a consequence of the disturbance, it can only be an apparent viscosity. This is consistent with the evolution of the concept of apparent viscosity in its more traditional application to turbulence modelling. The current model (yet to be shown to be capable of predicting the bounds of stability for the rotating cylinder problem) indicates that the apparent viscosity may be a universal parameter in fluid stability studies, arising in unstable flows other than turbulent ones. It can also be shown, quite simply, that Prandtl's formula for apparent viscosity can be derived from the model used here.

Consider the apparent stress term in equations (15) and (16) derived by solving the continuity equation for V_r. Clearly, $\mu\epsilon/r^2$ has dimensions of stress.

$$\text{Thus, } \mu\epsilon/r^2 \propto [\text{stress}]$$

We may generalize for some other flow situations that this apparent stress may be proportional to the shear stress in the fluid, or

$$\frac{\mu\epsilon}{\ell^2} = K\mu \left| \frac{du}{dy} \right| \tag{18}$$

where K is a proportionality constant and $\mu\,|du/dy|$ is a measure of the stress in the fluid, ℓ is a significant length in the flow, replacing r, the radial distance.

Then if $\mu_T = \rho\epsilon$ be an apparent dynamic viscosity, equation (18) becomes

$$\mu_T = K\rho\ell^2 \left| \frac{du}{dy} \right| \tag{19}$$

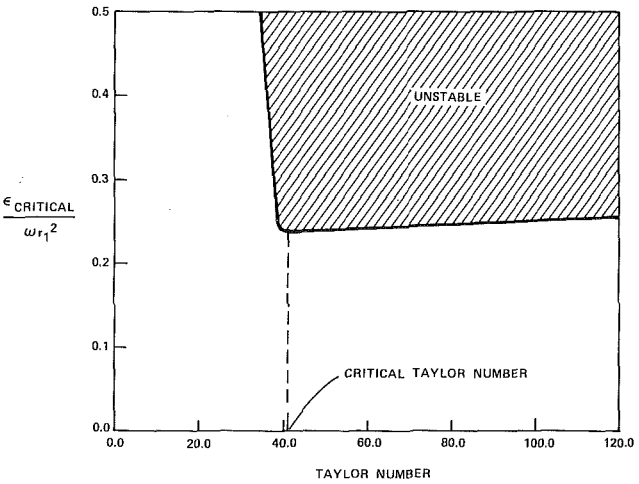


Fig. 2 The regime of stability for flow between concentric rotating cylinders (outer cylinder stationary)

or for $K = 1$

$$\mu_T = \rho\ell^2 \left| \frac{du}{dy} \right| \tag{20}$$

This is the formulation of the apparent dynamic viscosity derived by Prandtl (1967). Clearly, we are able to obtain his formula using a completely different approach, arising, as it were, from consideration of a problem in laminar instability.

C. Solution of the Stability Equation. We investigated the behavior of equation (17) by applying the following boundary and initial conditions to the fluid behavior.

$$u_2(r_1, t) = 0 \tag{21}$$

$$u_2(r_2, t) = 0 \tag{22}$$

$$u_2(r, o) = Y_o \sin \pi \frac{(r_2 - r)}{(r_2 - r_1)} \tag{23}$$

and

$$\frac{\partial u_2}{\partial t}(r, o) = 0 \tag{24}$$

In equation (23), Y_o is an arbitrarily small amplitude of the initial displacement. The choice of initial condition is arbitrary, the form shown being a reasonably simple one, consistent with those used in elasticity.

The pressure distribution is given by

$$p(r) = \frac{\rho\omega^2 r_1^4}{(r_2 - r_1)^2} \left[-\frac{r_2^4}{2r^2} + 2r_2^2 \ln \frac{r_2}{r} + \frac{r^2}{2} \right] \tag{25}$$

where we have used the boundary condition $p(r_2) = 0$ in solving for the pressure distribution from the Navier-Stokes equations.

Using equations (21)–(25), an implicit second order accurate time marching formulation was used to determine the fluid displacement at different time levels. The solution was deemed to be unstable when $u_2(r, t) \to \infty$ as $t \to \infty$ at some values of r.

4 Results and Discussions

Figure 2 shows a plot of $\epsilon_{\text{critical}}/\omega r_1^2$ versus Ta, the Taylor number defined as

$$T_a = \frac{\omega r_1{}^2}{\nu} \sqrt{\frac{(r_2 - r_1)}{r_1}}.$$

$\epsilon_{\text{critical}}$ is the numerically determined value of ϵ which ensures neutral stability. Thus, if $\epsilon > \epsilon_{\text{critical}}$ the displacements become

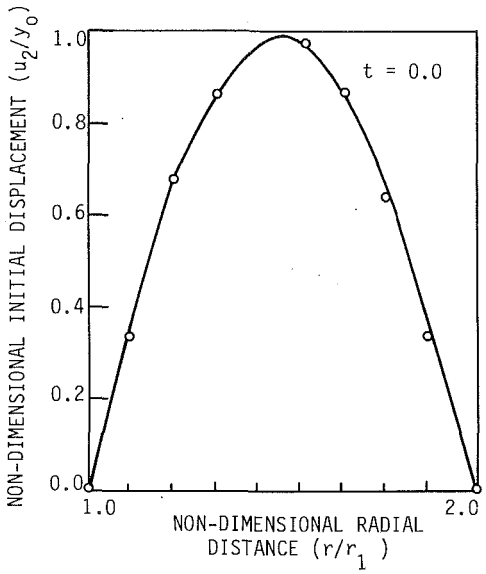


Fig. 3 Initial (assumed) displacement of fluid from plane

unstable but are stable for $\epsilon < \epsilon_{\text{critical}}$. It must be pointed out that as the flow approaches the point of neutral stability, it takes longer and longer for the flow to display either its stable or unstable character. Since, with the present numerical formulation, we cannot for practical reasons continue the solutions until time becomes infinite, we defined the flow as stable (or unstable) depending on whether instability has been attained after a reasonably long interval of time. There could be very slight variations in the values of $\epsilon_{\text{critical}}$ if the solutions were allowed to evolve to much higher time levels, but such variations are not expected to affect significantly the overall trends in Fig. 2.

From Fig. 2, it is clear that unstable flows can be induced at any Taylor number, however, it is also clear that as the level of the disturbance is increased, the first point of instability corresponds to the experimentally observed one. Points to the left of the critical point can be reached only by introducing large disturbances into the flow while at the same time keeping the Taylor number relatively low. Solutions of the displacement equation assuming an initial disturbance as shown in Fig. 3 show that under such conditions, the instability is *not* a centrifugal instability because it is found mostly near the inner wall. Figure 4 shows this type of instability arising from the initial disturbance shown in Fig. 3. Centrifugal type instability (that is, one where the instability is mostly near the outer wall) is observed only at Taylor numbers greater than the critical value indicated.

Figures 5 and 6 show typical unstable and stable solutions of equation (17). The unstable displacements of Fig. 5 were obtained to the right of the critical Taylor number. From Fig. 5, the fluid elements appear to "buckle" against the outer cylinder for the unstable modes given by $\epsilon/\omega r_1^2 = 0.458$. It was shown earlier that because of the external disturbances we may write:

$$\sigma_{rr} = -p - 2\mu \frac{\epsilon}{r^2} \sin\Omega t \tag{26}$$

$$\sigma_{\theta\theta} = -p + 2\mu \frac{\epsilon}{r^2} \sin\Omega t \tag{27}$$

where ϵ is an apparent kinematic viscosity. From these equations, a series of critical values of the apparent viscosity were obtained (Fig. 2).

A. The Concept of the Induced Critical Stress. We now

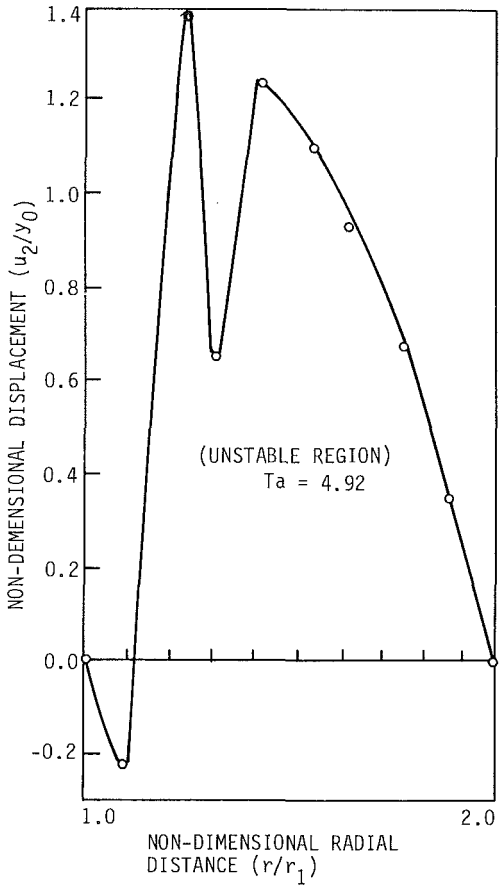


Fig. 4 Fluid displacement (unstable) for Ta < 41.3

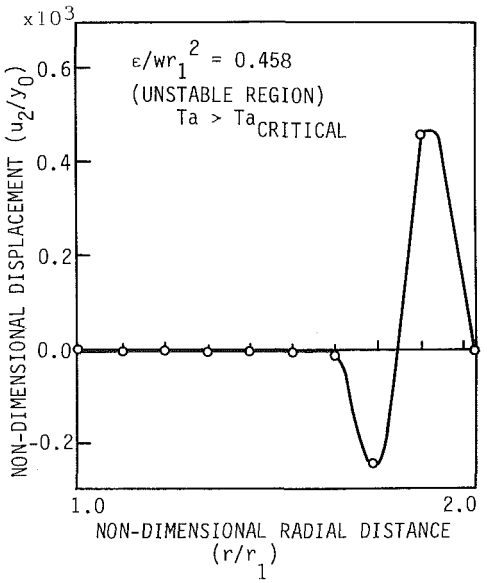


Fig. 5 Fluid displacement for unstable ($\epsilon/\omega r_1^2 = 0.458$) configuration Ta > 41.3

also propose the existence of corresponding induced critical stresses relating to these values of ϵ. This critical induced stress, σ_{critical}, would be the maximum value of the induced stress that exists at the onset of instability and is a natural follow-on from equations (15) and (16). Theoretical values of the induced critical stress will also be presented later for the case of small V_Θ. We define the critical induced stress arising at the onset of the instability as follows:

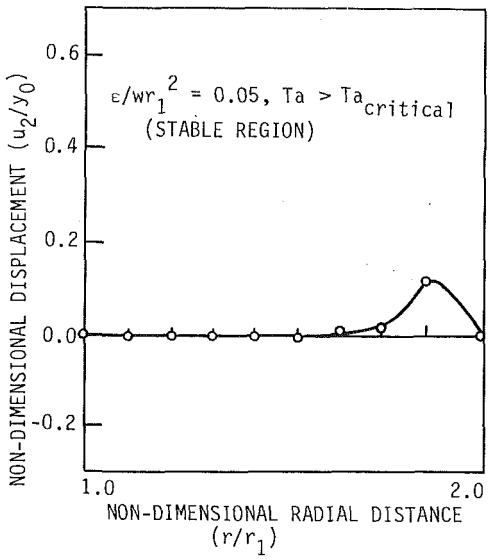


Fig. 6 A stable displacement configuration ($\epsilon/\omega r_1^2 = 0.05$): Ta > 41.3

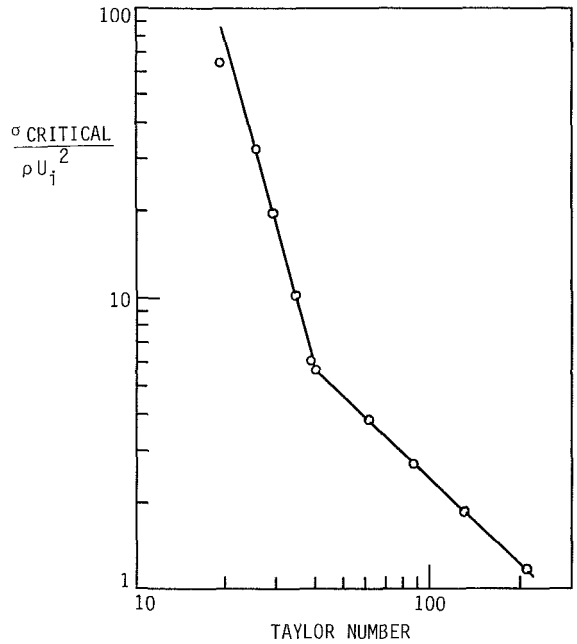


Fig. 7 The induced critical stress at onset of instability as a function of Taylor number

$$\sigma_{\text{critical}} = \mu \frac{\epsilon_{\text{critical}}}{r_1^2} \tag{28}$$

Figure 7 shows a plot of the induced critical stress coefficient as a function of the Taylor number. From Fig. 7, the following appear to be the relationships between the critical induced stress coefficient and the Taylor number for the fluid under study.

$$\text{Ta} < 41.3: \frac{\sigma_{\text{critical}}}{\rho\ U_i^2} = 7.5 \times 10^6 (\text{Ta})^{-3.82} \tag{29}$$

$$\text{Ta} > 41.3: \frac{\sigma_{\text{critical}}}{\rho\ U_i^2} = 208 (\text{Ta})^{-0.974} \tag{30}$$

where $U_i = \omega r_1$. As stated earlier, Fig. 4 shows a plot of the unstable displacements typically observed for Taylor numbers to the left of the critical Taylor number. It demonstrates that the instability that occurs in that region is of a more mechanical type. It is effectively a "sloshing" of the fluid due to the "irregular" motion of the inner cylinder.

As the critical Taylor number is approached, the instability no longer displays the clear-cut nature shown in these figures, that is, the instability is not confined mostly to one wall or the other of the cylinders, but is spread out over the whole plane. It appears that the critical Taylor number is the transition point between the "mechanical" instability of Fig. 4 and the "centrifugal" one of Fig. 5 ($\epsilon/\omega r_1^2 = 0.458$).

B. Analytical Solution of a Simplified Model. While it appears relativley difficult to generate analytical solutions of equation (17) for the general case, we can investigate the case where:

$$V_\Theta \ll 1 \tag{31}$$

Using equation (31), equation (17) asymptotically reduces to

$$\frac{\partial^2 u_2}{\partial t^2} + 2\ \nu\epsilon\ \sin\Omega t \frac{\partial}{r\partial r}\left(\frac{1}{r}\frac{\partial u_2}{\partial r}\right) = 0 \tag{32}$$

In equation (32) we have arbitrarily set the pressure, which is approximately a constant, for small V_Θ, to zero to simplify the analysis. The goal here is to attempt to demonstrate the underlying physics of the approach adopted in this paper through the development of an analytic solution for a relatively simple case, a physics which may not be obvious from the numerical solution generated from the more general approach.

Let $u_2(r, t) = R(r^2/2)\ T(t)$, then using the method of separation of variables, we can show that equation (32) reduces to the following:

$$\ddot{T} + (-2\ \nu\epsilon\lambda^2\ \sin\Omega t)\, T = 0 \tag{33}$$

$$\text{where } \ddot{T} = \frac{d^2 T}{dt^2}$$

$$R(r^2/2) \text{ is given by } R'' + \lambda^2 R = 0 \tag{34}$$

Using the boundary conditions discussed earlier, we can show that:

$$\lambda = \frac{2n\pi}{(r_2^2 - r_1^2)} \quad n = 0,1,2,3\ldots \tag{35}$$

Equation (33) is a Mathieu stability equation, hence for small values of V_Θ the underlying Mathieu stability problem inherent in this approach is brought into sharper focus, and hence the consistency with the Bolotin method for the stability analysis of elastic systems is established. We can solve analytically for the value of $\epsilon_{\text{critical}}$ for the case $V_{\Theta<<1}$.

Typically, the Mathieu Equation is written as follows:

$$\frac{d^2 f}{dt^2} + (K - h^2 \sin\Omega t) f = 0 \tag{36}$$

from equations (33) and (36), we have:

$$h^2 = 2\nu\epsilon\lambda^2 \tag{37}$$

and thus we can obtain the critical values of ϵ for the limiting case considered here, since we know the value of h^2 for the case $K = 0$ from the stability diagram for the Mathieu equation.

From White (1974), $h^2 \sim 1.5$, hence for this limiting case the apparent kinematic viscosity becomes:

$$\epsilon_{\text{critical}} = \frac{(1.5)}{2\nu\ \lambda^2} \tag{38}$$

where λ (for $n = 1$) is given by: $\lambda = 2\pi/(r_2^2 - r_1^2)$

The induced critical stress for the limiting case is given by:

$$\sigma_{cr} = \rho(1.5) \frac{(r_2^2 - r_1^2)^2}{8\ \pi^2 r_1^2} \tag{39}$$

The simplified model shown above indicates the nature of

the stability problem formulated here. It is, in the limit, a Mathieu stabilty problem.

5 Conclusions

A simple approach to flow instability based on the displacement of the fluid particles has been presented. It has been shown that an analysis based on the physical displacement of the fluid particles yields a stability criterion that is indistinguishable from the well known result. Furthermore, the existence of an apparent viscosity for this low Reynolds number flow has been established suggesting that apparent viscosity is perhaps a universal property of disturbed flows. The generalization of this apparent viscosity leads to an equation that is consistent with the Prandtl formulation.

This relatively simple model appears to indicate that stability models based on a philosophical approach analogous to that used for elastic systems could lead to a deeper understanding of the complex physical processes that lead to flow instabilty. It could also, possibly, open up the vast literature in existence on the stability of elastic systems for extension to fluid systems encouraging a useful integration of these two currently disparate fields.

References

Bejan, A., 1981, "On the Buckling Property of Inviscid Jets and the Origins of Turbulence," *Letters in Heat and Mass Transfer*, Vol. 8, p. 187.

Blake, K. R., and Bejan, A., 1984, "Experiments on the Buckling of Thin Fluid Layers Undergoing End Compression," ASME *Journal of Fluids Engineering*, Vol. 106, No. 1, pp. 74–78.

Bolotin, V. V., 1964a, *The Dynamic Stability of Elastic Systems*, Holden-Day, San Francisco, CA.

Bolotin, V. V., 1964b, *The Dynamic Stability of Elastic Systems*, Holden-Day, San Francisco, CA, pp. 196–198.

Cruickshank, J. O., 1980, "Viscous Fluid Buckling: A Theoretical and Experimental Analysis with Extensions to General Fluid Stability," Ph.D. Thesis, Iowa State University, Ames, Iowa.

Cruickshank, J. O., and Munson, B. R., 1981, "Viscous Fluid Buckling of Plane and Axisymmetric Jets," *Journal of Fluid Mechanics*, Vol. 113, p. 221.

Prandtl, L., 1967, *Essentials of Fluid Dynamics*, Blackie and Son, Ltd., London, p. 118.

Reynolds, O., 1883, "An Experimental Investigation of the Circumstances Which Determine Whether the Motion of Water Shall be Sinuous, and of the Law of Resistance in Parallel Channels," *Phil. Trans. Roy. Soc.*, Vol. 174, pp. 935–982.

Schlichting, H., 1968, *Boundary Layer Theory*, 6th Ed., McGraw-Hill, New York, pp. 438–444.

Suleiman, S. M., and Munson, B. R., 1981, "Viscous Fluid Buckling of Thin Fluid Layers," *Phys. Fluids*, Vol. 24, p. 1.

Swope, R. D., and Ames, W. F., 1963, "Vibrations of a Moving Threadline," *J. Franklin Inst.*, Vol. 275, No. 1, pp. 37–55.

White, F. M., 1974, *Viscous Fluid Flow*, McGraw-Hill, New York, pp. 392–397.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N. Y. 10017, or to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

Simplified Rigid-Plastic Beam Analysis

R. B. Schubak,[1,3] M. D. Olson,[2,3] and D. L. Anderson[2,3]

1 Introduction

Limit analysis involving rigid-plastic material idealization and plastic hinging has been applied to laterally loaded rectangular beams with axially constrained ends by Haythornthwaite (1957), Gürkök and Hopkins (1981), and Vaziri et al. (1987). While their solutions are academically valuable, they are of limited applicability. To apply them to nonrectangular beams (notably I-beams and T-beams), would require rederivation for each case. Furthermore, they are very complicated and their extensions to dynamic problems would be impractical. It is therefore desirable to have a simplified solution which is insensitive to section geometry and, when applied to dynamic problems, results in simple, preferably linear, differential equations of motion. The purpose of this note is to present such a solution based on a recent study by Schubak (1986).

2 Assumptions and Analysis

Beams of at least singly symmetric cross section subjected to a uniformly distributed load are considered. The ends of the beam are symmetrically supported at the section's centroidal axis with constraints against axial displacements. The beam material is assumed to be rigid perfectly plastic. Beam deflections are finite but small compared to the beam span, so that the axial force can be taken as constant along the beam span.

Consider a pinned beam of span 2ℓ subjected to a load per unit length p. As p increases beyond the linear collapse load p_o, a plastic hinge is formed at the midspan. The beam begins to deform, and is assumed to continue to deform, in the mechanism of Fig. 1(*a*) with a midspan displacement of w_o. Defining M_o, N_o as the ultimate plastic moment and axial force capacities, respectively, then $p_o = 2M_o/\ell^2$, and moment equilibrium gives

$$\frac{p}{p_o} = m + \left(\frac{N_o w_o}{M_o}\right) n \tag{1}$$

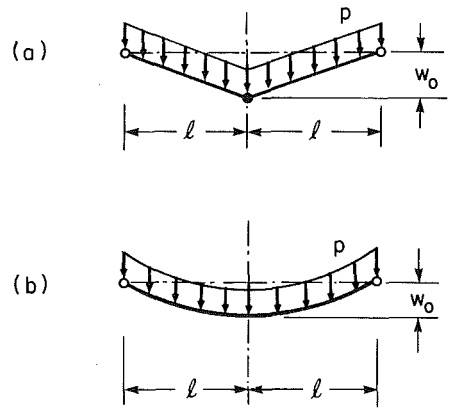


Fig. 1 Deflection modes: (*a*) single midspan hinge mechanism; (*b*) plastic string

where $m = M/M_o$, $n = N/N_o$, M is the moment in the midspan hinge and N is the axial force. To complete the solution, the section's moment-axial force capacity interaction relation (yield condition) and flow rule must be defined.

The plastic flow rule states that the plastic strain increment vector $(N_o\delta\epsilon, M_o\delta\psi)$, where $\delta\epsilon$ and $\delta\psi$ are the increments of longitudinal strain and curvature, respectively, is outwardly normal to the interaction curve. In equation form this becomes

$$\frac{M_o\delta\psi}{N_o\delta\epsilon}\frac{dm}{dn} = -1. \tag{2}$$

For the present mechanism, $\delta\epsilon/\delta\psi = w_o$, and hence,

$$\frac{dm}{dn} = -\frac{N_o w_o}{M_o}. \tag{3}$$

After some finite deformation, the axial force will reach the axial capacity N_o and the beam will no longer have any moment resistance. The beam will respond thereafter as a plastic string (Fig. 1(*b*)) with tension, N_o, for which the midspan displacement is given by

$$w_o = \frac{p\ell^2}{2N_o}. \tag{4}$$

Reorganizing equation (4) and dividing both sides by p_o, the load capacity for large deflections becomes

[1]Graduate Student.
[2]Professor.
[3]Department of Civil Engineering, University of British Columbia, Vancouver, B.C., Canada V6T 1W5.
Manuscript received by ASME Applied Mechanics Division, November 5, 1986; final revision, December 10, 1986.

$$\frac{p}{p_o}=\frac{N_o w_o}{M_o} \tag{5}$$

(*a*) *Rectangular Sections.* For a rectangular cross section, the interaction relation is

$$m=1-n^2. \tag{6}$$

Combining this with equations (1), (2), and (3) gives the load capacity for the bending phase as

$$\frac{p}{p_o}=1+\frac{1}{4}\left(\frac{N_o w_o}{M_o}\right)^2 \tag{7}$$

The preceding analysis was based on the deformation assumption of a single midspan hinge mechanism. Although this mechanism is kinematically admissible and equilibrium has been satisfied, the presence of the axial force requires that the bending moment in the rigid half-beam exceed the plastic moment M_o, a violation of the beam's plasticity condition. However, the results compare very well with those obtained by Gürkök and Hopkins using the correct displacement fields. This comparison is shown in Fig. 2.

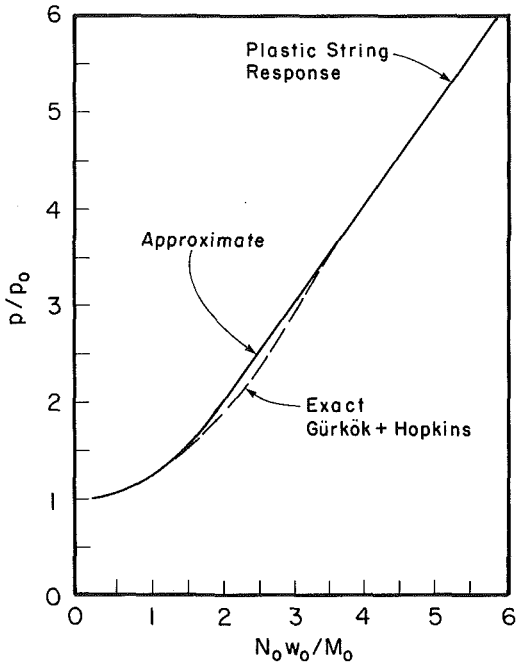


Fig. 2 Load capacity of rectangular beams

(*b*) *Doubly Symmetric, Nonrectangular Sections.* The interaction relations for a rectangular section and an open-web section comprised of two equal concentrated areas are shown in Fig. 3. These two sections may be considered to be the limiting cases of practical bending sections—the interaction relations of all doubly symmetric I-beams and box-beams lie between the aforementioned pair.

Consider the linear interaction relation of the open-web section. At the stress states $(n, m) = (0, 1)$ and $(1, 0)$, the normal to the interaction curve is undefined and thus the plastic strain increment vector may lie anywhere within the appropriate fan. Putting $m = 1$ and $n = 0$ into equation (1) gives the load capacity as $p = p_o$, which is valid for $0 \le w_o \le M_o/N_o$. For larger deflections, the stress state "travels" instantaneously along the interaction relation to $(n, m) = (1, 0)$ and the plastic string solution of equation (5) then applies.

The result is plotted in Fig. 4 along with the solution for the rectangular section. The solution for practical bending sections such as I-beams and box-beams will then lie between these two cases.

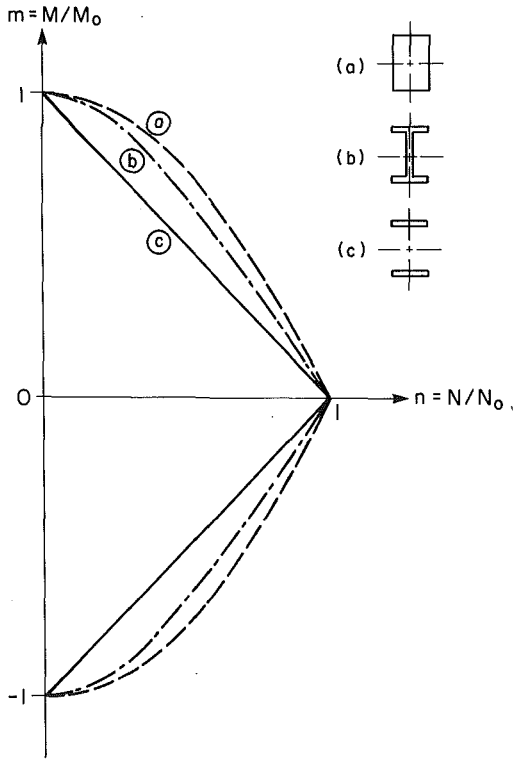


Fig. 3 Interaction curves for doubly symmetric sections: (a) rectangular; (b) I or box; (c) open-web

(*c*) *Singly Symmetric Sections.* Consider a T-beam with axial constraint applied at the centroid. The section properties and its interaction diagram are shown in Fig. 5. The usual plastic moment capacity (no axial force) for this section is $135\,\sigma_o$, where σ_o is the yield stress. However, the ultimate moment capacity defined as M_o herein is $162\,\sigma_o$ which occurs when the neutral axis coincides with the centroidal axis, resulting in a net axial force.

The interaction relation for this section is

$$m=\begin{cases}\dfrac{5}{9}(1-4n-5n^2), & \text{if } n<-1/3;\\[2ex] \dfrac{5}{18}(3-2n-n^2), & \text{if } n\ge -1/3.\end{cases} \tag{8}$$

Following the same procedure as for the rectangular section, the load capacity is found to be

$$\frac{p}{p_o}=\begin{cases}1-\dfrac{2}{5}\left(\dfrac{N_o w_o}{M_o}\right)+\dfrac{9}{100}\left(\dfrac{N_o w_o}{M_o}\right)^2, & \text{if } 0\le \dfrac{N_o w_o}{M_o}\le \dfrac{10}{27};\\[2ex] \dfrac{10}{9}-\dfrac{N_o w_o}{M_o}+\dfrac{9}{10}\left(\dfrac{N_o w_o}{M_o}\right)^2, & \text{if } \dfrac{10}{27}<\dfrac{N_o w_o}{M_o}\le \dfrac{10}{9};\\[2ex] \dfrac{N_o w_o}{M_o}, & \text{if } \dfrac{N_o w_o}{M_o}>\dfrac{10}{9}.\end{cases} \tag{9}$$

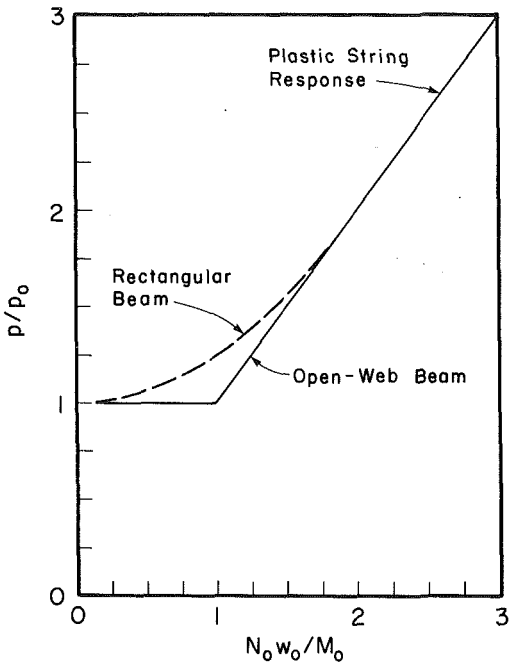


Fig. 4 Load capacity of doubly symmetric beams

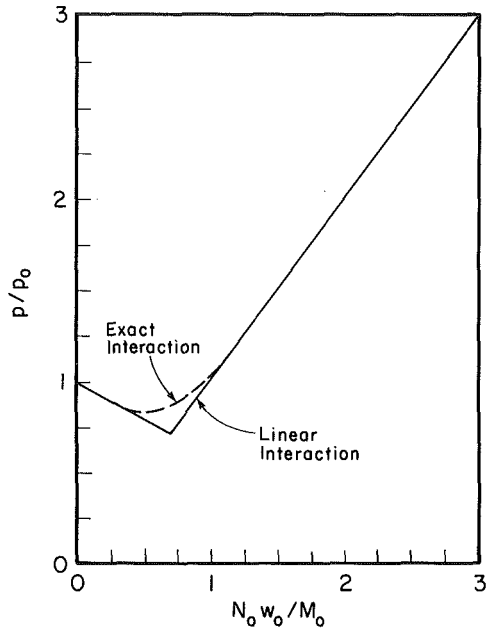


Fig. 6 Load capacity of T-beam

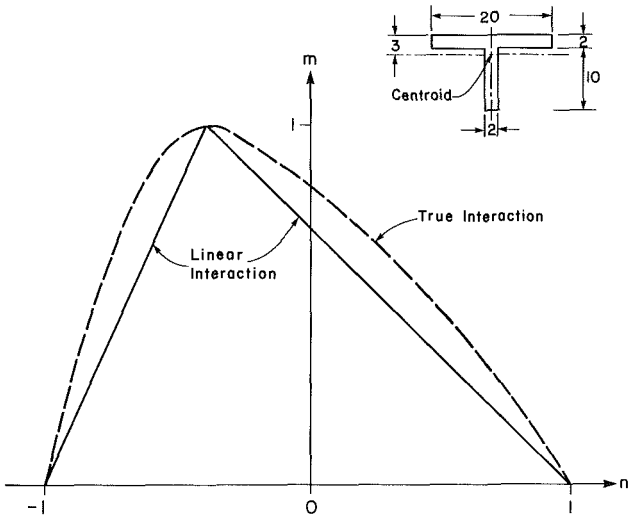


Fig. 5 True and approximate interaction curves for T-beam section

The interaction relation can be approximated by linear segments connecting nodes at $(n, m) = (-1, 0)$, $(-2/5, 1)$ and $(1, 0)$ (see Fig. 5). From consideration of the flow rule it is seen that the stress state of a plastic section must always be at one of the nodes. As a rsult, $(n, m) = (-2/5, 1)$ for $0 \le N_o w_o/M_o \le 5/7$ and $(n, m) = (1, 0)$ for $N_o w_o/M_o > 5/7$. Putting the first stress state into equation (1) and recalling equation (5), the approximate load capacity becomes

$$\frac{p}{p_o} = \begin{cases} 1 - \frac{2}{5}\left(\frac{N_o w_o}{M_o}\right), & \text{if } 0 \le \frac{N_o w_o}{M_o} \le \frac{5}{7}; \\ \frac{N_o w_o}{M_o}, & \text{if } \frac{N_o w_o}{M_o} > \frac{5}{7}. \end{cases} \tag{10}$$

The load capacity of the pinned T-beam is plotted in Fig. 6 for both the true interaction relation and the linear approximation. The initial negative slope in Fig. 6 is due to the compressive axial force which develops instantaneously as a consequence of the rigid-plastic assumption. This part of the curve represents an equilibrium solution but obviously not a stable one, and would be changed by the inclusion of elastic effects. However, the main purpose here was to show the effect of the linear interaction approximation.

(*d*) *Clamped Beams.* For a doubly symmetric clamped beam, equations (1) to (7), and all other terms such as the definition of p_o and the labelling of the axes in Figs. 2 and 4 are valid if M_o is simply replaced by $2M_o$.

3 Conclusions

The static response of axially restrained beams to uniformly distributed loads has been studied. An approximate deformation mechanism was used to derive simple results for the load capacity in terms of the moment and axial load in the plastic hinge(s). Solutions were derived for beams of various doubly symmetric and singly symmetric cross sections. It was found that replacing the interaction relation of a section with a linear approximation thereof results in a good estimate of the beam's load capacity and greatly simplifies the analysis. This simplification is particularly beneficial when applied to the dynamic response of axially restrained beams as shown by Schubak (1986). This work will be presented in a future paper.

Acknowledgment

This work has been supported by the Canadian Department of National Defence through a contract with the Defence Research Establishment Suffield.

References

Gürkök, A., and Hopkins, H. G., 1981, "Plastic Beams at Finite Deflection Under Transverse Load with Variable End-Constraint," *Journal of the Mechanics and Physics of Solids*, Vol. 29, No. 5/6, pp. 447-476.

Haythornthwaite, R. M., 1957, "Beams with Full End Fixity," *Engineering*, Vol. 25, Jan., pp. 110-112.

Schubak, R. B., 1986, "Finite Deflection Dynamic Response of Axially Restrained Beams," M.A.Sc. Thesis, Department of Civil Engineering, University of British Columbia, Vancouver, Canada.

Vaziri, R., Olson, M. D., and Anderson, D. L., 1987, "Dynamic Response of Axially Constrained Plastic Beams to Blast Loads," *Int. J. Solids and Structures*, Vol. 23, No. 1, pp. 153-174.

Stress Concentration in Fiber Composite Sheets Including Matrix Extension

J. N. Rossettos[4,6] and M. Shishesaz[5,6]

Introduction

The study of stress concentration which occurs near broken fibers in filamentary composites is important in failure prediction. The first analysis of this problem was carried out by Hedgepeth (1961), who used a shear lag model (SLM), where the fibers carry only axial loads and the matrix carries only shear loads. Recent work along these lines can be found in references by Reedy (1984). The present work is a direct extension of Hedgepeth's analysis in that the matrix now also carries axial load. This would be appropriate for a relatively stiff matrix. As it turns out, the results show that the SLM is surprisingly useful even for typical metal matrix materials. Similar trends are indicated by Reedy (1984), who used 3-D finite element calculations.

In our analysis, the matrix adjacent to the last fiber break (of a series of in-line fiber breaks or crack) is assumed to either develop a line yield zone perpendicular to the fibers, or to break itself. Both cases are compared, and of particular interest is the fact that they bracket the SLM. The analysis deals with resultant loads, so that a structural rather than a continuum approach is used.

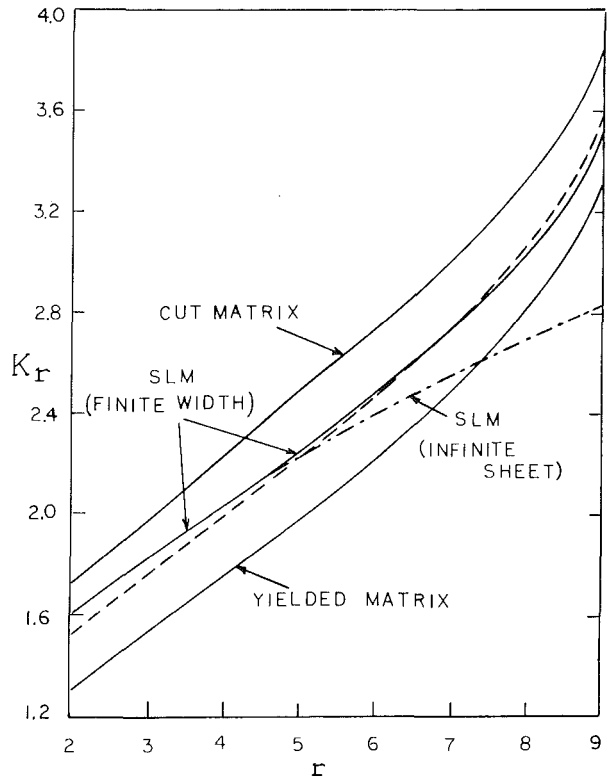


Fig. 1 Stress concentration factor in the first unbroken fiber versus number of broken filaments, *r*. Total number of fibers is 21; $V_f = V_m = 0.5$; $\epsilon = 0.088$; $\omega = 0.70$.

Analysis

We consider a finite width sheet consisting of $2q+1$ fibers, parallel to the x axis, which are numbered so that the center fiber is the zeroeth, and the fibers above it are numbered 1 to q. Below it they are numbered -1 to $-q$. The lamina is loaded in the x (or axial direction) and fiber breaks occur along the $x = 0$ axis. A quadratic displacement, u, is assumed in a given matrix bay as,

$$u = u^m_{n,n-1} + (y/d)(u_n - u_{n-1}) + 2(y/d)^2(u_n + u_{n-1} - 2u^m_{n,n-1}) \tag{1}$$

In equation (1) the origin of the local $x-y$ system is at the center of each matrix bay; $u^m_{n,n-1}$ is the displacement at the center of the $(n, n-1)$ bay, u_n and u_{n-1} are the nth and $n-1$ fiber displacements, and d is the matrix width. Using equation (1), the shear stress in a matrix bay is $\tau_{xy} = G\,\partial u/\partial y$. Its change, $\Delta\tau_{xy}$, over a bay in the y direction is

$$(\Delta\tau_{xy})_{n,n-1}) = \int_{-d/2}^{d/2} (\partial\tau_{xy}/\partial y)_{n,n-1} dy = 4Ge_{n,n-1}/d \tag{2}$$

where

$$e_{n,n-1} = u_n + u_{n+1} - 2u^m_{n,n-1} \tag{3}$$

so that equilibrium of the matrix bay below, say, the nth fiber yields, $dp^m_{n,n-1}/dx + h(\Delta\tau_{xy})_{n,n-1} = 0$, where the matrix net axial load is

$$p^m_{n,n-1} = \int_{-d/2}^{d/2} E_m (du/dx) h dy$$
$$= E_m A_m (du^m_{n,n-1}/dx + de_{n,n-1}/6dx) \tag{4}$$

h is the lamina thickness and E_m is matrix modulus. The axial load in the nth fiber is $p_n = E_f A_f du_n/dx$. With these relations, the equilibrium equations for a typical fiber and matrix bay are easily derived in terms of displacements (Shishesaz, 1986). Note from equations (2) and (3) that $e_{n,n-1} = 0$ corresponds to the SLM. Dimensionless coordinates, loads, displacements are next defined by

$$x = (E_f A_f d/Gh)^{1/2}\xi;\ (p_n, p^m_{n,n-1}) = p(P_n, P^m_{n,n-1})$$
$$(u_n, e_{n,n-1}, u^m_{n,n-1}) = p(d/E_f A_f Gh)^{1/2}(U_n, E_{n,n-1}, U^m_{n,n-1}) \tag{5}$$

where p is the fiber load value at $x = \infty$. The equilibrium equations for all fibers and matrix bays are then

$$d^2U_n/d\xi^2 + (U_{n+1} - 2U_n + U_{n-1}) - 2(E_{n+1,n} + E_{n,n-1}) = 0$$
$$(-q+1 \le n \le q-1) \tag{6}$$

$$\epsilon(d^2U_n/d\xi^2 + d^2U_{n-1}/d\xi^2 - 2d^2E_{n,n-1}/3d\xi^2) + 4E_{n,n-1} = 0$$
$$(-q+1 \le n \le q-1) \tag{7}$$

$$d^2U_n/d\xi^2 + (U_{n-1} - U_n) - 2E_{n,n-1} = 0 \quad (n = q) \tag{8}$$

$$d^2U_n/d\xi^2 + (U_{n+1} - U_n) - 2E_{n+1,n} = 0 \quad (n = -q) \tag{9}$$

where the parameter, ϵ, is $\epsilon = A_m E_m/2A_f E_f = V_m E_m/2V_f E_f$, and V_m, V_f are matrix and fiber volume fractions. It is noted that as $\epsilon \to 0$ (i.e., E_m is very small compared to E_f), $E_{n,n-1} \to 0$ from equation (7), and equations (6)–(9) reduce to those of the SLM. The E's are expressed in terms of the U's using equations (3) and (5) so that the $4q+1$ equations (6)–(9) are written as

$$\mathbf{L}_1\mathbf{U}'' - \mathbf{L}_2\mathbf{U} = 0 \tag{10}$$

where

$$\mathbf{U}^T = [U_q, U^m_{q,q-1}, U_{q-1}, \ldots\ldots, U_{-q+1}, U^m_{-q+1,-q}, U_{-q}]$$

and matrices $\mathbf{L}_1$ and $\mathbf{L}_2$ are banded. It is convenient to assume a uniform strain at infinity, so that

[4]Professor, Mem. ASME.
[5]Graduate Student.
[6]Department of Mechanical Engineering, Northeastern University, Boston, MA 02115.
Manuscript received by ASME Applied Mechanics Division, August 5, 1986.

$$p^m_{n,n-1} = (A_m E_m / A_f E_f) p = 2\epsilon p \tag{11}$$

and the dimensionless loads have the behavior

$$P_n \to 1;\ P^m_{n,n-1} \to 2\epsilon \quad \text{as } \xi \to \infty \tag{12}$$

In the case where we assume a line yield zone, with yield stress, S_y, the net force in the matrix bay is $p^m_y = S_y hd$. If there are r broken fibers, say $n = 0, 1, 2. \ldots .r-1$, and $r-1$ cut matrix bays the boundary conditions are

$$\begin{aligned} U_n(0) &= 0 \quad n<0 \quad n\geq r \\ P_n(0) &= 0 \quad 0\leq n\leq r-1 \\ U^m_{n,n-1}(0) &= 0 \quad n\leq -1 \quad n\geq r+1 \\ P^m_{n,n-1}(0) &= 0 \quad n\leq r-1 \quad n\geq 1 \\ P^m_{n,n-1}(0) &= \omega \quad n=0 \quad n=r \end{aligned} \tag{13}$$

where $\omega = p^m_y/p$. In the case where the matrix next to the last fiber break is also cut, we set $\omega = 0$ in the last of equations (13). A solution to equation (10) is assumed in the form $\mathbf{U} = \mathbf{R}e^{\lambda\xi}$, where $\mathbf{U}$ and $\mathbf{R}$ are vectors of order $4q + 1$. The resulting eigenvalue problem, $(\mathbf{L}_2 - \lambda^2\mathbf{L}_1)\mathbf{R} = 0$. leads to eigenvalues, λ_i, and eigenvectors, $\mathbf{R}^i$. Positive λ_i are discarded to satisfy (12), and the solution to $\mathbf{U}$ can be written as the expansion

$$\mathbf{U} = \sum_{i=1}^{4q+1} C_i \mathbf{R}^i e^{\lambda i \xi} + \mathbf{U}_P \tag{14}$$

where the particular solution, $\mathbf{U}^T_P = [\xi, \xi, \ldots\ldots, \xi]$ is added in equation (14) so that (12) is satisfied. Since $\mathbf{U}$ contains both fiber, U_n, and matrix, $U^m_{n,n-1}$, displacements, differentiation and use of previous relations lead to the load quantities which also satisfy (12). Equations (13) represent $4q+1$ equations for the unknown constants C_i, and are easily solved (Shishesaz, 1986).

Results and Discussion

In Fig. 1, the stress concentration faction ($K_r = P_r$), is plotted against the number of broken fibers, r. The values of ϵ and ω are typical of boron-aluminum. The dashed curve is the average of the two bracket cases (i.e., upper and lower curves, which include matrix extension) and differs from the SLM (finite width case) by at most 5 percent. The smaller differences for larger r may be due to increased matrix shear action when more fibers break. Smaller values of ϵ bring results even closer to the SLM. For instance (Shishesaz, 1986), for $r = 3$, in the yielded matrix case, when $\epsilon = 0.056$ and 0.0064 (with $\omega = 0.444$ and 0.051) get $K_r = 1.63$ and 1.81 which approach the SLM value ($K_r = 1.83$). In all cases, the value of ω corresponds to a matrix force at infinity of 1/4 the matrix yield load, and assures that the matrix adjacent to the yielded or cut matrix does not yield. Finally, since we deal with a finite width sheet (i.e., 21 total fibers), note that in Fig. 1, for $r = 6$, the last break is 5 unbroken fibers from the edge and K_r (SLM, dash curve) is seen to deviate from the infinite sheet case (dash-dot curve) discussed by Hedgepeth (1961).

References

Hedgepeth, J. M., 1961, "Stress Concentrations in Filamentary Structures," NASA TN D - 882.

Reedy, E. D., Jr., 1984, "Fiber Stresses in a Cracked Monolayer: Comparison of Shear-Lag and 3-D Finite Element Predictions," *J. Composite Materials*, Vol. 18, pp. 595-607.

Shishesaz, M., 1986, "Stress Analysis of Fiber Composite Monolayers with Broken Filaments," Ph.D. Thesis, Northeastern University, Boston, MA.

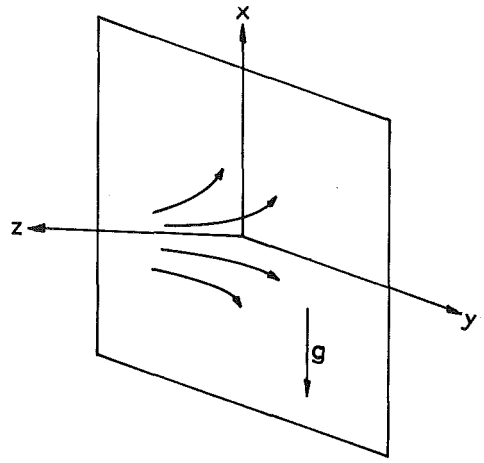


Fig. 1 The coordinate system

Effect of Natural Convection on the Axisymmetric Stagnation Flow on a Vertical Plate

C. Y. Wang[7]

Introduction

Consider a plate heated by a hot, impinging axisymmetric flow. The fluid dynamic problem near the stagnation point is the Homann (1936) exact solution and the forced heat convection problem was considered by Wang (1973). However, if the plate is vertical and the stagnation flow is weak, free convection effects may become important. This note considers the effect of free convection on the axisymmetric stagnation flow on a vertical plate. We shall use the full Navier-Stokes equations and the Boussinesq energy equation. The present note is an exact solution of these equations.

Formulation

Figure 1 shows the plate in the xy plane with x in the vertical direction. The axisymmetric jet is impinging on the plate at the origin. The plate has temperature T_w and the conditions far from the plate are

$$u = bx, \qquad v = by, \qquad w = -2bz$$

$$p = -\frac{\rho}{2} b^2(x^2+y^2) + p_o, \qquad T = T_\infty \tag{1}$$

where (u, v, w) are Cartesian velocity components, p is the pressure, ρ is the density, T is the temperature, and b is the strength of the axisymmetric stagnation flow.

For similarity solutions set

$$u = \frac{g\beta(T_\infty - T_w)}{b} M(\eta) + bx\varphi'(\eta) \tag{2}$$

$$v = by\varphi'(\eta) \tag{3}$$

$$w = -2\sqrt{b\nu}\varphi(\eta) \tag{4}$$

$$T - T_\infty = (T_w - T_\infty)H(\eta) \tag{5}$$

where ν is the kinematic viscosity, g is the gravitational acceleration, β is the coefficient of thermal expansion and

$$\eta \equiv \sqrt{\frac{b}{\nu}}\, z, \tag{6}$$

[7]Professor, Departments of Mathematics and Mechanical Engineering, Michigan State University, East Lansing, MI 48824. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, December 23, 1986; final revision, March 4, 1987.

$$p^m_{n,n-1} = (A_m E_m / A_f E_f) p = 2\epsilon p \tag{11}$$

and the dimensionless loads have the behavior

$$P_n \to 1;\ P^m_{n,n-1} \to 2\epsilon \quad \text{as } \xi \to \infty \tag{12}$$

In the case where we assume a line yield zone, with yield stress, S_y, the net force in the matrix bay is $p^m_y = S_y hd$. If there are r broken fibers, say $n = 0, 1, 2. \ldots .r-1$, and $r-1$ cut matrix bays the boundary conditions are

$$\begin{aligned} U_n(0) &= 0 \quad n<0 \quad n\geq r \\ P_n(0) &= 0 \quad 0\leq n\leq r-1 \\ U^m_{n,n-1}(0) &= 0 \quad n\leq -1 \quad n\geq r+1 \\ P^m_{n,n-1}(0) &= 0 \quad n\leq r-1 \quad n\geq 1 \\ P^m_{n,n-1}(0) &= \omega \quad n=0 \quad n=r \end{aligned} \tag{13}$$

where $\omega = p^m_y/p$. In the case where the matrix next to the last fiber break is also cut, we set $\omega = 0$ in the last of equations (13). A solution to equation (10) is assumed in the form $\mathbf{U} = \mathbf{R}e^{\lambda\xi}$, where $\mathbf{U}$ and $\mathbf{R}$ are vectors of order $4q + 1$. The resulting eigenvalue problem, $(\mathbf{L}_2 - \lambda^2\mathbf{L}_1)\mathbf{R} = 0$. leads to eigenvalues, λ_i, and eigenvectors, $\mathbf{R}^i$. Positive λ_i are discarded to satisfy (12), and the solution to $\mathbf{U}$ can be written as the expansion

$$\mathbf{U} = \sum_{i=1}^{4q+1} C_i \mathbf{R}^i e^{\lambda i \xi} + \mathbf{U}_P \tag{14}$$

where the particular solution, $\mathbf{U}^T_P = [\xi, \xi, \ldots\ldots, \xi]$ is added in equation (14) so that (12) is satisfied. Since $\mathbf{U}$ contains both fiber, U_n, and matrix, $U^m_{n,n-1}$, displacements, differentiation and use of previous relations lead to the load quantities which also satisfy (12). Equations (13) represent $4q+1$ equations for the unknown constants C_i, and are easily solved (Shishesaz, 1986).

Results and Discussion

In Fig. 1, the stress concentration faction ($K_r = P_r$), is plotted against the number of broken fibers, r. The values of ϵ and ω are typical of boron-aluminum. The dashed curve is the average of the two bracket cases (i.e., upper and lower curves, which include matrix extension) and differs from the SLM (finite width case) by at most 5 percent. The smaller differences for larger r may be due to increased matrix shear action when more fibers break. Smaller values of ϵ bring results even closer to the SLM. For instance (Shishesaz, 1986), for $r = 3$, in the yielded matrix case, when $\epsilon = 0.056$ and 0.0064 (with $\omega = 0.444$ and 0.051) get $K_r = 1.63$ and 1.81 which approach the SLM value ($K_r = 1.83$). In all cases, the value of ω corresponds to a matrix force at infinity of 1/4 the matrix yield load, and assures that the matrix adjacent to the yielded or cut matrix does not yield. Finally, since we deal with a finite width sheet (i.e., 21 total fibers), note that in Fig. 1, for $r = 6$, the last break is 5 unbroken fibers from the edge and K_r (SLM, dash curve) is seen to deviate from the infinite sheet case (dash-dot curve) discussed by Hedgepeth (1961).

References

Hedgepeth, J. M., 1961, "Stress Concentrations in Filamentary Structures," NASA TN D - 882.

Reedy, E. D., Jr., 1984, "Fiber Stresses in a Cracked Monolayer: Comparison of Shear-Lag and 3-D Finite Element Predictions," *J. Composite Materials*, Vol. 18, pp. 595-607.

Shishesaz, M., 1986, "Stress Analysis of Fiber Composite Monolayers with Broken Filaments," Ph.D. Thesis, Northeastern University, Boston, MA.

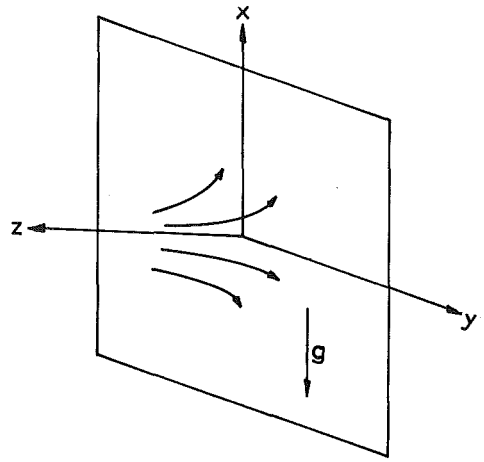


Fig. 1 The coordinate system

Effect of Natural Convection on the Axisymmetric Stagnation Flow on a Vertical Plate

C. Y. Wang[7]

Introduction

Consider a plate heated by a hot, impinging axisymmetric flow. The fluid dynamic problem near the stagnation point is the Homann (1936) exact solution and the forced heat convection problem was considered by Wang (1973). However, if the plate is vertical and the stagnation flow is weak, free convection effects may become important. This note considers the effect of free convection on the axisymmetric stagnation flow on a vertical plate. We shall use the full Navier-Stokes equations and the Boussinesq energy equation. The present note is an exact solution of these equations.

Formulation

Figure 1 shows the plate in the xy plane with x in the vertical direction. The axisymmetric jet is impinging on the plate at the origin. The plate has temperature T_w and the conditions far from the plate are

$$u = bx, \qquad v = by, \qquad w = -2bz$$

$$p = -\frac{\rho}{2} b^2(x^2+y^2) + p_o, \qquad T = T_\infty \tag{1}$$

where (u, v, w) are Cartesian velocity components, p is the pressure, ρ is the density, T is the temperature, and b is the strength of the axisymmetric stagnation flow.

For similarity solutions set

$$u = \frac{g\beta(T_\infty - T_w)}{b} M(\eta) + bx\varphi'(\eta) \tag{2}$$

$$v = by\varphi'(\eta) \tag{3}$$

$$w = -2\sqrt{b\nu}\varphi(\eta) \tag{4}$$

$$T - T_\infty = (T_w - T_\infty)H(\eta) \tag{5}$$

where ν is the kinematic viscosity, g is the gravitational acceleration, β is the coefficient of thermal expansion and

$$\eta \equiv \sqrt{\frac{b}{\nu}}\, z, \tag{6}$$

[7]Professor, Departments of Mathematics and Mechanical Engineering, Michigan State University, East Lansing, MI 48824. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, December 23, 1986; final revision, March 4, 1987.

Table 1

Pr	$H'(0)$	$M'(0)$
0.2	−0.4038	−0.6312
0.7	−0.6654	−0.4954
2.	−0.9862	−0.3912
7.	−1.5458	−0.2849
20.	−2.2297	−0.2131
70.	−3.4276	−0.1474

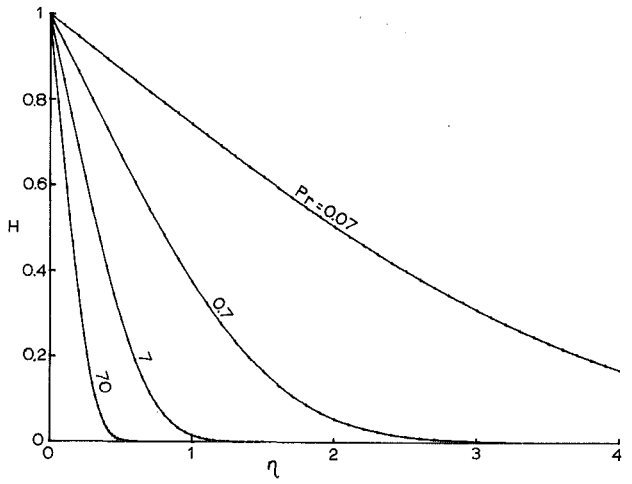


Fig. 2 Temperature distribution $H(\eta)$

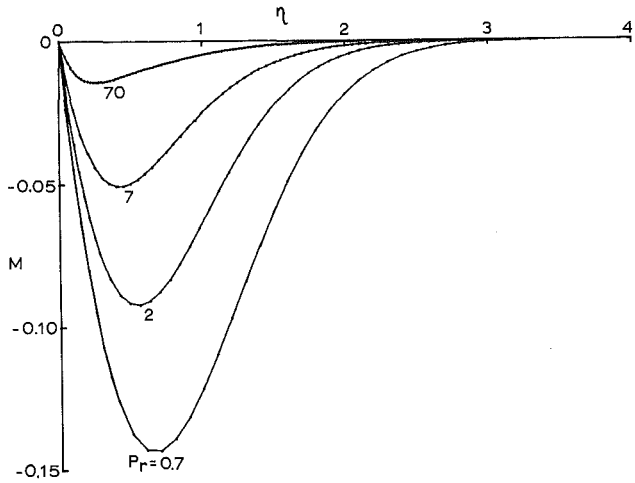


Fig. 3 Velocity due to free convection $M(\eta)$

The Navier-Stokes and energy equation yield

$$p = -\frac{\rho}{2}\left[b^2(x^2+y^2) + w^2 - 2\nu w_z\right] + p_0 \tag{7}$$

$$(\varphi')^2 - 2\varphi\varphi'' = 1 + \varphi''' \tag{8}$$

$$M\varphi' - 2\varphi M' = M'' - H \tag{9}$$

$$-2\mathrm{Pr}\varphi H' = H'' \tag{10}$$

Here Pr is the Prandtl number. The boundary conditions are

$$\varphi(0) = \varphi'(0) = 0, \qquad \varphi'(\infty) = 1 \tag{11}$$

$$M(0) = 0, \qquad M(\infty) = 0 \tag{12}$$

$$H(0) = 1, \qquad H(\infty) = 0 \tag{13}$$

Equations (8), (11) characterize the Homann stagnation point flow. Numerical solution gives an initial value of $\varphi''(0) = 1.1311937$. Equations (10), (13) represent the temperature distribution which is unaffected by the natural convection.

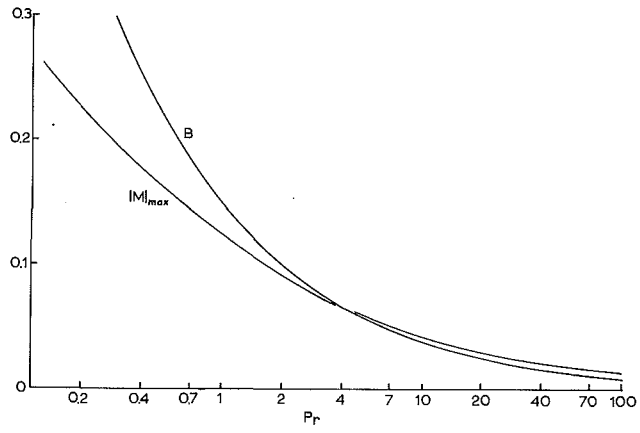


Fig. 4 Maximum velocity $|M|_{max}$ and net flow B as a function of Prandtl number

The effect of natural convection, not studied before, is felt in equation (9) which governs the vertical fluid velocity.

The Numerical Results and Discussion

The numerical scheme is as follows. Using the value of $\varphi''(0)$ and a given Prandtl number we guess $H'(0)$ and integrate equations (8), (10) as an initial value problem by the Runge-Kutta-Fehlberg algorithm. The value of $H'(0)$ is adjusted such that H remains zero for some large η (say, greater than 6). Then equations (8), (9), (10) are integrated again together with a guessed initial value $M'(0)$. The solution is obtained through successive one-parameter shooting. Notice the integral form for H from equation (10) is inconvenient in solving equation (9). Also equation (9) is linear but nonhomogeneous, making shooting necessary. The initial values thus obtained are listed in Table 1 for some representative Prandtl numbers. For air Pr = 0.7 and for water Pr = 7. The net heat transfer is proportional to $H'(0)$. In general $H(\eta)$ decays to zero in a thermal boundary layer (Fig. 2) with a thickness dependent on Pr.

The buoyancy driven flow profile $M(\eta)$ is shown in Fig. 3. The downward velocity (for a cool plate) decreases with increased Prandtl number. The net downward flow per width F is

$$F = g\beta(T_\infty - T_w)\nu^{1/2}b^{-3/2}B, \qquad B \equiv \int_0^\infty M d\eta \tag{14}$$

The values of B and $M_{\max}$ are plotted in Fig. 4. For small Prandtl numbers, free convection may become considerable. The shear stress on the plate is

$$\tau_x = \mu\sqrt{\frac{b}{\nu}}\left[\frac{g\beta(T_\infty - T_w)}{b}M'(0) + bx\varphi''(0)\right] \tag{15}$$

$$\tau_y = \mu\sqrt{\frac{b}{\nu}}\,by\varphi''(0) \tag{16}$$

The stagnation point is at

$$x = -\frac{g\beta(T_\infty - T_w)M'(0)}{b^2\varphi'(0)}, \quad y = z = 0 \tag{17}$$

Since $M'(0)$ is negative, the stagnation point is *above* the center line of the hot jet. The shear lines are rays from the stagnation point. The flow field, however, is highly three-dimensional.

References

Homann, F., 1936, "Der Einfluss Grosser Zähigkeit bei der Strömung um den Zylinder und um ide Kugel," *Zeit. angew. Math. Mech.*, Vol. 16, pp. 153–164.

Wang, C. Y., 1973, "Axisymmetric Stagnation Flow Towards a Moving Plate," *AIChE J.*, Vol. 19, pp. 1080–1081.

Will the Force Method Come Back?

C. A. Felippa[8]

Introduction

The force method of structural analysis was overtaken by the displacement method in the mid Sixties, and has disappeared from the scene except for few specialized applications. Virtually all existing general-purpose finite element programs are based on the *direct stiffness method* (DSM) introduced three decades ago by Turner et al. (1956) and Turner (1959). The DSM combines the displacement method of solution with the direct, element-by-element assembly of the stiffness equations. The simplicity and efficiency of the DSM for *general* applications have not been matched to date; it has the polished "black box" feeling of Lagrange's analytical mechanics.

There has been, however, a modest revival of interest in the force method as manifested in recent publications (Kaneko et al., 1982, 1985; Heath et al., 1984; Patnaik, 1986a, 1986b). This activity has been fueled by the hope that the force method may be competitive for a limited but important class of problems, namely those calling for a sequence of analyses of "perturbed" linear or materially-nonlinear structures, provided the perturbations *do not affect the discrete equilibrium equations*. This situation arises in fully-stressed design and optimization.

The main weakness of the conventional force method on the computer is the difficulty of automating the selection of force unknowns that optimize matrix sparseness while maintaining numerical stability. Twenty years ago Fraeijs de Veubeke (1965, p. 83) noted, "A great step forward in the automation of the (self-straining) calculations would be achieved if the computer itself could be taught to investigate the topology of the (connection) matrices and deduce the self-strainings confined to the smallest numbers of elements." Recent developments try to satisfy this goal by taking advantage of more advanced numerical techniques that were known in 1965 (for example, sparse orthogonal factorizations).

The purpose of this Note is to give the general formulation of the force method from the standpoint of the field of mathematical programming, and to call attention to links between this formulation and recent efforts in the field of matrix structural analysis. The relationship is noteworthy because great strides have been made in large-scale constrained optimization during the past 15 years while the force method has been neglected. It is hoped that the relations described here may spur further research work as well as development of computer-based applications using the abundant software now available in scientific libraries for linear and nonlinear programming.

The Lagrangian System

To exhibit the connections between mathematical programming (MP) and matrix structural analysis it is convenient to start from the Lagrangian system for the quadratic programming (QP) problem. The MP terminology used below follows the up-to-date textbooks of Fletcher (1981) and Gill et al. (1982). The matrix and vector notation, however, is more in line with that of structural mechanics.

The QP problem under equality constraints can be stated as

$$\begin{aligned} &\underset{\mathbf{f}}{\text{maximize}} \quad -\tfrac{1}{2}\mathbf{f}^T\mathbf{C}\mathbf{f}-\mathbf{b}^T\mathbf{f}, \\ &\text{subject to} \quad \mathbf{A}^T\mathbf{f}=\mathbf{p}. \end{aligned} \tag{1}$$

Here $\mathbf{b}$ and $\mathbf{p}$ are *given* vectors of dimension n and m, respectively, with $m \le n$; $\mathbf{C}$ is a $n \times n$ positive-definite symmetric Hessian matrix, $\mathbf{A}$ is an $n \times m$ constraint matrix of rank m, and $\mathbf{f}$ is the n vector of unknowns. The associated Lagrangian function is

$$L(\mathbf{f},\mathbf{u}) = -\tfrac{1}{2}\mathbf{f}^T\mathbf{C}\mathbf{f}-\mathbf{b}^T\mathbf{f}+\mathbf{u}^T(\mathbf{A}^T\mathbf{f}-\mathbf{p}), \tag{2}$$

where $\mathbf{u}$ is a vector of m Lagrange multipliers. The stationarity conditions $\partial L/\partial \mathbf{f} = \partial L/\partial \mathbf{u} = \mathbf{0}$ yield the system of linear algebraic equations

$$\begin{bmatrix} -\mathbf{C} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{f} \\ \mathbf{u} \end{Bmatrix} = \begin{Bmatrix} \mathbf{b} \\ \mathbf{p} \end{Bmatrix}. \tag{3}$$

Following Fletcher (1981, p. 84), choose any $n \times (n-m)$ matrix $\mathbf{V}$ such that the augmented $n \times n$ matrix $[\mathbf{A}\ \mathbf{V}]$ is nonsingular, and let its inverse be expressed in the partitioned form

$$[\mathbf{A}\ \mathbf{V}]^{-1} = \begin{bmatrix} \mathbf{S}^T \\ \mathbf{B}^T \end{bmatrix} \tag{4}$$

It follows that

$$\mathbf{S}^T\mathbf{A}=\mathbf{I}, \quad \mathbf{B}^T\mathbf{A}=\mathbf{0}. \tag{5}$$

This shows that $\mathbf{S}$ is a left generalized inverse of $\mathbf{A}$ (such inverse is of course nonunique if $m < n$) whereas $\mathbf{B}$ is a basis for the null space of $\mathbf{A}$. With these definitions the general solution of equation (3) can be expressed as

$$\begin{Bmatrix} \mathbf{f} \\ \mathbf{u} \end{Bmatrix} = \begin{bmatrix} \mathbf{H} & -\mathbf{T} \\ -\mathbf{T}^T & \mathbf{U} \end{bmatrix} \begin{Bmatrix} \mathbf{b} \\ \mathbf{p} \end{Bmatrix}, \tag{6}$$

in which

$$\mathbf{H}=\mathbf{B}\mathbf{F}^{-1}\mathbf{B}^T, \quad \mathbf{T}=(\mathbf{I}-\mathbf{H}\mathbf{C})\mathbf{S}, \quad \mathbf{U}=\mathbf{S}^T\mathbf{C}\mathbf{T}, \quad \mathbf{F}=\mathbf{B}^T\mathbf{C}\mathbf{B}. \tag{7}$$

The positive-definite symmetric matrix $\mathbf{F}$ is called the *reduced Hessian matrix* in the MP literature.

Staged Elimination

In practice the explicit inverse expressions (6) are rarely used to solve the Lagrangian system (3). Rather, some form of staged elimination of the unknowns $\mathbf{f}$ or $\mathbf{u}$ is explicitly or implicitly invoked.

To eliminate $\mathbf{f}$ in favor of $\mathbf{u}$ the first matrix equation is premultiplied by $\mathbf{A}^T\mathbf{C}^{-1}$ and added to the second, which yields

$$\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}\mathbf{u}=\mathbf{p}+\mathbf{A}^T\mathbf{C}^{-1}\mathbf{b}, \tag{8}$$

or

$$\mathbf{K}\mathbf{u}=\mathbf{p}+\mathbf{p}', \tag{9}$$

where $\mathbf{K} = \mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}$ is a positive definite $m \times m$ matrix, and $\mathbf{p}' = \mathbf{A}^T\mathbf{C}^{-1}\mathbf{b}$. Once $\mathbf{u}$ is calculated, it can be backsubstituted in the first of equation (3) to recover $\mathbf{f}$. In the MP literature, equation (8) is called a *range-space* equation, and is recommended when $m \ll n$. The formally equivalent technique in matrix structural analysis is the displacement method.

To eliminate $\mathbf{u}$ in favor of $\mathbf{f}$ the first matrix equation is premultiplied by $\mathbf{B}^T$, which upon taking equations (5) into account yields

$$\begin{bmatrix} -\mathbf{B}^T\mathbf{C} \\ \mathbf{A}^T \end{bmatrix} \mathbf{f} = \begin{Bmatrix} \mathbf{B}^T\mathbf{b} \\ \mathbf{p} \end{Bmatrix}. \tag{10}$$

This is the *tableau form* of the QP problem. The coefficient matrix of equation (10) is square but unsymmetric. A symmetric form can be obtained by observing that $\mathbf{f}' = \mathbf{S}\mathbf{p}$ and $\mathbf{f}'' = \mathbf{B}\mathbf{x}$, where $\mathbf{x}$ is a new unknown vector of dimension $n-m$,

[8]Department of Mechanical Engineering and Center for Space Structures and Controls, University of Colorado, Boulder, CO 80309-0427.

Manuscript received by ASME Applied Mechanics Division, March 10, 1986.

satisfy the constraint equation $\mathbf{A}^T\mathbf{f} = \mathbf{p}$. (These are sometimes called the particular and homogeneous solutions.) Insertion of

$$\mathbf{f} = \mathbf{f}' + \mathbf{f}'' = \mathbf{Bx} + \mathbf{Sp} \tag{11}$$

into equation (10) yields a new linear system in $\mathbf{x}$:

$$\mathbf{Fx} = -\mathbf{B}^T(\mathbf{b} + \mathbf{CSp}) = -\mathbf{v}. \tag{12}$$

In the MP literature, equation (12) is called the *null-space* or *range-reduced* system and $\mathbf{v}$ the *reduced gradient vector*. In structural analysis, equation (12) is the redundant-force or flexibility method. The coefficient matrix $\mathbf{F} = \mathbf{B}^T\mathbf{CB}$—which made its debut in equation (7)—is square, symmetric, and of order $n - m$. Once $\mathbf{f}$ is determined by solving either equation (10) or equations (11)-(12), the Lagrangian multipliers may be recovered from

$$\mathbf{u} = \mathbf{S}^T(\mathbf{b} + \mathbf{Cf}). \tag{13}$$

Expressing these solutions in terms of the original data of course reproduces equation (6).

The Structural Problem

Table 1 gives the mechanical meaning of the symbols used in equations (1) through (13). The key equations are now reinterpreted for the structural analysis problem.

Mixed Method. System (3), which combines force and displacement unknowns, is the mixed method of analysis (also called the combined method). The first matrix equation expresses compatibility of deformations whereas the second one expresses discrete force equilibrium.

This system may be derived by variational or physical arguments. Variational arguments are usually based on finite element discretizations of the Hellinger-Prange-Reissner principle, which leads directly to a symmetric form. Physical or nonvariational arguments generally lead to unsymmetric forms which may be symmetrized through appropriate transformation of the unknowns.

Displacement Method. Equation (9) is formally the stiffness equation of the displacement method. However, programs based on the DSM solve a similar equation

$$\mathbf{K}_u\mathbf{u} = \mathbf{f}_u, \tag{14}$$

where $\mathbf{K}_u$ is formed directly by assembling the "singular" element stiffness matrices. Displacement boundary conditions are applied *after* assembling $\mathbf{K}_u$. For conventional elements (e.g., bar, beams) $\mathbf{K} \equiv \mathbf{K}_u$, but for complicated 2D and 3D mixed elements $\mathbf{K}$ has wider connectivity than $\mathbf{K}_u$ so the latter is preferable. When finite elements are formulated by a mixed or stress-hybrid formulation, the DSM may be viewed as a condensation of the force (or stress) unknowns via equation (8) at the *element* level rather than at the systems level.

Table 1 Meaning of symbols in matrix structural analysis

Symbol	Interpretation
$\mathbf{A}$	Connection matrix (alias: leverage)
$\mathbf{A}^T$	Equilibrium matrix
$\mathbf{B}$	Self-stress matrix (alias: prestress, self-strain)
$\mathbf{B}^T$	Compatibility matrix
$\mathbf{b}$	Initial stress vector (alias: prestress, initial-strain, prescribed-displacement)
$\mathbf{C}$	Compliance matrix
$\mathbf{F}$	Flexibility matrix
$\mathbf{f}$	Unknown internal forces (or stresses)
$\mathbf{K}$	Mixed-method stiffness matrix
n	Number of force unknowns
m	Number of displacement unknowns
$\mathbf{p}$	Prescribed forces
$\mathbf{p}'$	Pseudo forces
$\mathbf{u}$	Unknown displacements
$\mathbf{x}$	Unknown redundant forces
S,H,T,U,v	No particular name

Force Method. Systems (10) and (12) furnish alternative expressions of the force method. In either case the crucial step is the determination of the self-stress matrix $\mathbf{B}$. Because of the arbitrariness in the choice of $\mathbf{V}$ in equation (4), the matrix is far from unique. A good procedure for computing $\mathbf{B}$ should take advantage of any sparseness present in the connection matrix $\mathbf{A}$ while maintaining numerical stability.

The best computer implementations of the force method in the Sixties were based on Gauss-Jordan factorization with pivoting, a procedure labelled "the rank technique" by Robinson (1973). Recent research has focused on triangular and orthogonal factorization schemes; in particular the "turnback" LU decomposition introduced by Kaneko et al. (1982) and variants thereof. Heath et al. (1984) give an excellent survey of candidate methods, but the problem remains open.

Once a $\mathbf{B}$ is available, $\mathbf{f}$ may be determined by solving the tableau system (10) or the range-reduced system (11)-(12). Patnaik (1986a, 1986b) prefers the former approach while Kaneko (1982), Heath et al. (1984), and Kaneko and Plemmons (1984) follow the latter.

Free Vibrations

Patnaik (1986b) notes that "the standard force method (meaning the range-reduced system (12)) has not been developed for vibration and buckling analysis." This statement is examined below for the free-vibration eigenproblem and shown to be justified. Set $\mathbf{b} = \mathbf{0}$ and $\mathbf{p} = \omega^2\mathbf{Mu}$, where ω is the circular frequency and $\mathbf{M}$ a structural mass matrix. Substitution into equations (10) and (13) yields

$$\begin{bmatrix} -\mathbf{B}^T\mathbf{C} \\ \mathbf{A}^T \end{bmatrix} \mathbf{f} = \omega^2 \begin{bmatrix} \mathbf{0} \\ \mathbf{MS}^T\mathbf{C} \end{bmatrix} \mathbf{f}. \tag{15}$$

This is an unsymmetric eigenproblem of order n which has $n - m$ improper roots. To symmetrize it, try the transformation $\mathbf{f} = \mathbf{Dz}$ and attempt to satisfy the first matrix equation in equation (15). A short calculation gives $\mathbf{D} = \mathbf{C}^{-1}\mathbf{A}$, which is the inverse of equation (13) when $\mathbf{b} = \mathbf{0}$. Therefore, $\mathbf{z} \equiv \mathbf{u}$ and we obtain

$$\mathbf{A}^T\mathbf{C}^{-1}\mathbf{Au} = \mathbf{Ku} = \omega^2\mathbf{Mu}. \tag{16}$$

This is the standard mass-stiffness eigenproblem of order m. Of course, DSM-based programs would use $\mathbf{K}_u$.

Eigensystems (15) and (16) correspond to equations (10) and (9), respectively. Is there a flexibility counterpart to equation (12), namely

$$\mathbf{Fx} = \omega^2\mathbf{Gx} \tag{17}$$

where $\mathbf{G}$ is symmetric and of order $n - m$? The question is nontrivial because $\mathbf{p}$ is no longer a given vector, which makes the range decomposition (11) frequency dependent:

$$\mathbf{f} = (\mathbf{I} - \omega^2\mathbf{SMS}^T\mathbf{C})^{-1}\mathbf{Bx} = [\mathbf{I} + \omega^2\mathbf{SMS}^T\mathbf{C} + \omega^4(\mathbf{SMS}^T\mathbf{C})^2 + \ldots]\mathbf{Bx}. \tag{18}$$

Substituting into $\mathbf{B}^T\mathbf{Cf} = \mathbf{0}$ we find that eigensystem (17), with $\mathbf{G} = -\mathbf{B}^T\mathbf{CSMS}^T\mathbf{CB}$, appears as the first two terms of an infinite eigenmatrix expansion in ascending powers of ω^2. The complexity of the matrices is unlikely to encourage practical applications. The buckling eigenproblem can be analyzed in a similar manner.

Concluding Remarks

A cautious answer to the title question would be: perhaps, but not as a general-purpose tool. Even after the arduous task of constructing a "good" self-stress matrix $\mathbf{B}$ is completed, we are not done. In linear stress analysis there remains the problem of solving either equation (10) or equations (11)-(12). In vibration analysis, the force based unsymmetric eigen-

problem (15) is computationally inferior to equation (16). As for relative matrix sizes, note that for two and three-dimensional continua discretized with mixed finite elements, $n \approx 1.5m$ and $n \approx 2m$, respectively.

But the case noted in the Introduction offers an applications niche. If the matrix equilibrium equation $\mathbf{A}^T\mathbf{f} = \mathbf{p}$ does not change as the sequence of related problems is solved, the self-stress matrix $\mathbf{B}$ can be reused. In structural optimization this happens if topology and element shapes are not design variables. Moreover, as the internal forces provided directly by the force method often have engineering significance, the calculation of design-sensitivity derivatives is simplified (Patnaik and Gallagher, 1986c).

The processing of very large systems, say $n > 10000$, calls for further research and numerical experimentation in substructuring and iterative methods. A comparative study of system-decomposition procedures presently favored in large-scale constrained optimization may be profitable regarding these two areas.

References

Fletcher, R., 1981, *Practical Methods of Optimization: Vol. 2, Constrained Optimization*, Wiley, New York.

Fraeijs de Veubeke, B. M., 1965, "Displacement and Equilibrium Models in the Finite Element Method," *Stress Analysis*, Zienkiewicz, O. C., and Hollister, G., eds., Wiley, London, Chapter 9.

Gill, P. E., Murray, W., and Wright, M. H., 1982, *Practical Optimization*, Academic Press, New York.

Heath, M. T., Plemmons, R. J., and Ward, R. C., 1984, "Sparse Orthogonal Schemes for Structural Optimization Using the Force Method," *SIAM J. Sci. Stat. Comput.*, Vol. 5, pp. 514–532.

Kaneko, I., Lawo, M., and Thierauf, G., 1982, "On Computational Procedures for the Force Method," *Int. J. Numer. Meth. Engrg.*, Vol. 18, pp. 1469–1495.

Kaneko, I., and Plemmons, R. J., 1984, "Minimum Norm Solutions to Linear Elastic Analysis Problems," *Int. J. Numer. Meth. Engrg.*, Vol. 20, pp. 983–998.

Patnaik, S. N., 1986a, "The Variational Formulation of the Integrated Force Method," *AIAA J.*, Vol. 24, pp. 129–137.

Patnaik, S. N., 1986b, "The Integrated Force Method Versus the Standard Force Method," *Computers and Structures*, Vol. 22, pp. 151–164.

Patnaik, S. N., and Gallagher, R. H., 1986c, "Gradients of Behavior Constraints and Reanalysis via the Integrated Force Method," *Int. J. Numer. Meth. Engrg.*, Vol. 23, pp. 2205–2212.

Robinson, J., 1973, *Integrated Theory of Finite Element Methods*, Wiley, New York.

Turner, M. J., Clough, R. W., Martin, H. C., and Topp, L. J., 1956, "Stiffness and Deflection Analysis of Complex Structures," *J. Aeron. Sci.*, Vol. 23, pp. 805–823.

Turner, M. J., 1959, "The Direct Stiffness Method of Structural Analysis," AGARD Meeting, Aachen, Germany.

Penalty Spring Stabilization of Singular Jacobians

C. A. Felippa[9]

Problem Description

Consider the system of nonlinear equations

$$\mathbf{r}(\mathbf{u},\lambda) = \mathbf{0}, \tag{1}$$

where $\mathbf{u}$ is the state vector of n unknowns, $\mathbf{r}$ is a vector of n residual components, and λ is a scalar control parameter. These equations are to be solved for varying values of λ. The additional equation that makes equation (1) determinate is the single constraint

$$c(\mathbf{u},\lambda) = 0. \tag{2}$$

Equations of the type (1)–(2) arise in many applications. For example, in nonlinear structural analysis by the finite element method, $\mathbf{u}$ collects the displacement degrees of freedom, $\mathbf{r}$ collects residual forces, λ is a loading parameter, and equation (2) is the "arclength" constraint.

Newton with Augmented Jacobian

We consider the solution of equations (1)–(2) in the context of a *continuation* procedure. Suppose that a solution $\mathbf{u}^*$, λ^* is available or predicted from a previous step. Starting with the initial approximation $\mathbf{u}^0 = \mathbf{u}^*$, $\lambda^0 = \lambda^*$, the conventional Newton method applied to equations (1)–(2) generates a sequence of iterates $\mathbf{u}^k$, λ^k, where $k = 0, 1, \ldots$ is an iteration step index. The corrections to the state and control variables

$$\mathbf{d} = \mathbf{u}^{k+1} - \mathbf{u}^k, \quad \eta = \lambda^{k+1} - \lambda^k, \tag{3}$$

are obtained by solving the linear algebraic system

$$\begin{bmatrix} \mathbf{K} & -\mathbf{q} \\ \mathbf{a}^T & b \end{bmatrix} \begin{Bmatrix} \mathbf{d} \\ \eta \end{Bmatrix} = - \begin{Bmatrix} \mathbf{r} \\ c \end{Bmatrix}, \tag{4}$$

where

$$\mathbf{K} = \frac{\partial \mathbf{r}}{\partial \mathbf{u}}, \quad \mathbf{q} = -\frac{\partial \mathbf{r}}{\partial \lambda}, \quad \mathbf{a}^T = \frac{\partial c}{\partial \mathbf{u}}, \quad b = \frac{\partial c}{\partial \lambda}, \tag{5}$$

and all known quantities are evaluated at $\mathbf{u}^k$, λ^k. (The negative sign for $\partial\mathbf{r}/\partial\lambda$ is chosen because $\mathbf{q}$ is conventionally the incremental loading vector in mechanical problems.)

The $n \times n$ matrix $\mathbf{K}$ is the Jacobian matrix of equation (1); also called the tangent stiffness matrix in structural mechanics. The coefficient matrix of equation (4) is called the augmented Jacobian. In many applications $\mathbf{K}$ is symmetric and sparse and it is of interest to solve the linear system (4) using techniques that preserve those attributes. The procedures described below make use of auxiliary systems of equations to achieve that goal. The number of auxiliary systems depends on whether the Jacobian $\mathbf{K}$ is singular (critical points) or nonsingular (regular points). Both cases are treated in the sequel.

Regular Points

Regular points of equations (1)–(2) are solutions at which the Jacobian matrix $\mathbf{K}$ is nonsingular. If so, performing forward Gauss elimination on equation (4) to get rid of $\mathbf{d}$ produces the scalar equation

$$(b + \mathbf{a}^T\mathbf{K}^{-1}\mathbf{q})\eta = -(c + \mathbf{a}^T\mathbf{K}^{-1}\mathbf{r}). \tag{6}$$

Let $\mathbf{d}_q$ and $\mathbf{d}_r$ denote the solution of the auxiliary symmetric linear systems

$$\mathbf{K}\mathbf{d}_r = -\mathbf{r}, \quad \mathbf{K}\mathbf{d}_q = \mathbf{q}. \tag{7}$$

Then

$$\eta = -(c + \mathbf{a}^T\mathbf{d}_r)/(b + \mathbf{a}^T\mathbf{d}_q), \quad \mathbf{d} = \mathbf{d}_r + \eta\mathbf{d}_q. \tag{8}$$

Two right-hand sides have to be generally solved at each Newton step. The number reduces to one for $k > 1$, however, if modified Newton is used so that $\mathbf{K}$ is held fixed for several iteration steps *and* $\mathbf{q}$ does not vary. The latter assumption holds in structural mechanics applications if the loading is conservative and proportional.

Limit Points

Limit points are characterized by the condition

$$\mathbf{K}\mathbf{z} = 0, \quad \mathbf{z}^T\mathbf{q} \neq 0, \tag{9}$$

where $\mathbf{z}$ is a null eigenvector (the "collapse mode" in structural mechanics) of $\mathbf{K}$. At limit points the Jacobian matrix is singular but the condition $\mathbf{z}^T\mathbf{q} \neq 0$ ensures that the augmented Jacobian is not. Although the detection of exact singularity may be masked by roundoff in actual computations, indirect

[9]Professor, Department of Mechanical Engineering and Center for Space Structures and Controls, University of Colorado, Boulder, CO 80309-0427.

Manuscript received by ASME Applied Mechanics Division, April 9, 1986.

problem (15) is computationally inferior to equation (16). As for relative matrix sizes, note that for two and three-dimensional continua discretized with mixed finite elements, $n \approx 1.5m$ and $n \approx 2m$, respectively.

But the case noted in the Introduction offers an applications niche. If the matrix equilibrium equation $\mathbf{A}^T\mathbf{f} = \mathbf{p}$ does not change as the sequence of related problems is solved, the self-stress matrix $\mathbf{B}$ can be reused. In structural optimization this happens if topology and element shapes are not design variables. Moreover, as the internal forces provided directly by the force method often have engineering significance, the calculation of design-sensitivity derivatives is simplified (Patnaik and Gallagher, 1986c).

The processing of very large systems, say $n > 10000$, calls for further research and numerical experimentation in substructuring and iterative methods. A comparative study of system-decomposition procedures presently favored in large-scale constrained optimization may be profitable regarding these two areas.

References

Fletcher, R., 1981, *Practical Methods of Optimization: Vol. 2, Constrained Optimization*, Wiley, New York.

Fraeijs de Veubeke, B. M., 1965, "Displacement and Equilibrium Models in the Finite Element Method," *Stress Analysis*, Zienkiewicz, O. C., and Hollister, G., eds., Wiley, London, Chapter 9.

Gill, P. E., Murray, W., and Wright, M. H., 1982, *Practical Optimization*, Academic Press, New York.

Heath, M. T., Plemmons, R. J., and Ward, R. C., 1984, "Sparse Orthogonal Schemes for Structural Optimization Using the Force Method," *SIAM J. Sci. Stat. Comput.*, Vol. 5, pp. 514–532.

Kaneko, I., Lawo, M., and Thierauf, G., 1982, "On Computational Procedures for the Force Method," *Int. J. Numer. Meth. Engrg.*, Vol. 18, pp. 1469–1495.

Kaneko, I., and Plemmons, R. J., 1984, "Minimum Norm Solutions to Linear Elastic Analysis Problems," *Int. J. Numer. Meth. Engrg.*, Vol. 20, pp. 983–998.

Patnaik, S. N., 1986a, "The Variational Formulation of the Integrated Force Method," *AIAA J.*, Vol. 24, pp. 129–137.

Patnaik, S. N., 1986b, "The Integrated Force Method Versus the Standard Force Method," *Computers and Structures*, Vol. 22, pp. 151–164.

Patnaik, S. N., and Gallagher, R. H., 1986c, "Gradients of Behavior Constraints and Reanalysis via the Integrated Force Method," *Int. J. Numer. Meth. Engrg.*, Vol. 23, pp. 2205–2212.

Robinson, J., 1973, *Integrated Theory of Finite Element Methods*, Wiley, New York.

Turner, M. J., Clough, R. W., Martin, H. C., and Topp, L. J., 1956, "Stiffness and Deflection Analysis of Complex Structures," *J. Aeron. Sci.*, Vol. 23, pp. 805–823.

Turner, M. J., 1959, "The Direct Stiffness Method of Structural Analysis," AGARD Meeting, Aachen, Germany.

Penalty Spring Stabilization of Singular Jacobians

C. A. Felippa[9]

Problem Description

Consider the system of nonlinear equations

$$\mathbf{r}(\mathbf{u},\lambda) = \mathbf{0}, \tag{1}$$

where $\mathbf{u}$ is the state vector of n unknowns, $\mathbf{r}$ is a vector of n residual components, and λ is a scalar control parameter. These equations are to be solved for varying values of λ. The additional equation that makes equation (1) determinate is the single constraint

$$c(\mathbf{u},\lambda) = 0. \tag{2}$$

Equations of the type (1)–(2) arise in many applications. For example, in nonlinear structural analysis by the finite element method, $\mathbf{u}$ collects the displacement degrees of freedom, $\mathbf{r}$ collects residual forces, λ is a loading parameter, and equation (2) is the "arclength" constraint.

Newton with Augmented Jacobian

We consider the solution of equations (1)–(2) in the context of a *continuation* procedure. Suppose that a solution $\mathbf{u}^*$, λ^* is available or predicted from a previous step. Starting with the initial approximation $\mathbf{u}^0 = \mathbf{u}^*$, $\lambda^0 = \lambda^*$, the conventional Newton method applied to equations (1)–(2) generates a sequence of iterates $\mathbf{u}^k$, λ^k, where $k = 0, 1, \ldots$ is an iteration step index. The corrections to the state and control variables

$$\mathbf{d} = \mathbf{u}^{k+1} - \mathbf{u}^k, \quad \eta = \lambda^{k+1} - \lambda^k, \tag{3}$$

are obtained by solving the linear algebraic system

$$\begin{bmatrix} \mathbf{K} & -\mathbf{q} \\ \mathbf{a}^T & b \end{bmatrix} \begin{Bmatrix} \mathbf{d} \\ \eta \end{Bmatrix} = - \begin{Bmatrix} \mathbf{r} \\ c \end{Bmatrix}, \tag{4}$$

where

$$\mathbf{K} = \frac{\partial \mathbf{r}}{\partial \mathbf{u}}, \quad \mathbf{q} = -\frac{\partial \mathbf{r}}{\partial \lambda}, \quad \mathbf{a}^T = \frac{\partial c}{\partial \mathbf{u}}, \quad b = \frac{\partial c}{\partial \lambda}, \tag{5}$$

and all known quantities are evaluated at $\mathbf{u}^k$, λ^k. (The negative sign for $\partial\mathbf{r}/\partial\lambda$ is chosen because $\mathbf{q}$ is conventionally the incremental loading vector in mechanical problems.)

The $n \times n$ matrix $\mathbf{K}$ is the Jacobian matrix of equation (1); also called the tangent stiffness matrix in structural mechanics. The coefficient matrix of equation (4) is called the augmented Jacobian. In many applications $\mathbf{K}$ is symmetric and sparse and it is of interest to solve the linear system (4) using techniques that preserve those attributes. The procedures described below make use of auxiliary systems of equations to achieve that goal. The number of auxiliary systems depends on whether the Jacobian $\mathbf{K}$ is singular (critical points) or nonsingular (regular points). Both cases are treated in the sequel.

Regular Points

Regular points of equations (1)–(2) are solutions at which the Jacobian matrix $\mathbf{K}$ is nonsingular. If so, performing forward Gauss elimination on equation (4) to get rid of $\mathbf{d}$ produces the scalar equation

$$(b + \mathbf{a}^T\mathbf{K}^{-1}\mathbf{q})\eta = -(c + \mathbf{a}^T\mathbf{K}^{-1}\mathbf{r}). \tag{6}$$

Let $\mathbf{d}_q$ and $\mathbf{d}_r$ denote the solution of the auxiliary symmetric linear systems

$$\mathbf{K}\mathbf{d}_r = -\mathbf{r}, \quad \mathbf{K}\mathbf{d}_q = \mathbf{q}. \tag{7}$$

Then

$$\eta = -(c + \mathbf{a}^T\mathbf{d}_r)/(b + \mathbf{a}^T\mathbf{d}_q), \quad \mathbf{d} = \mathbf{d}_r + \eta\mathbf{d}_q. \tag{8}$$

Two right-hand sides have to be generally solved at each Newton step. The number reduces to one for $k > 1$, however, if modified Newton is used so that $\mathbf{K}$ is held fixed for several iteration steps *and* $\mathbf{q}$ does not vary. The latter assumption holds in structural mechanics applications if the loading is conservative and proportional.

Limit Points

Limit points are characterized by the condition

$$\mathbf{K}\mathbf{z} = 0, \quad \mathbf{z}^T\mathbf{q} \neq 0, \tag{9}$$

where $\mathbf{z}$ is a null eigenvector (the "collapse mode" in structural mechanics) of $\mathbf{K}$. At limit points the Jacobian matrix is singular but the condition $\mathbf{z}^T\mathbf{q} \neq 0$ ensures that the augmented Jacobian is not. Although the detection of exact singularity may be masked by roundoff in actual computations, indirect

[9]Professor, Department of Mechanical Engineering and Center for Space Structures and Controls, University of Colorado, Boulder, CO 80309-0427.

Manuscript received by ASME Applied Mechanics Division, April 9, 1986.

effects may be felt in the form of ill-conditioning when $\mathbf{K}$ is factored to solve the linear systems (7).

This problem can be stably solved by combining the solution of three auxiliary systems. The coefficient matrix of these systems is rendered nonsingular by adding a fictitious "penalty" spring stiffness s to the ith equation:

$$(\mathbf{K}+s\mathbf{E}_i)\mathbf{d}_r=-\mathbf{r},\quad (\mathbf{K}+s\mathbf{E}_i)\mathbf{d}_q=\mathbf{q},\quad (\mathbf{K}+s\mathbf{E}_i)\mathbf{d}_e=s\mathbf{e}_i, \tag{10}$$

where $\mathbf{e}_i$ is the elementary vector of order n, all of whose entries are zero except the ith entry which is one, and $\mathbf{E}_i = \mathbf{e}_i\mathbf{e}_i^T$. The choice of i is discussed below. As the only modification to $\mathbf{K}$ is on a diagonal element, its sparseness and symmetry are not affected. Let us express the solution of equation (4) as

$$\mathbf{d}=\mathbf{d}_r+\alpha\mathbf{d}_q+\beta\mathbf{d}_e,\quad \eta=-(c+\mathbf{a}^T\mathbf{d})/b. \tag{11}$$

Solving for η from the second equation assumes that $b \neq 0$, a restriction to be removed shortly. Coefficient α is dimensionless whereas β has the physical dimensions of $\mathbf{d}$. Inserting (11) into the first matrix equation of (4) yields the right-hand side

$$\mathbf{K}\mathbf{d}-\mathbf{q}\eta=-\mathbf{r}+\left[\alpha+\frac{\mathbf{a}^T\mathbf{d}+c}{b}\right]\mathbf{q}+s[\beta-\mathbf{e}_i^T\mathbf{d}]\mathbf{e}_i. \tag{12}$$

The values of α and β can be determined by requiring that the expressions in square brackets vanish. Defining the six inner products

$$a_r=\mathbf{a}^T\mathbf{d}_r,\quad a_q=\mathbf{a}^T\mathbf{d}_q,\quad a_e=\mathbf{a}^T\mathbf{d}_e,$$
$$d_{ri}=\mathbf{e}_i^T\mathbf{d}_r,\quad d_{qi}=\mathbf{e}_i^T\mathbf{d}_q,\quad d_{ei}=\mathbf{e}_i^T\mathbf{d}_e, \tag{13}$$

the following 2×2 unsymmetric linear system results:

$$\begin{bmatrix} b+a_q & a_e \\ -d_{qi} & 1-d_{ei}\end{bmatrix}\begin{Bmatrix}\alpha\\ \beta\end{Bmatrix}=\begin{Bmatrix}-c-a_r\\ d_{ri}\end{Bmatrix}. \tag{14}$$

The first equation in (14) has been multiplied through by b to avoid breakdown for $b = 0$. Once α and β are computed, the increment $\mathbf{d}$ is formed from the first of equations (11). The value of the fictitious spring stiffness s enters equation (14) only indirectly through d_r, d_q, and d_e. It is recommended that s be taken as large as possible while guarding against overflow; a detailed error analysis of this problem is given in Felippa (1977).

Equation (10) shows that three right-hand sides, $\mathbf{r}$, $\mathbf{q}$, and $s\mathbf{e}_i$, have to be solved at each step in the general case. But if modified Newton is used and $\mathbf{q}$ is constant, only one right-hand side ($\mathbf{r}$) has to be processed for $k > 1$, assuming of course that i remains fixed.

The vanishing of $\mathbf{a}$ because $c = c(\lambda)$ causes no difficulty. On the other hand, if b vanishes because $c = c(\mathbf{u})$ some care must be taken as the calculation of η from the second of equations (11) breaks down. It is then preferable to use the Rayleigh-quotient-like formula

$$\eta=\frac{\mathbf{d}^T\mathbf{K}\mathbf{d}+\mathbf{d}^T\mathbf{r}}{\mathbf{d}^T\mathbf{q}}, \tag{15}$$

which results on premultiplying the first of equation (4) by $\mathbf{d}^T$. The expensive matrix-vector product $\mathbf{K}\mathbf{d}$ can be circumvented by using an equivalent expression that involves only vector operations:

$$\mathbf{K}\mathbf{d}=-\mathbf{r}+\alpha\mathbf{q}+\beta s\mathbf{e}_i-s\mathbf{E}_i\mathbf{d}. \tag{16}$$

Substituting this expression into equation (15) yields

$$\eta=\alpha+s\,\frac{\beta d_i-d_i^2}{\mathbf{d}^T\mathbf{q}} \tag{17}$$

in which $d_i = \mathbf{e}_i^T\mathbf{d}$. Recovery of η from this equation is recommended even if b is nonzero.

Choosing *i*

For choosing i several strategies are possible, and only one is mentioned here. Keep an index list of the m components of $\mathbf{u}$ that have varied most rapidly in previous iterations; for example $m \approx n/100$. (This is a disguised way of monitoring the fundamental eigenvector of the Jacobian.) If ill-conditioning is detected at equation j when factoring $\mathbf{K}$ to solve the auxiliary systems (7), backtract to the closest largest-change index less than j, and switch to the limit-point procedure. This strategy tries to save most of the factorization work since the factors of $\mathbf{K}$ and $\mathbf{K} + s\mathbf{E}_i$ are the same up to the ith equation.

Bifurcation Points

A bifurcation point of equation (1) is characterized by a variant of equations (9) in which the second condition is $\mathbf{z}^T\mathbf{q} = 0$, this $\mathbf{z}$ is called the buckling mode. Both $\mathbf{K}$ and the augmented Jacobian are singular, and the second row of the coefficient matrix in equation (14) vanishes. The occurrence of *simple* bifurcation (the case in which the buckling mode is unique) can be detected by monitoring changes of sign of the determinant of the augmented Jacobian. A simple bifurcation point can be stably treated by solving a 3×3 unsymmetric system that incorporates an estimate of $\mathbf{z}$. The procedure is discussed in Felippa (1987). The multiple bifurcation case remains a frontier research topic.

Concluding Remarks

The limit-point treatment presented here is a synthesis of two earlier techniques. To traverse limit points, Sharifi and Popov (1970) introduced fictitious stiffnesses in the form of rank-one updates but did not use auxiliary systems. Rheinboldt (1981) discussed the use of auxiliary systems in conjunction with the partition of the Jacobian matrix. The procedure presented here avoids the need for partitioning and consequent special treatment of the elements in the ith row and column of $\mathbf{K}$; only the diagonal entry is changed. It enjoys the added advantage that none of the auxiliary system right-hand sides requires access to elements of the $\mathbf{K}$ matrix. This would be computationally inconvenient should $\mathbf{K}$ be stored by blocks on peripheral storage, or if its elements are not explicitly available as in frontal solution schemes. The connection between the partitioned-equation and penalty spring treatments may be established through the Sherman-Morrison modified-inverse formula.

Extension to several penalty springs placed on additional equations follows a similar procedure. If m springs are introduced, $m + 2$ right-hand sides (m of which are elementary vectors) have to be solved for at each iteration, although a considerable reduction of work is possible if modified Newton is used. Such an "overkill" stabilization may be called for if $\mathbf{K}$ becomes highly rank deficient. Two examples: flow-like behavior of a structure undergoing deep plastification or high-temperature creep, and cable structures with initially unstressed members.

References

Felippa, C. A., 1977, "Error Analysis of Penalty Function Techniques for Constraint Definition in Linear Algebraic Systems," *Int. J. Numer. Meth. Engrg.*, Vol. 11, pp. 709–728.

Felippa, C. A., 1987, "Traversing Critical Points," to be presented at NUMETA 1987 Conference, Swansea, UK; Proceedings to be published by Pineridge Press.

Rheinboldt, W. C., 1981, "Numerical Analysis of Continuation Methods for Nonlinear Structural Problems," *Computers and Structures*, Vol. 13, pp. 103–113.

Sharifi, P., and Popov, E. P., 1970, "Nonlinear Buckling Analysis of Sandwich Arches," *ASCE J. Engrg. Div.*, Vol. 97, pp. 1397–1411.

On the Fracture of Pencil Points

H. Petroski[10]

Introduction

To ask how and why a pencil point breaks is essentially to ask the same fundamental questions that Galileo (1638) did in his seminal work on the strength of materials three and a half centuries ago. Yet the problem of the fracture of a pencil point seems to have a sparse literature. In a 1979 paper, Cronquist observed that broken-off conical pencil points always appear to be virtually identical in size and shape, and he presented an elementary strength-of-materials analysis to explain the phenomenon. Cronquist's observation was discussed in a popular vein by Walker (1979) and was extended by Cowin (1983). In his note on broken pencil points Cowin takes into account a more general loading than did Cronquist, but while still working within the context of strength of materials. These last three references appear to be the only literature explicitly on the problem of predicting the size and shape of broken-off pencil points. The purpose of the present note is to give some background on the problem, to explain an aspect of the fracture that remains unanswered, and to extend the analysis of the fracture of pencil points to a broader class of points.

The kinds of pencil points that have heretofore been treated have been the truncated circular cylindrical cones of the hardened mixture of graphite and clay that we find in common wood-cased pencils and that we are all familiar with from school days. The geometry and notation employed by Cowin (1983) is shown in Fig. 1. Cronquist (1979) treated only the case where the force components F and R combine to give a force transverse to the pencil axis, noting that an equal axial component of force would change his result by only about ten percent, whereas Cowin allowed F and R to be completely arbitrary. In both analyses, the normal stress across plane x = const was calculated using familiar strength of materials equations for axial and bending stresses in beams. The maximum value of this normal stress was found by Cronquist to occur where the ratio of the fracture diameter to the pencil point diameter at the writing tip is 3/2, and Cowin essentially confirmed this to be an average value under a broad range of loading conditions. Perhaps because the actual fracture surface is slanted and makes it difficult to measure the diameter where a conical pencil point fractures, Walker introduced the ratio N of slant length to tip diameter of the broken-off point. In terms of the parameters defined in Fig. 1,

$$N = [(x_{\max} - \ell)/\cos\alpha]/2\tan\alpha \qquad (1)$$

Walker collected data in a desk-top experiment and found reasonable agreement with a predicted value of N of approximately 2.5.

The results of Cronquist and Cowin are consistent with our experience that sharper points tend to break more easily and closer to the point, where the cross-sectional area is small. On the other hand, blunt pencil points, of the kinds children seem to prefer, are less prone to breaking. But when they do, under larger forces than children would normally exert in writing, large pieces of the point break off to give the appropriate value of N. (Mechanical pencils, which have essentially right-cylindrical lead points, will always have their lead break off where it enters the metal case, of course, because the maximum tensile stress increases linearly with distance from the writing surface. Thus, turning out too long a lead from our mechanical pencils brings us the same frustrations as making our wood-cased pencils too sharp.)

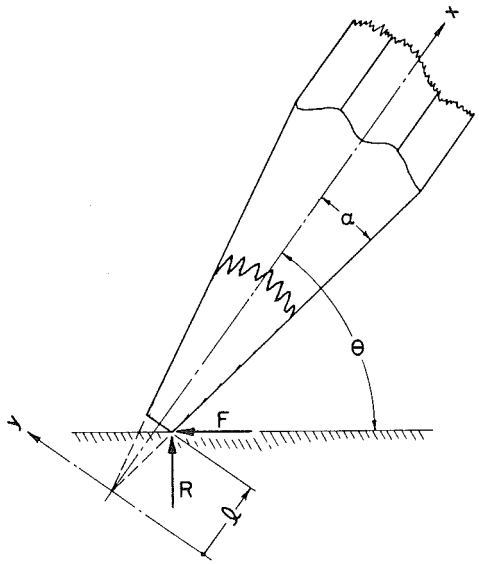


Fig. 1 Geometry of a pencil point (Cowin, 1983)

The characteristic of the broken-off pencil points that was inexplicable to Cronquist is that the fracture surface is not exactly perpendicular to the pencil axis. In fact, the fracture surface always slants back toward the pencil shaft as it grows from the edge of the pencil point closest to the writing surface. Cronquist believed the reason for this behavior to be beyond the reach of his simple analysis, and he did not pursue the point further. We shall show that the characteristic slant of the fracture surface is readily explained in the context of both Cronquist's and Cowin's analyses when one looks at the maximum principal stress in the pencil lead and not just the maximum axial tensile stress.

Some Historical Background

The problem of the resistance of a pencil point to fracture is essentially that of the strength of a cantilever beam loaded at its end, and this is precisely the problem that Galileo considered in the Second Day of his 1638 discourses on the strength of materials, which is generally agreed to be the work with which the history of the theory of elasticity and of the rational determination of the strength of materials properly begins (cf Todhunter and Pearson, 1886; Timoshenko, 1953).

While Galileo incorrectly assumed a uniform tensile stress at the root of the beam to resist the moment of the weight supported, he did correctly predict that the strength of a uniformly thick beam varies as the square of its depth. Galileo went on to argue that the profile of a beam of constant strength, or a "solid of equal resistance," would have a parabola as its generating curve. Such an optimized beam would be no more or no less likely to break at one location than at any other along its length. However, since the need to sharpen continually a pencil point to such an optimal shape would be more trouble than it was worth, pencils and the shape of their points have developed independent of Galileo's insights about optimization of strength.

The first "pencils" are believed to have been pieces of actual lead formed in convenient shapes. These were used to scribe guidelines such as engineering students used to do before lettering by hand their mechanical or structural drawings. The modern pencil had its origins in the discovery of a graphite mine in Borrowdale, England, in 1565. At first, prismatic pieces of solid graphite were cut for use as pencils, and later were encased in protective wood. The wood not only

[10]Department of Civil and Environmental Engineering, Duke University, Durham, N.C. 27706. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, January 27, 1987; final revision April 20, 1987.

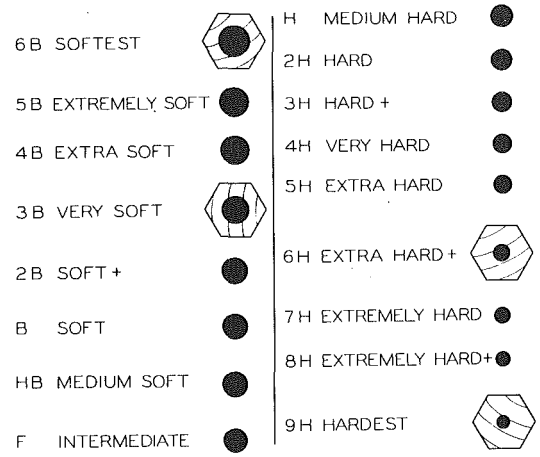


Fig. 2 Lead size as a function of hardness (Svensen and Street, 1962)

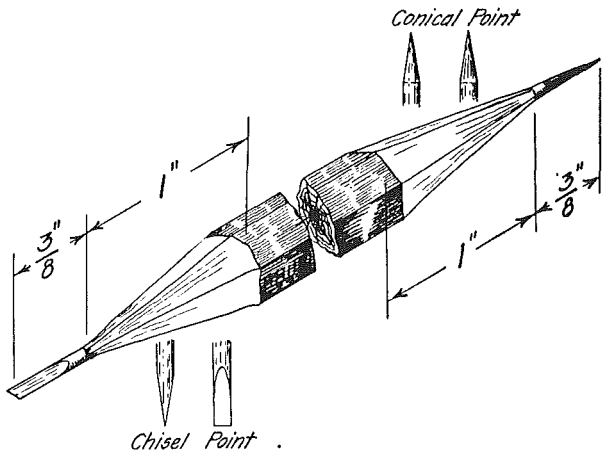


Fig. 3 Conical and wedge points (Kirby, 1925)

kept the writer's fingers clean, but also it strengthened the graphite prism so that pieces smaller in cross section could be used in pencils. When the graphite mine at Borrowdale was becoming worked out, alternatives were sought to using solid graphite for pencils. In 1790 Nicolas Conte, a French mechanic, and Josef Hardmuth of Vienna, perfected a process that enabled a mixture of pulverized graphite and clay to be used in the manufacture of pencils. This made it possible not only to use graphite dust instead of solid graphite, but also to make pencils of variable hardness. Soon a variety of pencils with circular cylindrical leads were being offered to writers, artists, and engineers (Fleming and Guptill, 1936). This variety is desirable to this day because paper, being composed of a weblike mass of interlaced fibers, exerts a file action on the pencil point. Especially in architectural drawing, where texture is so important, the rougher the paper being drawn on, the harder the pencil to be used (Halse, 1960). And in engineering drawing, of course, the hardness of the pencil is often matched to the fineness of the line being drawn, with harder pencils being able to take and hold a sharper point as long as they are not pressed too hard against the writing surface.

Problems of Strength

With a variety of products often comes a variety of manufacturing problems, however, and the resistance of pencil leads to breaking is a function of the mixture of graphite and clay that they contain. So different writing hardnesses of pencils meant that, all other things being equal, different pencils would break at different writing pressures, and this would

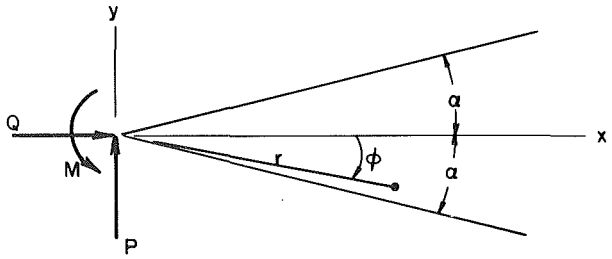


Fig. 4 Equivalent loading for wedge-pointed pencil

mean that those using the pencils would have to adjust their touch to the degree of the pencil in their hand. Furthermore, the pencil manufacturers would have to adjust their processes to take into account different strengths of pencil lead. One economical solution to the problem is to make the lead of different diameters in pencils of different hardnesses (and strengths) as discussed, e.g., by Svensen and Street (1962). Properly adjusting the lead diameter, a whole line of pencils could be marketed with essentially the same fracture strength. Fig. 2 shows how the lead diameters of a whole range of drafting pencils varies from hard to soft.

One of the disadvantages of a conventional wood-case pencil is that it must be constantly sharpened in order to produce a uniform line. The conical point is clearly an easy to one to make with a piece of sandpaper, and easier with a mechanical pencil sharpener. But, as we all know, and, as discussed above, as the analyses of Cronquist and Cowin confirm, the sharper we make a conical point, the easier it will break. The empirical evidence long ago taught draftsmen that they could gain advantages in strength by using another shape for their pencils: the wedge or chisel point. This point is illustrated in Fig. 3, and it too is easily formed with a sandpaper pad. According to Kirby (1925): "For mechanical drawing (line work) use a 6H pencil, sharpened to a chisel (wedge) point at one end and to a conical point at the other end. (See Fig. 3.) Use the chisel point in ruling lines, and the conical point in marking points, as in laying off distances."

The chisel point has the advantage that it keeps a more constant thickness when used to draw thick lines, such as in architectural rendering when the point is pulled along on its wide side, as well as to draw thin lines when the point is pulled along on its thin side. We shall analyze this pencil point shape to determine if it is indeed stronger than the conical shape analyzed by Cronquist and Cowin. As with their analyses, we shall assume the pencil point is geometrically ideal and without nicks or other imperfections. The propensity of a brittle pencil point to break when nicked with the sharpening knife was cautioned against in drafting text books (e.g., Hoelscher and Springer, 1956) long before fracture mechanics became current.

Stress Analysis of a Wedge-Pointed Pencil

Timoshenko (1956) has shown that for a transversely loaded cantilever having the form of a wedge, with half angle, α, as shown in Fig. 4 (with $Q = M = 0$), the strength of materials beam formula can be corrected by a factor β, where

$$\beta = \frac{4}{3} \frac{\tan^3 \alpha \cos^4 \alpha}{2\alpha - \sin 2\alpha} \tag{2}$$

to give the exact elasticity solution for the maximum normal stress, i.e.,

$$(\sigma_x)_{\max} = -\beta \frac{My}{I} \tag{3}$$

(Note that since the normal stress is of the order r^{-1}, the stresses would blow up at an infinitely sharp wedge-shaped

pencil point as soon as it touched the paper, and the point would be immediately broken. This is true for a conical point as well (cf Love, 1927).)

For a typical sharpened pencil point, $\alpha = 6$ deg, $\beta = 0.99$, and we see that the method of calculating the maximum normal stress σ_x is not critical in establishing the *location* of the maximum normal stress. However, the exact shear stress on a section perpendicular to the pencil axis is given by

$$\tau_{xy} = \frac{P}{bh}\,\frac{16y^2}{h^2}\,\frac{\tan^3\alpha\cos^4\phi}{2\alpha-\sin 2\alpha} \tag{4}$$

where b is the (constant) width and h the (varying) thickness of the section of interest. Thus the shear stress differs from that predicted by strength of materials analysis of a beam of uniform section in two important aspects: (a) The maximum value of the shear stress occurs at the top and bottom of the wedge, rather than at the neutral axis; and (b) the maximum value of the shear stress is numerically greater than the average by a factor of 3β:

$$(\tau_{xy})_{\max} = 3\beta\,\frac{P}{bh} \tag{5}$$

The implications of this unbeam-like behavior of the shear stress means that at the top and bottom of the wedge the plane normal to the x axis is not a principal plane and is threfore not a plane of absolute maximum normal stress.

Since both Cronquist and Cowin adopt a maximum normal stress failure criterion in interpreting their calculations for the cone, which suffer from the same inability to give a good approximation to the shear stress behavior at the cone boundary, their analysis cannot be expected to predict the *inclination* of the fracture plane. In fact, the actual initial fracture plane can be expected to be a plane perpendicular to the stress-free surface of the pencil point. That plane will always have the maximum normal stress, and illustrations in the papers of both Cronquist and Walker suggest that the fractures of their pencil points did indeed initiate across such principal planes.

Although beam theory is thus still sufficiently accurate to predict the *location* along the pencil axis where the normal stresses are maximum, we shall employ the elasticity solutions for a wedge loaded as shown in Fig. 4 to predict where the absolute maximum normal stress will occur in a wedge point. The elasticity problem of a wedge loaded at its tip by the point loads P and Q is a generalization of Michell's Problem, while the problem of a moment acting on the vertex of a wedge is Inglis' Problem (see Volterra and Gaines 1971), and these loads are statically equivalent to the loads Cowin considered to act on the truncated cone of Fig. 1 if we take

$$P = R\sin\theta - F\cos\theta \tag{6}$$

$$Q = R\cos\theta + F\sin\theta \tag{7}$$

$$M = \ell R(\tan\alpha\sin\theta + \cos\theta) - \ell F(\tan\alpha\cos\theta - \sin\theta) \tag{8}$$

According to Timoshenko and Goodier (1970), the normal and shear stresses across surface r = const in the wedge are given by

$$\sigma_r = \frac{-2P\cos\phi}{r(2\alpha+\sin 2\alpha)} + \frac{2Q\sin\phi}{r(2\alpha-\sin 2\alpha)} - \frac{2M\sin 2\phi}{r^2(\sin 2\alpha - 2\alpha\cos 2\alpha)} \tag{9a}$$

$$\tau_{r\phi} = \frac{M(\cos 2\phi - \cos 2\alpha)}{r^2(\sin 2\alpha - 2\alpha\cos 2\alpha)} \tag{9b}$$

The shear is clearly zero where $\phi = \alpha$, as the boundary conditions require, so that we can find the location $r_{\max}$ of the maximum *principal* stress along the edge $\phi = \alpha$ of the wedge point by the condition $\partial\sigma_r/\partial r = 0$. This procedure gives

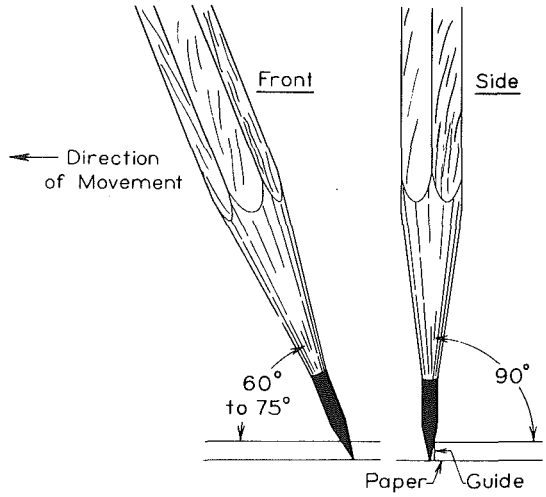


Fig. 5 Pencil inclined to draw a line (Svensen and Street, 1962)

$$r_{\max} = \frac{MA(\alpha)}{P\cos\alpha(2\alpha-\sin 2\alpha) - Q\sin\alpha(2\alpha+\sin 2\alpha)} \tag{10}$$

where

$$A(\alpha) = \frac{2\sin 2\alpha(\sin^2 2\alpha - 4\alpha^2)}{\sin 2\alpha - 2\alpha\cos 2\alpha} \tag{11}$$

For the case where the pencil is simply being pushed down on the writing surface, so that $F = 0$, we have

$$\frac{r_{\max}}{\ell} = \frac{A(\alpha)(\tan\alpha\sin\theta + \cos\theta)}{\sin\theta\cos\alpha(2\alpha-\sin 2\alpha) - \cos\theta\sin\alpha(2\alpha+\sin 2\alpha)} \tag{12}$$

For $\alpha = 6$ deg, this gives a positive value of $r_{\max}/\ell > 2$ for all $\theta < 88$ deg, and this corresponds to the parameter $N > 9.5$. For the standard drawing and drafting practice of exposing 3/8 in. of lead before sharpening (see, e.g., Halse, 1960, or Hoelscher and Springer, 1956), this would mean that a wedge point would have to be less than 0.04 in. thick to cause the maximum stresses to occur outside the wood case.

For the situation where the force $F \to \infty$ relative to R, we can take

$$P = -F\cos\theta \tag{13}$$

$$Q = F\sin\theta \tag{14}$$

$$M = \ell F(\sin\theta - \tan\alpha\cos\theta) \tag{15}$$

Inserting these values into equation (11) gives

$$\frac{r_{\max}}{\ell} = -\frac{A(\alpha)(\sin\theta - \tan\alpha\cos\theta)}{\cos\theta\cos\alpha(2\alpha-\sin 2\alpha) + \sin\theta\sin\alpha(2\alpha+\sin 2\alpha)} \tag{16}$$

For $\alpha = 6$ deg, this gives $r_{\max}/\ell > 0$ for $\theta > 6$ deg, and for $\theta > 30$ deg, this gives $r_{\max}/\ell > 1.5$, which corresponds to $N > 2.3$.

If we assume a pencil with a wedge-shaped point is being used at the recommended angle of 60 deg to 75 deg, as indicated in Fig. 5 (see, e.g., Svensen and Street, 1962, or Hoelscher and Springer, 1956), then the location of a possible fracture can be predicted more precisely. For 65 deg, for example, we have $1.87 < r_{\max}/\ell < 2.64$, which corresponds to a range of sizes for broken pencil points of:

$$4.1 < N_{\text{wedge}} < 7.8 \text{ (av.} = 5.95) \tag{17}$$

which compares with Cowin's

$$1.9 < N_{\text{cone}} < 7.0 \text{ (av.} = 4.45) \quad (18)$$

Since the size of the pencil tip, characterized by ℓ and α, determines the thickness or weight of the line drawn with the sharpened pencil, it is meaningful to compare the two results from equations (17) and (18). On the average, we can see that a given pencil, when overloaded in drawing a line of given weight, will have larger broken pencil points, as measured by the parameter N, when sharpened into a wedge than the same pencil sharpened into a conical point. Since the larger N, the greater the fracture area, we can also conclude that it would take on the average a greater effort to break a wedge-pointed pencil. Hence, not only does the wedge-pointed pencil have the advantage that it does not have to be twirled to keep line weight uniform, but also it can generally withstand a heavier hand on the part of the draftsperson.

Conclusion

The wood-case pencil is a common technological artifact whose size, shape, and composition have no doubt evolved with very little, if any, mathematical analysis. Equations flow from pencils, but pencils do not come of equations. Indeed, at one time during the nineteenth century, arguably the best pencils made in America were manufactured by John Thoreau and Co., after a process perfected by John Thoreau's son, Henry David, who is remembered neither as a mathematician nor as an engineer (Harding, 1965).

The use of the pencil by engineers and others, in particular the customary nature of the point with which drafting and sketching has traditionally been executed, also probably evolved more through trial and error and serendipity than through any deliberate, rational mathematical analysis. But, regardless of its origins and use, the pencil is as proper an object of analysis as is the natural world and universe. By asking why and how a pencil point breaks in the way it does, we are not only led to a better understanding of the tools of stress analysis and their limitations, but we are also led to a fuller appreciation of the wonders of technology when we analyze the aptness of such a manufactured product as the common pencil.

References

Cowin, S. C., 1983, "A Note on Broken Pencil Points," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 453-454.

Cronquist, D., 1979, "Broken-off Pencil Points," *American Journal of Physics*, Vol. 47, pp. 653-655.

Fleming, C. C., and Guptill, A. L., 1936, *The Pencil: Its History, Manufacture and Use*. Koh-I-Noor Pencil Co., Bloomsbury, N.J.

Galilei, G., 1638, *Dialogues Concerning Two New Sciences*, translated by H. Crew and A. de Salvio, reprinted 1954, Dover Publications, New York.

Halse, A. O., 1960, *Architectural Rendering: The Techniques of Contemporary Presentation*, McGraw-Hill Book Company, New York, Ch. 10.

Harding W., 1965, *The Days of Henry Thoreau*, Alfred A. Knopf, New York, pp. 56-57.

Hoelscher, R. P., and Springer, C. H., 1956, *Engineering Drawing and Geometry*, Wiley, New York, Ch. 3.

Kirby, R. S., 1925, *The Fundamentals of Mechanical Drawing*, Wiley, New York, Ch. 1.

Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, Fourth Edition, reprinted 1944, Dover Publications, New York, pp. 202-203.

Svensen, C. L., and Street, W. E., 1962, *Engineering Graphics*, D. Van Nostrand Company, Princeton, NJ, Ch. 2.

Timoshenko, S. P., 1953, *History of Strength of Materials: With a Brief Account of the History of Theory of Elasticity and Theory of Structures*, reprinted 1983, Dover Publications, New York.

Timoshenko, S., 1956, *Strength of Materials: Part II*, Advanced Theory and Problems, Third Edition, D. Van Nostrand Company, Princeton, NJ, pp. 60-64.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, Third Edition, McGraw-Hill Book Company, New York, pp. 109-113.

Todhunter, I., and Pearson, K., 1886, *A History of the Theory of Elasticity and of the Strength of Materials from Galilei to Lord Kelvin*, reprinted 1960, Dover Publications, New York.

Volterra, E., and Gaines, J. H., 1971, *Advanced Strength of Materials*, Prentice-Hall, Englewood Cliffs, NJ, pp. 163-169, pp. 175-177.

Walker, J., 1979, "The Amateur Scientist," *Scientific American*, Vol. 240, Feb., pp. 156-158. (See also Vol. 241, November, pp. 200-204.)

Inclusion Effects on Stress Measurement in Geological Materials

A. L. Florence[11]

Introduction

Stress measurement in geologic materials by embedded gauges is complicated because the gauge forms an inclusion with properties different from those of the surrounding medium. The sensing element measures the stress or pressure in the inclusion, which then has to be related to the far-field stress component we wish to measure. The analytical results presented here provide this relationship for elastic axisymmetric quasi-static stress fields.

Our analysis was motivated by a need to support our laboratory experiments to develop and apply miniature gauges for obtaining compressive stress measurements in rocks, rock simulants, and soils. As background, we describe briefly the design of the stress gauge shown in Fig. 1. It consists of a thin disk of elastic material (2.0 cm diameter and 0.2 cm thick) cast around a very thin flat foil (0.3 cm square and 0.002 cm thick) of piezoresistive material, such as ytterbium. The gauge is thin to allow the stress normal to the face to have the dominant effect on the inclusion stress field. This normal stress is the component the gauge is trying to measure. The sensing element is a piezoresistive conductor with a scalar response, consisting of a resistance change, to applied tensorial stress and strain (Gupta 1983, 1984). We, therefore, have to restrict the inclusion stress-strain state so the resistive change can be related to that state. Such states are generally hydrostatic or uniaxial strain and they are calibratable. The hydrostatic state is provided by a gauge of fluid-like material (low shear modulus) around the foil. The uniaxial strain state is provided by bonding the foil to a thin, relatively stiff, material, such as steel. Thus, we can relate the resistance change to the pressure in a fluid inclusion or to the inclusion stress component normal to the surface of the sensing foil. Our task is to relate the inclusion pressure or stress normal to the disk face to the medium far-field stress normal to the disk face.

Our analytical approach to the elastic inclusion problem is first to replace the thin disk with a thin oblate spheroid having the same aspect ratio, so that we can employ Eshelby's theory of inclusions (1958) to relate the inclusion and the far-field stress loading. We then examine the results to obtain the effects on this relationship of the loading stress ratio, the inclusion aspect ratio, and the elastic properties of the inclusion and medium. We assume that the very thin foil embedded in the inclusion has no effect on the stress distribution throughout most of the inclusion. In fact, the foil forms an additional inclusion problem. The analysis of an oblate spheroidal inclusion under far-field loading has been treated (Edwards, 1951; Shibata and Kanji Ono, 1978), but our results were derived in applicable form by Eshelby's method because we are employing the method in a more general context of gauge design.

Oblate Spheroidal Inclusion

We consider an oblate spheroidal inclusion at the origin of axes (x_1, x_2, x_3) and x_3 the axis of symmetry. Let the semiaxes have lengths a, b, and c such that $c < b = a$. The applied stress σ_{ij}^A is also symmetric about the x_3 axis, so $\sigma_{11}^A = \sigma_{22}^A$, and $\sigma_{ij}^A = 0$ $(i \neq j)$. Application of Eshelby's method leads to the inclusion stress formulas

[11]Engineering Mechanics Department, SRI International, Menlo Park, CA 94025. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, May 8, 1985; final revision April 7, 1987.

$$1.9 < N_{\text{cone}} < 7.0 \text{ (av.} = 4.45) \quad (18)$$

Since the size of the pencil tip, characterized by ℓ and α, determines the thickness or weight of the line drawn with the sharpened pencil, it is meaningful to compare the two results from equations (17) and (18). On the average, we can see that a given pencil, when overloaded in drawing a line of given weight, will have larger broken pencil points, as measured by the parameter N, when sharpened into a wedge than the same pencil sharpened into a conical point. Since the larger N, the greater the fracture area, we can also conclude that it would take on the average a greater effort to break a wedge-pointed pencil. Hence, not only does the wedge-pointed pencil have the advantage that it does not have to be twirled to keep line weight uniform, but also it can generally withstand a heavier hand on the part of the draftsperson.

Conclusion

The wood-case pencil is a common technological artifact whose size, shape, and composition have no doubt evolved with very little, if any, mathematical analysis. Equations flow from pencils, but pencils do not come of equations. Indeed, at one time during the nineteenth century, arguably the best pencils made in America were manufactured by John Thoreau and Co., after a process perfected by John Thoreau's son, Henry David, who is remembered neither as a mathematician nor as an engineer (Harding, 1965).

The use of the pencil by engineers and others, in particular the customary nature of the point with which drafting and sketching has traditionally been executed, also probably evolved more through trial and error and serendipity than through any deliberate, rational mathematical analysis. But, regardless of its origins and use, the pencil is as proper an object of analysis as is the natural world and universe. By asking why and how a pencil point breaks in the way it does, we are not only led to a better understanding of the tools of stress analysis and their limitations, but we are also led to a fuller appreciation of the wonders of technology when we analyze the aptness of such a manufactured product as the common pencil.

References

Cowin, S. C., 1983, "A Note on Broken Pencil Points," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 453-454.

Cronquist, D., 1979, "Broken-off Pencil Points," *American Journal of Physics*, Vol. 47, pp. 653-655.

Fleming, C. C., and Guptill, A. L., 1936, *The Pencil: Its History, Manufacture and Use*. Koh-I-Noor Pencil Co., Bloomsbury, N.J.

Galilei, G., 1638, *Dialogues Concerning Two New Sciences*, translated by H. Crew and A. de Salvio, reprinted 1954, Dover Publications, New York.

Halse, A. O., 1960, *Architectural Rendering: The Techniques of Contemporary Presentation*, McGraw-Hill Book Company, New York, Ch. 10.

Harding W., 1965, *The Days of Henry Thoreau*, Alfred A. Knopf, New York, pp. 56-57.

Hoelscher, R. P., and Springer, C. H., 1956, *Engineering Drawing and Geometry*, Wiley, New York, Ch. 3.

Kirby, R. S., 1925, *The Fundamentals of Mechanical Drawing*, Wiley, New York, Ch. 1.

Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, Fourth Edition, reprinted 1944, Dover Publications, New York, pp. 202-203.

Svensen, C. L., and Street, W. E., 1962, *Engineering Graphics*, D. Van Nostrand Company, Princeton, NJ, Ch. 2.

Timoshenko, S. P., 1953, *History of Strength of Materials: With a Brief Account of the History of Theory of Elasticity and Theory of Structures*, reprinted 1983, Dover Publications, New York.

Timoshenko, S., 1956, *Strength of Materials: Part II*, Advanced Theory and Problems, Third Edition, D. Van Nostrand Company, Princeton, NJ, pp. 60-64.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, Third Edition, McGraw-Hill Book Company, New York, pp. 109-113.

Todhunter, I., and Pearson, K., 1886, *A History of the Theory of Elasticity and of the Strength of Materials from Galilei to Lord Kelvin*, reprinted 1960, Dover Publications, New York.

Volterra, E., and Gaines, J. H., 1971, *Advanced Strength of Materials*, Prentice-Hall, Englewood Cliffs, NJ, pp. 163-169, pp. 175-177.

Walker, J., 1979, "The Amateur Scientist," *Scientific American*, Vol. 240, Feb., pp. 156-158. (See also Vol. 241, November, pp. 200-204.)

Inclusion Effects on Stress Measurement in Geological Materials

A. L. Florence[11]

Introduction

Stress measurement in geologic materials by embedded gauges is complicated because the gauge forms an inclusion with properties different from those of the surrounding medium. The sensing element measures the stress or pressure in the inclusion, which then has to be related to the far-field stress component we wish to measure. The analytical results presented here provide this relationship for elastic axisymmetric quasi-static stress fields.

Our analysis was motivated by a need to support our laboratory experiments to develop and apply miniature gauges for obtaining compressive stress measurements in rocks, rock simulants, and soils. As background, we describe briefly the design of the stress gauge shown in Fig. 1. It consists of a thin disk of elastic material (2.0 cm diameter and 0.2 cm thick) cast around a very thin flat foil (0.3 cm square and 0.002 cm thick) of piezoresistive material, such as ytterbium. The gauge is thin to allow the stress normal to the face to have the dominant effect on the inclusion stress field. This normal stress is the component the gauge is trying to measure. The sensing element is a piezoresistive conductor with a scalar response, consisting of a resistance change, to applied tensorial stress and strain (Gupta 1983, 1984). We, therefore, have to restrict the inclusion stress-strain state so the resistive change can be related to that state. Such states are generally hydrostatic or uniaxial strain and they are calibratable. The hydrostatic state is provided by a gauge of fluid-like material (low shear modulus) around the foil. The uniaxial strain state is provided by bonding the foil to a thin, relatively stiff, material, such as steel. Thus, we can relate the resistance change to the pressure in a fluid inclusion or to the inclusion stress component normal to the surface of the sensing foil. Our task is to relate the inclusion pressure or stress normal to the disk face to the medium far-field stress normal to the disk face.

Our analytical approach to the elastic inclusion problem is first to replace the thin disk with a thin oblate spheroid having the same aspect ratio, so that we can employ Eshelby's theory of inclusions (1958) to relate the inclusion and the far-field stress loading. We then examine the results to obtain the effects on this relationship of the loading stress ratio, the inclusion aspect ratio, and the elastic properties of the inclusion and medium. We assume that the very thin foil embedded in the inclusion has no effect on the stress distribution throughout most of the inclusion. In fact, the foil forms an additional inclusion problem. The analysis of an oblate spheroidal inclusion under far-field loading has been treated (Edwards, 1951; Shibata and Kanji Ono, 1978), but our results were derived in applicable form by Eshelby's method because we are employing the method in a more general context of gauge design.

Oblate Spheroidal Inclusion

We consider an oblate spheroidal inclusion at the origin of axes (x_1, x_2, x_3) and x_3 the axis of symmetry. Let the semiaxes have lengths a, b, and c such that $c < b = a$. The applied stress σ_{ij}^A is also symmetric about the x_3 axis, so $\sigma_{11}^A = \sigma_{22}^A$, and $\sigma_{ij}^A = 0$ $(i \neq j)$. Application of Eshelby's method leads to the inclusion stress formulas

[11]Engineering Mechanics Department, SRI International, Menlo Park, CA 94025. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, May 8, 1985; final revision April 7, 1987.

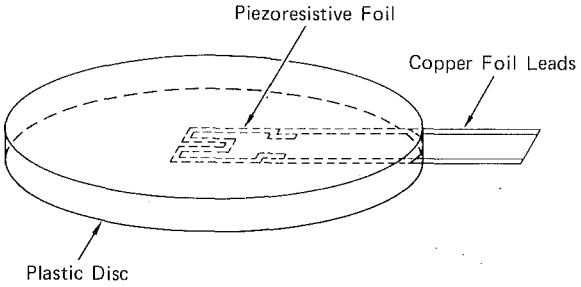


Fig. 1 Schematic of a miniature stress gage

$$3H\sigma_{11} = \{n[(m-1)B'+1] + m(n-1)A'/\alpha\}p^A - \{n(m-1)\alpha B + m[(n-1)A+1]\}\tau^A \quad (1)$$

$$3H\sigma_{33} = \{n[(m-1)B'+1] - m(n-1)2A'/\alpha\}p^A - \{n(m-1)\alpha B - 2m[(n-1)A+1]\}\tau^A \quad (2)$$

where

$$H = [(n-1)A+1][(m-1)B'+1] - (m-1)(n-1)A'B$$
$$A = 1-(16\pi/3)R \quad B = 1-A-4RI_a$$
$$A' = A-(Q-R)I_a$$
$$B' = 1-A'-RI_a - Q\{(4\pi/3)-I_a\}3k^2/(1-k^2)$$
$$Q = 3/8\pi(1-\nu) \quad R = (1-2\nu)/8\pi(1-\nu)$$
$$I_a = 2\pi k\{(\cos^{-1}k)/(1-k^2)^{1/2} - k\}/(1-k^2)$$
$$k = \frac{c}{a} \quad (0<k<1)$$
$$p^A = \sigma_{33}^A + 2\sigma_{11}^A \quad \tau^A = \sigma_{33}^A - \sigma_{11}^A$$
$$\alpha = 3K/2\mu \quad m = \mu^*/\mu \quad n = K^*/K \quad (3)$$

In these formulas K, μ, and ν are the bulk modulus, shear modulus, and Poisson's ratio. An asterisk indicates the inclusion material property.

Because of our interest in thin gauges in order to minimize the difference between the inclusion stress σ_{33} and the free-field stress component σ_{33}^A to be measured, we can simplify stress formula (2) by expanding in powers of the inverse aspect ratio $k = c/a$ and neglect terms with powers of k greater than unity. This process leads to

$$\sigma_{33}/\sigma_{33}^A = 1-(\pi/16)(B_o/H_o)\{3(m-1)(n-1)(3A-1) - 2(m-1) - 4(n-1) - 2(n-m)\Lambda\}k \quad (4)$$

where

$$H_o = nA + mB_o \quad B_o = 1-A \quad \Lambda = \sigma_{11}^A/\sigma_{33}^A \quad (5)$$

In equations (5), the constants A, B_o, and H_o apply to an infinity thin disk, A being the same as that given by equations (3). The far-field stress ratio Λ is called here the loading ratio and we shall restrict our attention to the uniaxial to hydrostatic range, $0 < \Lambda < 1$. The design formula becomes inaccurate when both $m = \mu^*/\mu$ and $n = K^*/K$ are small because of the binomial approximation made for H^{-1}; a guide is that we should have $k << 2n$ when $\nu = 1/3$ and $m = n$. The case of both m and n small is not of practical interest.

If we want to subject the sensing foil to a hydrostatic state by using an inclusion material of very low shear modulus ($m \approx 0$), we can approximate design formula (4) by

$$\sigma_{33}/\sigma_{33}^A = 1-(\pi/4\alpha n)\{1-(n-1)(9A+1)/2 - n\Lambda\}k \quad (6)$$

in which α and A are functions of ν for the medium according to equations (3).

Discussion of Results

Our discussion on the results of the analysis is focussed on gauge inclusions of fluid-like material, that is, material with a low shear modulus. Under quasi-static and static conditions the fluid will not allow an undesirable stretch of the piezoresistive foil. Under dynamic conditions, however, stretching may be induced by viscous drag.

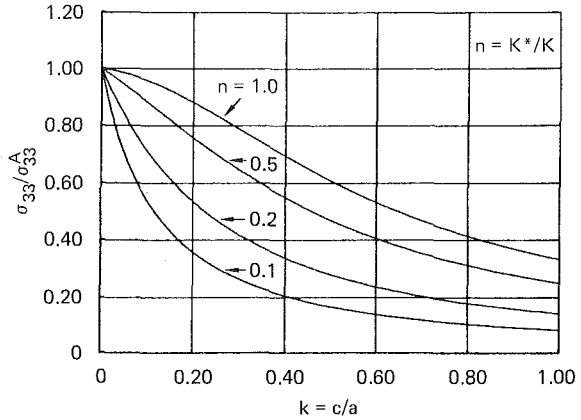


Fig. 2 Influence on registration factor of aspect ratio for several bulk modulus ratios ($m = \mu^*/\mu = 0$, $\sigma_{11}^A = 0$)

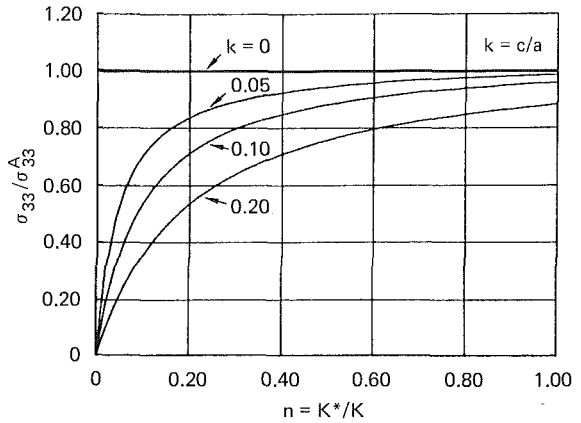


Fig. 3 Influence on registration factor of bulk modulus ratio for several thin inclusions ($m = \mu^*/\mu = 0$, $\sigma_{11}^A = 0$)

The results are presented to examine the ratio, $\sigma_{33}/\sigma_{33}^A$, of the inclusion stress being measured to the far-field stress we want to measure. We shall call this ratio the registration factor. We either require this factor to be close to unity or relatively invariant under the range of loading factors, $0 < \Lambda < 1$. Our calculations have shown that the worst conditions are caused by uniaxial stress described by $\Lambda = 0$.

Figure 2 shows the effect on the registration factor, $\sigma_{33}/\sigma_{33}^A$, of the bulk modulus mismatch n over the full range of oblate spheroidal aspect ratios for a fluid inclusion subjected to uniaxial stress. It is immediately evident that the inclusion should be thin with inverse aspect ratios below $k = 0.05$. Then the registration factor is close to unity and fairly insensitive to the bulk modulus ratio (for $0.2 < n < 1.0$).

Figure 3 shows how the registration factor varies with the bulk modulus ratio for a few thin fluid inclusions under uniaxial stress conditions. If for practical reasons the inclusion must have an inverse aspect ratio larger than desired, we see that the bulk modulus ratio must be large enough. For example, if the minimum practical value of inverse aspect ratio is $k = 0.1$ we should try to select a fluid-like material to give $n \gtrsim 0.4$.

Should the inclusion material have some shear resistance, and the oblate spheroid is slender, say $k = 0.1$, the influence of small values of the shear modulus ratio, m, is fairly small, as shown in Fig. 4. The registration factor increases from about 0.7 to 0.75 when $k = 0.1$, $n = 0.2$, and m increases

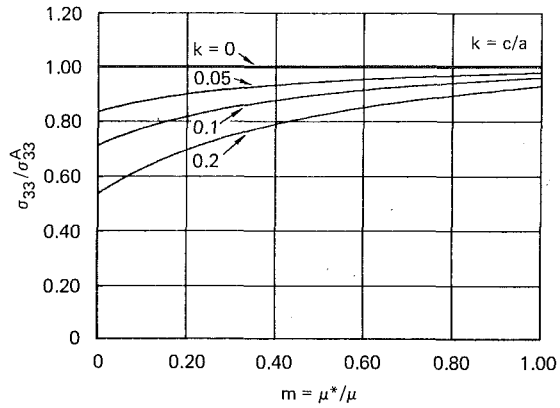


Fig. 4 Influence on registration factor of shear modulus ratio for several thin inclusions ($n = K^*/K = 0.2$, $\sigma_{11}^A = 0$)

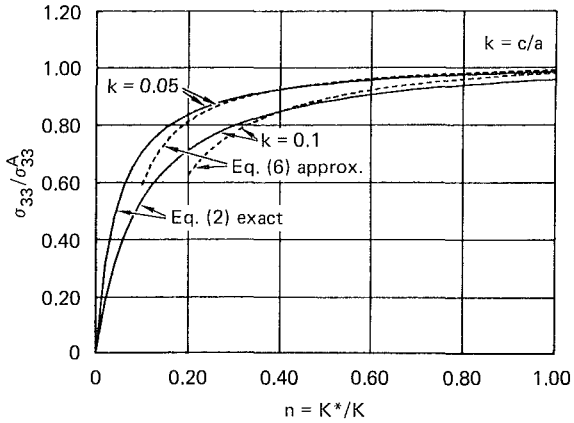


Fig. 5 Comparison of exact and approximate formulas for the registration factor ($m = \mu^*/\mu = 0$, $\sigma_{11}^A = 0$)

from zero to 0.1. This influence is less if the bulk modulus ratio is increased.

Whenever the oblate spheroid is slender ($k < 0.1$) and the bulk modulus ratio is large enough ($n \gtrsim 0.2$) the change in registration factor with loading ratio Λ is small, which is a very desirable feature during nonproportional loading.

Figure 5 shows a comparison of the pressures in a thin fluid gauge inclusion under uniaxial stress loading. The full lines were obtained from equation (2) with $m = 0$, $\Lambda = 0$, $k = 0.05$ and 0.10, and the dashed lines were obtained from the approximate formula (6) for a thin fluid inclusion, also with $\Lambda = 0$, $k = 0.05$ and 0.10. The comparison for $\Lambda = 1$ (not shown) is even better. Provided $n \gtrsim 2k$, Fig. 5 shows that (6) is a useful design formula that can be used to obtain the influence on stress measurement with a fluid gauge of loading ratio, bulk modulus ratio, medium Poisson's ratio, and low inverse aspect ratios. For example, as $n \to 1$ we see immediately from (6) that the registration factor tends to

$$\sigma_{33}/\sigma_{33}^A = 1 - (1-\Lambda)(\pi k/4\alpha)$$

For $\nu = 1/3$, and hence $\alpha = 4$, we have $\sigma_{33}/\sigma_{33}^A = 1 - (1-\Lambda)\pi k/16$. Because $\pi k/16$ is small, the stress measurement is insensitive to the loading ratio. Examination of (6) and Fig. 5 shows that satisfactory registration factors are obtained if the bulk modulus ratio lies in the range $0.3 < n < 1$.

Onset of Plasticity

We assume that as loading increases the fluid-like inclusion material remains elastic and so yielding first occurs in the medium. The location of yielding will be at the inclusion/matrix interface. To find the stress there we again call upon Eshelby's method. Generally, yielding will occur first at the edge of a slender inclusion, a point on which can be described by $x_1 = a$, $x_2 = 0$, $x_3 = 0$. If we let the stress difference there be $\sigma_{33}^M - \sigma_{11}^M = \tau^M$, having a Tresea yield condition in mind, we find that

$$H\tau^M = \{-(n-1)[(m-1)B' + 1] + (m-1)(A'/\alpha)\}p^A A/\alpha + (m-1)\{(n-1)\alpha B + [(n-1)A + 1]\}\tau^A A/\alpha - (n-1)(A'/\alpha)p^A + [(n-1)A + 1]\tau^A \quad (7)$$

As a special case of interest a slender fluid inclusion ($k \approx 0$, $m = 0$) subjected to a far-field uniaxial stress loading ($\sigma_{33}^A = p^A = \tau^A$, $\Lambda = 0$) reduces equation (7) to

$$\tau^M/\sigma_{33}^A = (1-n)(1-\alpha)(\alpha - 3A)/\alpha^2 n - 1/\alpha + 1/An \quad (8)$$

After setting $\tau^M = \sigma_o$, the uniaxial compressive strength of the medium material (8) gives the uniaxial loading for incipient plastic deformation around the edge of the slender fluid inclusion. If the material has $\nu = 1/3$, equation (8) gives $\sigma_{33}^A = 8n\sigma_o/(9 + n)$. As expected, if the bulk modulus of the fluid is much lower than that of the medium the loading range for elastic response ($\sigma_{33}^A \approx n\sigma_o$) becomes limited.

Conclusions

We have obtained theoretical results to support our laboratory experiments for developing a miniature stress gauge for measuring compressive stresses in geologic materials under axisymmetric quasistatic elastic conditions. The results have focussed on a gauge inclusion made of fluid-like material to simplify the interpretation of the piezoresistive foil resistance change due to stress and strain.

The results show that the best stress registration factors ($\sigma_{33}/\sigma_{33}^A$) are obtained if the inclusion has a high aspect ratio ($k < 0.1$), and a bulk modulus ratio that is not too low ($0.3 < n < 1.0$). These requirements ensure a registration factor that is not too far below unity, ($0.8 < \sigma_{33}/\sigma_{33}^A < 1.0$), and insensitive to the far-field loading ratio ($0 < \Lambda < 1$). Above all, the results determine the factor in terms of the parameters so that stress gauge development is guided.

Acknowledgments

The author wishes to acknowledge C. E. Keller, D. D. Keough, and G. Nagy for helpful discussions on stress measurements in continua.

References

Edwards, R. H., 1951, "Stress Concentrations Around Spheroidal Inclusions and Cavities," ASME Journal of Applied Mechanics, Vol. 18.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proc. Royal Soc. London*, Vol. A241, pp. 376-396.

Gupta, Y. M., 1983, "Stress Measurements Using Piezoresistance Gages: Modeling the Gage as an Elastic-Plastic Inclusion," *J. Appl. Phys.*, Vol. 54, No. 11, pp. 6256-6266.

Gupta, Y. M., 1984, "Quasistatic Experiments to Determine Material Constants for the Piezoresistance Foils Used in Shock Wave Experiments," *J. Appl. Phys.*, Vol. 55, No. 11, pp. 3984-3993.

Shibata, M., and Kanji, O., 1978, "Internal Stresses Due to an Oblate Spheroidal Inclusion: Misfit, Inhomogeneity, and Plastic Deformation Effects," *Acta Metallurgica*, Vol. 26, pp. 921-932.

An Approximate Green's Function for Beams and Application to Contact Problems

B. V. Sankar[12]

[12]Assistant Professor, Department of Engineering Sciences, University of Florida, Gainesville, FL 32611. Assoc. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, December 17, 1986; final revision February 21, 1987.

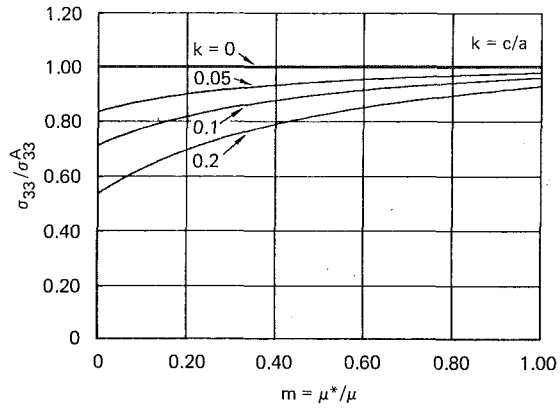


Fig. 4 Influence on registration factor of shear modulus ratio for several thin inclusions ($n = K^*/K = 0.2$, $\sigma_{11}^A = 0$)

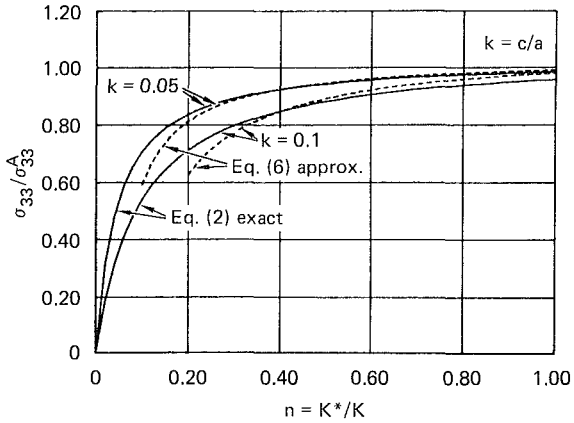


Fig. 5 Comparison of exact and approximate formulas for the registration factor ($m = \mu^*/\mu = 0$, $\sigma_{11}^A = 0$)

from zero to 0.1. This influence is less if the bulk modulus ratio is increased.

Whenever the oblate spheroid is slender ($k < 0.1$) and the bulk modulus ratio is large enough ($n \gtrsim 0.2$) the change in registration factor with loading ratio Λ is small, which is a very desirable feature during nonproportional loading.

Figure 5 shows a comparison of the pressures in a thin fluid gauge inclusion under uniaxial stress loading. The full lines were obtained from equation (2) with $m = 0$, $\Lambda = 0$, $k = 0.05$ and 0.10, and the dashed lines were obtained from the approximate formula (6) for a thin fluid inclusion, also with $\Lambda = 0$, $k = 0.05$ and 0.10. The comparison for $\Lambda = 1$ (not shown) is even better. Provided $n \gtrsim 2k$, Fig. 5 shows that (6) is a useful design formula that can be used to obtain the influence on stress measurement with a fluid gauge of loading ratio, bulk modulus ratio, medium Poisson's ratio, and low inverse aspect ratios. For example, as $n \to 1$ we see immediately from (6) that the registration factor tends to

$$\sigma_{33}/\sigma_{33}^A = 1 - (1-\Lambda)(\pi k/4\alpha)$$

For $\nu = 1/3$, and hence $\alpha = 4$, we have $\sigma_{33}/\sigma_{33}^A = 1 - (1-\Lambda)\pi k/16$. Because $\pi k/16$ is small, the stress measurement is insensitive to the loading ratio. Examination of (6) and Fig. 5 shows that satisfactory registration factors are obtained if the bulk modulus ratio lies in the range $0.3 < n < 1$.

Onset of Plasticity

We assume that as loading increases the fluid-like inclusion material remains elastic and so yielding first occurs in the medium. The location of yielding will be at the inclusion/matrix interface. To find the stress there we again call upon Eshelby's method. Generally, yielding will occur first at the edge of a slender inclusion, a point on which can be described by $x_1 = a$, $x_2 = 0$, $x_3 = 0$. If we let the stress difference there be $\sigma_{33}^M - \sigma_{11}^M = \tau^M$, having a Tresea yield condition in mind, we find that

$$H\tau^M = \{-(n-1)[(m-1)B' + 1] + (m-1)(A'/\alpha)\}p^A A/\alpha + (m-1)\{(n-1)\alpha B + [(n-1)A + 1]\}\tau^A A/\alpha - (n-1)(A'/\alpha)p^A + [(n-1)A + 1]\tau^A \quad (7)$$

As a special case of interest a slender fluid inclusion ($k \approx 0$, $m = 0$) subjected to a far-field uniaxial stress loading ($\sigma_{33}^A = p^A = \tau^A$, $\Lambda = 0$) reduces equation (7) to

$$\tau^M/\sigma_{33}^A = (1-n)(1-\alpha)(\alpha - 3A)/\alpha^2 n - 1/\alpha + 1/An \quad (8)$$

After setting $\tau^M = \sigma_o$, the uniaxial compressive strength of the medium material (8) gives the uniaxial loading for incipient plastic deformation around the edge of the slender fluid inclusion. If the material has $\nu = 1/3$, equation (8) gives $\sigma_{33}^A = 8n\sigma_o/(9 + n)$. As expected, if the bulk modulus of the fluid is much lower than that of the medium the loading range for elastic response ($\sigma_{33}^A \approx n\sigma_o$) becomes limited.

Conclusions

We have obtained theoretical results to support our laboratory experiments for developing a miniature stress gauge for measuring compressive stresses in geologic materials under axisymmetric quasistatic elastic conditions. The results have focussed on a gauge inclusion made of fluid-like material to simplify the interpretation of the piezeoresistive foil resistance change due to stress and strain.

The results show that the best stress registration factors ($\sigma_{33}/\sigma_{33}^A$) are obtained if the inclusion has a high aspect ratio ($k < 0.1$), and a bulk modulus ratio that is not too low ($0.3 < n < 1.0$). These requirements ensure a registration factor that is not too far below unity, ($0.8 < \sigma_{33}/\sigma_{33}^A < 1.0$), and insensitive to the far-field loading ratio ($0 < \Lambda < 1$). Above all, the results determine the factor in terms of the parameters so that stress gauge development is guided.

Acknowledgments

The author wishes to acknowledge C. E. Keller, D. D. Keough, and G. Nagy for helpful discussions on stress measurements in continua.

References

Edwards, R. H., 1951, "Stress Concentrations Around Spheroidal Inclusions and Cavities," ASME Journal of Applied Mechanics, Vol. 18.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proc. Royal Soc. London*, Vol. A241, pp. 376-396.

Gupta, Y. M., 1983, "Stress Measurements Using Piezoresistance Gages: Modeling the Gage as an Elastic-Plastic Inclusion," *J. Appl. Phys.*, Vol. 54, No. 11, pp. 6256-6266.

Gupta, Y. M., 1984, "Quasistatic Experiments to Determine Material Constants for the Piezoresistance Foils Used in Shock Wave Experiments," *J. Appl. Phys.*, Vol. 55, No. 11, pp. 3984-3993.

Shibata, M., and Kanji, O., 1978, "Internal Stresses Due to an Oblate Spheroidal Inclusion: Misfit, Inhomogeneity, and Plastic Deformation Effects," *Acta Metallurgica*, Vol. 26, pp. 921-932.

An Approximate Green's Function for Beams and Application to Contact Problems

B. V. Sankar[12]

[12]Assistant Professor, Department of Engineering Sciences, University of Florida, Gainesville, FL 32611. Assoc. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, December 17, 1986; final revision February 21, 1987.

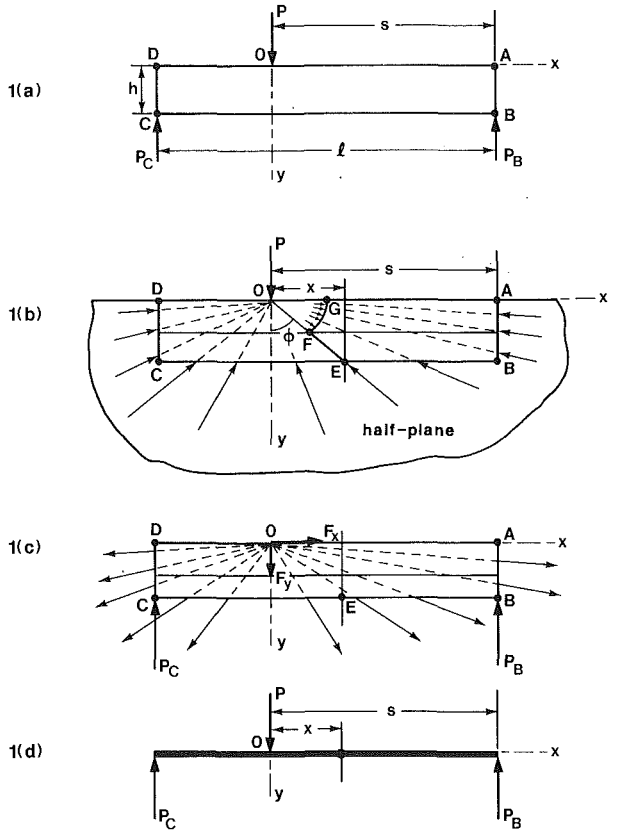


Fig. 1 Principle of superposition for calculating normal displacements on the surface of a beam

1 Introduction

The problem of indentation of a beam supported at both ends has been solved by Sankar and Sun (1983) and Keer and Miller (1983). The basic principle in both methods was using the elasticity solution to describe displacements and stresses in the vicinity of contact, and using the beam-theory equations to describe the global behavior. They differed in the numerical technique of solving the contact problem, but reached essentially the same conclusions. It may be noted that the basic information needed to solve a contact problem is the Green's function for surface displacements. Then, the problem may be formulated in terms of an integral equation, which may be solved numerically. Once the contact area and the contact stresses beneath the indenter are found, the stress field in the contacting bodies can be solved by using the equations of elasticity.

In this paper an approximate Green's function for normal displacements on the surface of a beam is proposed. The integral equation for frictionless contact between a beam and a rigid cylindrical indenter is formed in terms of the Green's function. The integral equation is solved by a least squares approximation procedure. A numerical example is given for the problem of central indentation of a simply supported isotropic beam by a smooth, rigid cylinder. The results for the contact stresses beneath the indenter are compared with those given by Keer and Miller (1983).

2 An Approximate Green's Function

Consider the problem of a simply supported beam of rectangular cross section and unit width subjected to a concentrated force P as shown in Fig. 1(a). In the context of contact problems our interest is in determining the surface displacements $v(x, 0)$ in the y direction. As described in Timoshenko and Goodier (1970) the solution to the problem in Fig. 1(a) can be obtained as the superposition of solutions of systems shown in Figs. 1(b) and 1(c). In the above reference such a superposition procedure has been used to calculate stresses in a beam subjected to a concentrated force. We shall extend the same method for determining the displacements as

$$v(x,0) = v_h(x,0) + v_{hb}(x,0), \tag{1}$$

where v_h and v_{hb} are the displacements of systems in Figs. 1(b) and 1(c), respectively.

In Fig. 1(b) the force P acts on a half-plane. The expression for surface displacements in a half-plane of unit thickness under plane stress is given by (Timoshenko and Goodier, 1970)

$$v_h(x,0) = -(2P/\pi E)\ \log\ |x| + \text{constant}. \tag{2}$$

It will be shown later that in contact problems we need only relative displacements, and there is no need to evaluate the constant term.

In Fig. 1(c) radial tensile stresses act on the sides of the rectangular beam supported at the ends. The magnitude of these radial stresses are equal to the magnitude of the compressive stresses on face $ABCD$ in the half-plane in Fig. 1(b). Displacements $v_{hb}(x, 0)$ for the problem shown in Fig. 1(c) can be obtained from the beam theory. It will be further shown that the displacement v_{hb} is approximately equal to the deflection v_b in a beam subjected to a concentrated force P as shown in Fig. 1(d).

The bending moment about the centroid at any section of the beam in Fig. 1(c) can be easily computed if we note that the radial tractions on face ABE are statically equivalent to the radial pressure over the circular arc FG of an arbitrary radius r in Fig. 1(b). The expression for this radial pressure distribution is given in Timoshenko and Goodier (1970) as

$$\sigma_{rr} = -(2P\cos\theta)/\pi r, \tag{3}$$

where θ is measured in a counter-clockwise sense from the y axis. The resultant of this radial pressure is equivalent to two forces F_x and F_y (Fig. 1(c)) acting at O given by

$$F_x = P(1+\cos 2\phi)/2\pi$$

and (4)

$$F_y = P(\pi - 2\phi - \sin 2\phi)/2\pi,$$

where $\tan\phi = x/h$.

The bending moment $M(x)$ at any section can be calculated as

$$M(x) = \lambda P(s-x) - F_x(h/2) + F_y x, \tag{5}$$

where $\lambda = 1 - s/\ell$.

The displacement v_{hb} is then obtained by integrating $M(x)$ twice as follows:

Table 1 Comparison of displacement V_{hb} and deflections V_b

x/h	$(EI/Ph^3)v_b$	$(EI/Ph^3)v_{hb}$	$(v_b - v_{hb})/v_{hb}$
0.0	8.5333	8.4612	0.00853
0.1	8.2055	8.1347	0.00869
0.2	7.8624	7.7942	0.00875
0.3	7.5049	7.4401	0.00871
0.4	7.1339	7.0729	0.00862
0.5	6.7500	6.6932	0.00849
0.6	6.3541	6.3016	0.00833
0.7	5.9471	5.8989	0.00816
0.8	5.5296	5.4858	0.00798
0.9	5.1025	5.0630	0.00781
1.0	4.6667	4.6313	0.00764

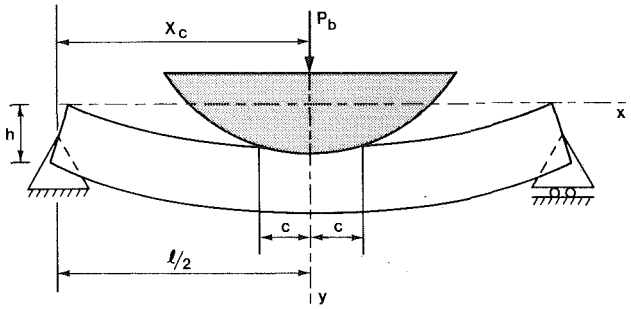


Fig. 2 Central indentation of a simply supported beam

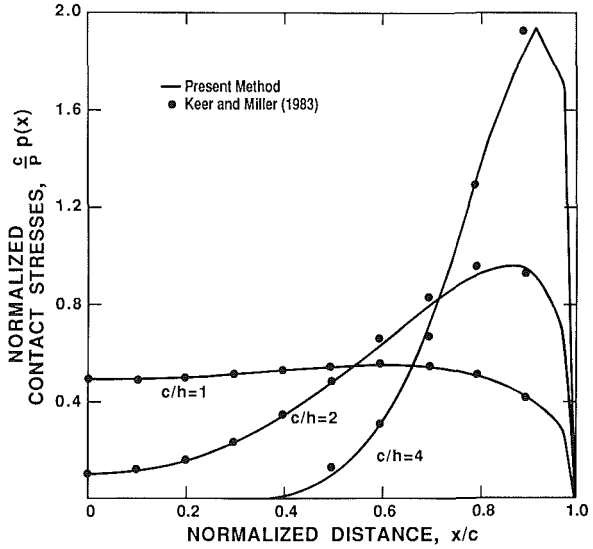


Fig. 3 Contact stress variation in a simply supported beam

$$(EI/Ph^3)V_{hb}(x) = \tilde{x}^2\left(\frac{1}{6\pi} - \frac{\lambda s}{2h}\right) + \left(\frac{\lambda - 0.5}{6}\right)\tilde{x}^3 + \frac{1}{12\pi}\left[2\tilde{x}^3 \tan^{-1}\tilde{x} + \log(1+\tilde{x}^2)\right] + c_1\tilde{x} + c_2, \quad (6)$$

where $\tilde{x} = x/h$.

In equation (6) E is the Young's modulus, I is the moment of inertia of beam-cross section, and s is the distance of force P from the right support. The constants of integration c_1 and c_2 can be determined from the conditions that displacements are zero at $x = s$ and $x = (s - \ell)$.

The displacements thus obtained were compared with the deflection $v_b(x)$ of a beam subjected to a concentrated force P calculated using the elementary beam formula. The agreement was excellent for large ℓ/h ratios, and also when the load was not very close to either support. The comparison for a worst case ($\ell/h = 10$ and $s/\ell = 0.2$) is given in Table 1. It may be noted that the maximum difference in displacements occurs near the point of application of the concentrated force, but it is still less than 1 percent. An interpretation of this agreement can be that the resultant of F_x and F_y in Fig. 1(c) pass very close to the centroid at the section through E, and does not contribute much to the bending moment, making $M(x)$ approximately equal to $\lambda P(s - x)$, which is identical to the problem in Fig. 1(d). It should be remembered that v_{hb} or v_b is only part of the solution to which v_h has to be added.

In conclusion, it has been shown that the transverse displacements on the surface of a beam due to a concentrated force can be calculated by superposing the displacements on the surface of the half-plane and the beam-theory deflections. The solution for a concentrated force can then be used as a Green's function for computing displacements due to any other type of loading on the beam.

3 Central Indentation of a Simply Supported Beam

Consider a simply supported beam of length ℓ and thickness h (Fig. 2). The beam is assumed to be of unit width and is in a state of plane stress parallel to the $x - y$ plane. The beam is indented by a rigid cylindrical indenter with a parabolic profile given by $y = -x^2/2R$. The contact is assumed to be smooth. Our interest is in determining the contact stress distribution $p(x) = -\sigma_{yy}(x, 0)$ beneath the indenter.

We start with a known contact length $2c$ symmetrical about the center of the beam. The contact length is divided into N_d number of divisions. The unknown contact stresses are assumed to be uniform over each division, that is, over the jth division $p(x) = p_j$. The p_j's are determined from the condition that the deformed shape of the beam in the contact region should conform to the shape of the indenter. This is achieved by choosing N_c number of collocation points x_i's including $|x| = c$, and requiring that

$$v(x_r,0) - v(x_i,0) = (x_i^2 - x_r^2)/2R, \; i = 1, N_c, \quad (7)$$

where x_r is a reference point in the contact region. The left-hand side of equation (7) can be found as a linear function of p_j's using the superposition principle described in Section 2. Thus, the system of equation (7) can be written in the form

$$\sum_j A_{ij}p_j = (x_i^2 - x_r^2)/2R, \; i = 1, N_c. \quad (8)$$

The number N_c has to be at least equal to N_d, but it was found that the variation of contact stresses would be smooth if $N_c > N_d$, and the least squares solution procedure was used to solve for the p_j's. In the numerical examples N_c was equal to 25 and N_d was equal to 20. The IMSL subroutine LLSQF was used in a VAX-11/780 computer to solve the system of linear equations.

In order to compute A_{ij}'s, the solution for relative normal displacements on the boundary of the half-plane due to a uniform load, say p over $-t < x < t$, is needed. For the case of plane stress the relative displacements can be expressed as (Timoshenko and Goodier, 1970)

$$v_h(x,0) - v_h(0,0) = (-2pt/\pi E)[(1-\bar{x})\log|1-\bar{x}| + (1+\bar{x})\log|1+\bar{x}|], \quad (9)$$

where $\bar{x} = x/t$.

The contact stress distribution shown in Fig. 3 corresponds to the case $\ell = 50.8$ mm, $h = 2.54$ mm, $R = 25.4$ mm, and $E = 6.8971$ GPa (10^6 psi). The results agree well with those given in Sankar and Sun (1983). In Fig. 3 the symbols represent the results obtained by Keer and Miller (1983). Again, the agreement is quite good.

4 Summary

The normal displacements on the surface of a beam can be obtained by superposing the beam-theory deflections and the corresponding half-plane solutions. The restriction is that the load should not be very close to a support. This method of superposition simplifies the formulation of the problem of contact between a beam and a rigid indenter. Although the numerical example was concerned with the indentation of a simply supported beam, this method can be easily extended to other boundary conditions, and also to the case of asymmetrical indentation, where the contact stresses may not be symmetrical about the indenter.

References

Keer, L. M., and Miller, G. R., 1983, "Smooth Indentation of a Finite Layer," *ASCE J. Eng. Mech. Div.*, Vol. 109, pp. 706–717.

Sankar, B. V., and Sun, C. T., 1983, "Smooth Indentation of a Beam by a Rigid Cylinder," *Int. J. Solids Structures*, Vol. 19, pp. 293–303.

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

On Singular Solutions for Inclusion Problems in Plane Elasticity

G. R. Miller[13,15] and R. P. Young[14,15]

Introduction

In a series of papers (Dundurs and Hetenyi, 1961; Hetenyi and Dundurs, 1962; Dundurs and Mura, 1964; Dundurs and Sendeckyj, 1965) Dundurs and coworkers have presented solutions to the class of problems involving edge dislocations and concentrated force type singularities interacting with circular elastic inclusions. As later discussed by Dundurs (Dundurs, 1968), these types of singularities are closely related, having similar Airy stress function representations. The purpose of this note is to solve again this class of singular problems using a unified complex variables approach, which, beyond providing a means of deriving the appropriate complex potentials in a direct manner, further emphasizes the close relationship between the dislocation and concentrated force solutions. A similar approach has recently been used by Rubinstein (1986) for the limiting cases of a rigid inclusion and a void.

Analysis

The starting point of the analysis is the well-known formulation of Muskhelishvili (1954) written in polar form:

$$\sigma_{rr}+i\tau_{r\theta}=\Phi_k(z)+\overline{\Phi_k(z)}-\frac{\bar{z}}{z}\left[z\overline{\Phi_k'(z)}+\overline{\Psi_k(z)}\right] \tag{1}$$

$$\frac{\partial}{\partial\theta}\left[u_x+iv_y\right]=\frac{iz}{2\mu_k}\left\{\kappa_k\Phi_k(z)-\overline{\Phi_k(z)}+\frac{\bar{z}}{z}\left[z\overline{\Phi_k'(z)}+\overline{\Psi_k(z)}\right]\right\} \tag{2}$$

in which: $z = x+iy$; Φ_k and Ψ_k are complex functions of z, with $k = 1, 2$ designating the region of interest: $|z| > a$ defines region 1; $|z| < a$ defines region 2; and the respective material constants are μ and κ (μ is the shear modulus, while $\kappa = 3-4\nu$ for plane strain, and $\kappa = (3-\nu)/(1+\nu)$ for plane stress). The solution sought must contain the proper singularity at a given point, say $z = z_0$, and the tractions and displacements must be continuous across the interface between the two materials at $|z| = a$. To meet this latter condition the new set of complex potentials $\Omega_k(z)$ are introduced as follows:

$$\Omega_1(z)=\begin{cases}\Phi_1(z)-\bar{\Phi}_2\left(\dfrac{a^2}{z}\right)+\dfrac{a^2}{z}\,\bar{\Phi}_2'\left(\dfrac{a^2}{z}\right)+\dfrac{a^2}{z^2}\,\bar{\Psi}_2\left(\dfrac{a^2}{z}\right) & |z|>a\\[2ex] \Phi_2(z)-\bar{\Phi}_1\left(\dfrac{a^2}{z}\right)+\dfrac{a^2}{z}\,\bar{\Phi}_1'\left(\dfrac{a^2}{z}\right)+\dfrac{a^2}{z^2}\,\bar{\Psi}_1\left(\dfrac{a^2}{z}\right) & |z|<a\end{cases} \tag{3}$$

$$\Omega_2(z)=\begin{cases}\dfrac{\kappa_1}{2\mu_1}\,\Phi_1(z)+\dfrac{1}{2\mu_2}\left[\bar{\Phi}_2\left(\dfrac{a^2}{z}\right)-\dfrac{a^2}{z}\,\bar{\Phi}_2'\left(\dfrac{a^2}{z}\right)-\dfrac{a^2}{z^2}\,\bar{\Psi}_2\left(\dfrac{a^2}{z}\right)\right] & |z|>a\\[2ex] \dfrac{\kappa_2}{2\mu_2}\,\Phi_2(z)+\dfrac{1}{2\mu_1}\left[\bar{\Phi}_1\left(\dfrac{a^2}{z}\right)-\dfrac{a^2}{z}\,\bar{\Phi}_1'\left(\dfrac{a^2}{z}\right)-\dfrac{a^2}{z^2}\,\bar{\Psi}_1\left(\dfrac{a^2}{z}\right)\right] & |z|<a\end{cases} \tag{4}$$

in which the usual notation $\bar{\Psi}(z) = \overline{\Psi(\bar{z})}$ is used. These new potentials have been chosen such that for $z \to t^+ = ae^{i\theta}$ (i.e., z approaches the circular boundary from the outside), the following relations hold:

$$[\sigma_{rr}+i\tau_{r\theta}]_1^+-[\sigma_{rr}+i\tau_{r\theta}]_2^-=\Omega_1^+(t)-\Omega_1^-(t) \tag{5}$$

$$\frac{\partial}{\partial\theta}\left\{[u_x+iv_y]_1^+-[u_x+iv_y]_2^-\right\}=aie^{i\theta}\left\{\Omega_2^+(t)-\Omega_2^-(t)\right\} \tag{6}$$

That is, the jumps in the boundary tractions and displacements correspond to the jumps in the potentials across the interface. Conversely, if the potentials defined in equations (3) and (4) are analytic across the interface, then there will be no jump in the corresponding tractions and displacements. To determine the appropriate potentials, consider the well-known singular potentials corresponding to either a point force or an edge dislocation acting at a point, z_0:

$$\begin{aligned}{}_0\Phi_1(z)&=\frac{A}{z-z_0}\\ {}_0\Psi_1(z)&=\frac{B}{z-z_0}+\frac{A\bar{z}_0}{(z-z_0)^2}\end{aligned} \tag{7}$$

These potentials give rise to logarithmic displacements or to net forces acting on closed contours, depending on the complex constants, A and B. In particular, for a dislocation we have:

$$A=\bar{B}=\frac{\mu\{[u_x]+i[v_y]\}}{\pi i(\kappa+1)} \tag{8}$$

while for a concentrated force the appropriate values are:

$$A=-\bar{B}/\kappa=-\frac{X+iY}{2\pi(\kappa+1)} \tag{9}$$

In equation (8) the square brackets denote displacement jumps, while in equation (9) X and Y represent the concentrated force components in the x and y directions, respectively. Equations (7) have been written assuming the singularity exists in region 1, but the reverse choice could be made equally well. For the case at hand, however, the corresponding starting potentials for region 2 may be taken as ${}_0\Phi_2 = {}_0\Psi_2 = 0$. Transforming these starting potentials according to equations (3) and (4) gives rise to the following expressions for the corresponding ${}_0\Omega_k$:

$${}_0\Omega_1(z)=\begin{cases}\dfrac{A}{(z-z_0)} & |z|>a\\[2ex] -\dfrac{\bar{A}}{a^2/z-\bar{z}_0}-\dfrac{a^2}{z}\left[\dfrac{\bar{A}}{(a^2/z-\bar{z}_0)^2}\right]+\dfrac{a^2}{z^2}\left[\dfrac{\bar{B}}{a^2/z-\bar{z}_0}+\dfrac{z_0\bar{A}}{(a^2/z-\bar{z}_0)^2}\right] & |z|<a\end{cases} \tag{10}$$

[13] Assistant Professor of Civil Engineering.
[14] Graduate Student.
[15] University of Washington, Seattle, WA 98195.

Manuscript received by ASME Applied Mechanics Division, November 3, 1986; final revision May 5, 1987.

$$
{}_0\Omega_2(z) = \begin{cases} \dfrac{\kappa_1}{2\mu_1}\dfrac{A}{(z-z_0)} & |z|>a \\ \dfrac{1}{2\mu_1}\left\{\dfrac{\bar{A}}{a^2/z-\bar{z}_0}+\dfrac{a^2}{z}\left[\dfrac{\bar{A}}{(a^2/z-\bar{z}_0)^2}\right]\right. & \\ \left.-\dfrac{a^2}{z^2}\left[\dfrac{\bar{B}}{a^2/z-\bar{z}_0}+\dfrac{z_0\bar{A}}{(a^2/z-\bar{z}_0)^2}\right]\right\} & |z|<a \end{cases} \tag{11}
$$

Since for these potentials ${}_0\Omega_j^+ - {}_0\Omega_j^- \neq 0$, the interface condition is not satisfied. To correct this the additional potentials ${}_R\Omega_j$ are introduced as follows:

$$
{}_R\Omega_1(z) = \begin{cases} -\dfrac{\bar{A}}{a^2/z-\bar{z}_0}-\dfrac{a^2}{z}\left[\dfrac{\bar{A}}{(a^2/z-\bar{z}_0)^2}\right] & \\ +\dfrac{a^2}{z^2}\left[\dfrac{\bar{B}}{a^2/z-\bar{z}_0}+\dfrac{z_0\bar{A}}{(a^2/z-\bar{z}_0)^2}\right] & |z|>a \\ \dfrac{A}{(z-z_0)} & |z|<a \end{cases} \tag{12}
$$

$$
{}_R\Omega_2(z) = \begin{cases} \dfrac{1}{2\mu_1}\left\{\dfrac{\bar{A}}{a^2/z-\bar{z}_0}+\dfrac{a^2}{z}\left[\dfrac{\bar{A}}{(a^2/z-\bar{z}_0)^2}\right]\right. & \\ \left.-\dfrac{a^2}{z^2}\left[\dfrac{\bar{B}}{a^2/z-\bar{z}_0}+\dfrac{z_0\bar{A}}{(a^2/z-\bar{z}_0)^2}\right]\right\} & |z|>a \\ \dfrac{\kappa_1}{2\mu_1}\dfrac{A}{(z-z_0)} & |z|<a \end{cases} \tag{13}
$$

Noting that these additional potentials are nonsingular, it can be seen that the total potentials $\Omega_k = {}_0\Omega_k + {}_R\Omega_k$ will be analytic everywhere, thus satisfying the interface condition, and will possess the proper singular behavior as embodied in the ${}_0\Omega_k$ terms.

The remaining conditions the solution must satisfy are those at infinity and at the origin. Requiring that the stresses vanish at infinity and be bounded at the origin leads to the addition of a constant term to each of the total potentials, Ω_k. To see this, first note that the relations (3) and (4) may be inverted, giving the following matrix expressions for the Φ_k and Ψ_k in terms of the Ω_k:

for $z>a$:

$$
\begin{bmatrix} \Phi_1(z) \\ \bar{\Phi}_2\left(\dfrac{a^2}{z}\right)-\dfrac{a^2}{z}\bar{\Phi}_2'\left(\dfrac{a^2}{z}\right)-\dfrac{a^2}{z^2}\bar{\Psi}_2\left(\dfrac{a^2}{z}\right) \end{bmatrix} = \frac{2\mu_1\mu_2}{\mu_1+\kappa_1\mu_1}\begin{bmatrix} 1/2\mu_2 & 1 \\ -\kappa_1/2\mu_1 & 1 \end{bmatrix}\begin{bmatrix} \Omega_1(z) \\ \Omega_2(z) \end{bmatrix} \tag{14}
$$

for $z<a$:

$$
\begin{bmatrix} \Phi_2(z) \\ \bar{\Phi}_1\left(\dfrac{a^2}{z}\right)-\dfrac{a^2}{z}\bar{\Phi}_1'\left(\dfrac{a^2}{z}\right)-\dfrac{a^2}{z^2}\bar{\Psi}_1\left(\dfrac{a^2}{z}\right) \end{bmatrix} = \frac{2\mu_1\mu_2}{\mu_2+\kappa_2\mu_1}\begin{bmatrix} 1/2\mu_1 & 1 \\ -\kappa_2/2\mu_2 & 1 \end{bmatrix}\begin{bmatrix} \Omega_1(z) \\ \Omega_2(z) \end{bmatrix} \tag{15}
$$

The potentials Φ_k can be determined from these expressions (for z in the appropriate region), while the Ψ_k can be determined in the appropriate region by noting that for $|z| > a$, $|a^2/z| < a$, and vice versa. Thus for $\Psi_1(z)$ the correct expression for $|z| > a$ can be obtained from equation (15):

$$
\Psi_1(z) = \frac{a^2}{z^2}\left[\Phi_1(z) - z\Phi_1'(z) + \frac{\kappa_2\mu_1}{\mu_2+\kappa_2\mu_1}\bar{\Omega}_1\left(\frac{a^2}{z}\right) - \frac{2\mu_1\mu_2}{\mu_2+\kappa_2\mu_1}\bar{\Omega}_2\left(\frac{a^2}{z}\right)\right] \tag{16}
$$

The analogous expression for $\Psi_2(z)$ for $|z| < a$ is determined from equation (14):

$$
\Psi_2(z) = \frac{a^2}{z^2}\left[\Phi_2(z) - z\Phi_2'(z) + \frac{\kappa_1\mu_2}{\mu_1+\kappa_1\mu_2}\bar{\Omega}_1\left(\frac{a^2}{z}\right) - \frac{2\mu_1\mu_2}{\mu_1+\kappa_1\mu_2}\bar{\Omega}_2\left(\frac{a^2}{z}\right)\right] \tag{17}
$$

Thus the stress conditions at the origin and at infinity can be cast in terms of conditions on the potentials through equations (1) and (2). In the present case these conditions become:

$$
\Omega_k(z) \sim 0(1/z) \qquad \text{as } |z|\to\infty \tag{18}
$$

and

$$
\left\{\Phi_2(z) - z\Phi_2'(z) + \frac{\kappa_1\mu_2}{\mu_1+\kappa_1\mu_2}\bar{\Omega}_1\left(\frac{a^2}{z}\right) - \frac{2\mu_1\mu_2}{\mu_1+\kappa_1\mu_2}\bar{\Omega}_2\left(\frac{a^2}{z}\right)\right\} \sim 0(z^2) \quad \text{as } |z|\to 0 \tag{19}
$$

Substituting the total potentials $\Omega_k = {}_0\Omega_k + {}_R\Omega_k$ into these expressions then makes it possible to determine the additional constant terms required. The final expressions for the potentials Φ_k and Ψ_k can then be derived from equations (10)–(17), thus making it possible to calculate the stresses through relations such as those in equations (1) and (2) (alternatively, expressions such as equations (1) and (2) may be derived in terms of the Ω potentials directly). The final potentials are as follows:

$$
\Phi_1(z) = \frac{A}{z-z_0} - \frac{\beta-\alpha}{1+\beta}\left\{\frac{\bar{A}}{a^2/z-\bar{z}_0}+\frac{a^2}{z}\left[\frac{\bar{A}}{(a^2/z-\bar{z}_0)^2}\right] - \frac{a^2}{z^2}\left[\frac{\bar{B}}{a^2/z-\bar{z}_0}+\frac{z_0\bar{A}}{(a^2/z-\bar{z}_0)^2}\right]+\frac{\bar{A}}{\bar{z}_0}\right\} \tag{20}
$$

$$
\Psi_1(z) = \frac{B}{z-z_0}+\frac{A\bar{z}_0}{(z-z_0)^2}+\frac{a^2}{z^2}\left\{\Phi_1(z) - z\Phi_1'(z) + \left(\frac{\alpha+\beta}{\beta-1}\right)\frac{\bar{A}}{a^2/z-\bar{z}_0} - \frac{A}{z-z_0} - \frac{zA}{(z-z_0)^2} - \frac{A}{z_0}+\left(\frac{1-\alpha}{1-\beta}\right)\bar{M}\right\} \tag{21}
$$

$$
\Phi_2(z) = \left(\frac{1+\alpha}{1-\beta}\right)\frac{A}{z-z_0}+\left(\frac{\alpha-\beta}{1-\beta}\right)M \tag{22}
$$

$$\Psi_2(z) = \left(\frac{1+\alpha}{1+\beta}\right)\left[\frac{B}{z-z_0}+\frac{A\bar{z}_0}{(z-z_0)^2}\right] + \frac{a^2}{z^2}\left\{\Phi_2(z)-z\Phi_2'(z)-\left(\frac{1+\alpha}{1+\beta}\right)\left[\frac{A}{z-z_0} + \frac{zA}{(z-z_0)^2}+\frac{A}{z_0}\right]+\bar{M}\right\} \quad (23)$$

in which

$$M=\frac{(1+\alpha)}{(\alpha-1)(1+\alpha-2\beta)}\left[(\alpha-\beta)\frac{A}{z_0}-(1-\beta)\frac{\bar{A}}{\bar{z}_0}\right] \quad (24)$$

These potentials have been written in terms of Dundur's constants (Dundurs, 1968):

$$\alpha=[\mu_2(\kappa_1+1)-\mu_1(\kappa_2+1)]/[\mu_2(\kappa_1+1)+\mu_1(\kappa_2+1)]$$
$$\beta=[\mu_2(\kappa_1-1)-\mu_1(\kappa_2-1)]/[\mu_2(\kappa_1+1)+\mu_1(\kappa_2+1)] \quad (25)$$

and it should be noted that the stresses outside the inclusion are to be calculated using Φ_1 and Ψ_1, while the stresses inside the circle are determined by Φ_2 and Ψ_2.

References

Dundurs, J., and Hetenyi, M., 1961, "The Elastic Plane With a Circular Insert, Loaded by a Radial Force," ASME JOURNAL OF APPLIED MECHANICS, Vol. 83, pp. 103-111.

Dundurs, J., and Mura, T., 1964, "Interaction Between an Edge Dislocation and Circular Inclusion," *Journal of the Mechanics and Physics of Solids*, Vol. 12, pp. 177-189.

Dundurs, J., and Sendeckyj, G. P., 1965, "Edge Dislocation Inside a Circular Inclusion," *Journal of the Mechanics and Physics of Solids*, Vol. 13, pp. 141-147.

Dundurs, J., 1968, Discussion, ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 650-652.

Hetenyi, M., and Dundurs, J., 1962, "The Elastic Plane With a Circular Insert, Loaded by a Tangentially Directed Force," ASME JOURNAL OF APPLIED MECHANICS, Vol. 84, pp. 1-7.

Muskhelishvili, N. I., 1954, *Some Basic Problems in the Theory of Elasticity*, Noordhoff, Leyden, The Netherlands.

Rubinstein, A. A., 1986, "Macrocrack-Microcrack Interaction," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 505-510.

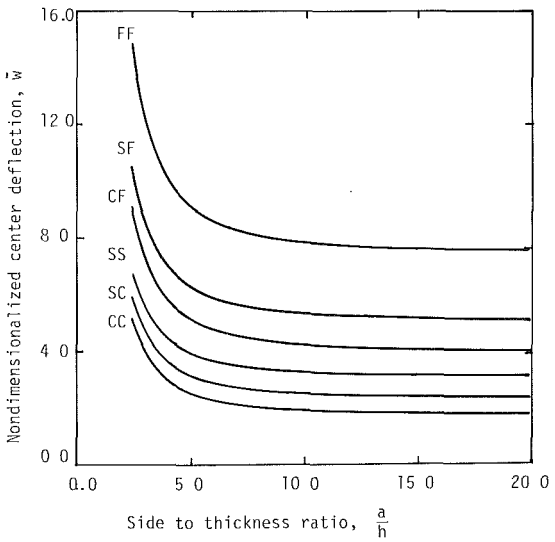


Fig. 1 Nondimensionalized center deflection versus side to thickness ratio of orthotropic plates under uniformly distributed, q_o

Lévy Type Solutions for Symmetrically Laminated Rectangular Plates Using First-Order Shear Deformation Theory

J. N. Reddy,[16,19] A. A. Khdeir,[17,19] and L. Librescu[18,19]

Introduction

Closed-form solutions for the bending case of the classical laminated plate theory and the first-order shear deformation theory were developed by Whitney and Leissa (1969), Whitney and Pagano (1970), Bert and Chen (1978), and Reddy and Chao (1981) for two types of simply supported boundary conditions and certain lamination schemes (Reddy, 1984). These served as excellent references for comparison by numerical analysis (Reddy, 1980). Such closed-form solutions of the first-order shear deformation theory for composite laminates with other types of boundary conditions are not reported in the literature. The present study deals with the development of Lévy-type solutions originated by Lévy (1899) and approached here in conjunction with the state-space concept. The analysis concerns the solution of the first-order transverse shear deformation theory of symmetrically laminated rectangular plates with two opposite edges simply supported and the remaining edges subjected to a combination of free, simply supported, and clamped boundary conditions.

Governing Equations

Consider a rectangular $(a \times b)$ laminated plate composed of N orthotropic layers, symmetrically located with respect to the midplane of the laminate. For such symmetric composite flat plates the stretching and bending states of stress are uncoupled. The equations governing the transverse bending are given by (Reddy, 1984)

$$A_{55}\frac{\partial}{\partial x}\left(\psi_x+\frac{\partial w}{\partial x}\right)+A_{44}\frac{\partial}{\partial y}\left(\psi_y+\frac{\partial w}{\partial y}\right)+q=0$$

$$\frac{\partial}{\partial x}\left(D_{11}\frac{\partial \psi_x}{\partial x}+D_{12}\frac{\partial \psi_y}{\partial y}\right)+D_{66}\frac{\partial}{\partial y}\left(\frac{\partial \psi_x}{\partial y}+\frac{\partial \psi_y}{\partial x}\right) - A_{55}\left(\psi_x+\frac{\partial w}{\partial x}\right)=0$$

$$D_{66}\frac{\partial}{\partial x}\left(\frac{\partial \psi_x}{\partial y}+\frac{\partial \psi_y}{\partial x}\right)+\frac{\partial}{\partial y}\left(D_{12}\frac{\partial \psi_x}{\partial x}+D_{22}\frac{\partial \psi_y}{\partial y}\right) - A_{44}\left(\psi_y+\frac{\partial w}{\partial y}\right)=0 \quad (1)$$

Here w denotes the transverse displacement, ψ_x and ψ_y denote the rotations of a transverse normal about the y and x axes, respectively, q is the distributed transverse load, while A_{ij} $(i, j = 4, 5)$ and D_{ij} $(i, j = 1, 2, 6)$ denote the transverse shear and bending rigidities, respectively.

The Solution Procedure

The Lévy method of solution is applicable to rectangular plates with two opposite edges simply supported and the remaining ones having arbitrary boundary conditions. The method reduces the partial differential equations to ordinary ones. The plate edges $x = 0$, a are considered simply supported (SS) while the other two edges $(y = \pm b/2)$ have arbitrary edge conditions.

[16]Clifton C. Garvin Professor, Mem. ASME.

[17]Assistant Professor.

[18]Professor.

[19]Department of Engineering Science and Mechanics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

Manuscript received by ASME Applied Mechanics Division, May 16, 1986; final revision January 22, 1987.

$$\Psi_2(z) = \left(\frac{1+\alpha}{1+\beta}\right)\left[\frac{B}{z-z_0}+\frac{A\bar{z}_0}{(z-z_0)^2}\right] + \frac{a^2}{z^2}\left\{\Phi_2(z)-z\Phi_2'(z)-\left(\frac{1+\alpha}{1+\beta}\right)\left[\frac{A}{z-z_0} + \frac{zA}{(z-z_0)^2}+\frac{A}{z_0}\right]+\bar{M}\right\} \quad (23)$$

in which

$$M=\frac{(1+\alpha)}{(\alpha-1)(1+\alpha-2\beta)}\left[(\alpha-\beta)\frac{A}{z_0}-(1-\beta)\frac{\bar{A}}{\bar{z}_0}\right] \quad (24)$$

These potentials have been written in terms of Dundur's constants (Dundurs, 1968):

$$\alpha=[\mu_2(\kappa_1+1)-\mu_1(\kappa_2+1)]/[\mu_2(\kappa_1+1)+\mu_1(\kappa_2+1)]$$
$$\beta=[\mu_2(\kappa_1-1)-\mu_1(\kappa_2-1)]/[\mu_2(\kappa_1+1)+\mu_1(\kappa_2+1)] \quad (25)$$

and it should be noted that the stresses outside the inclusion are to be calculated using Φ_1 and Ψ_1, while the stresses inside the circle are determined by Φ_2 and Ψ_2.

References

Dundurs, J., and Hetenyi, M., 1961, "The Elastic Plane With a Circular Insert, Loaded by a Radial Force," ASME JOURNAL OF APPLIED MECHANICS, Vol. 83, pp. 103-111.

Dundurs, J., and Mura, T., 1964, "Interaction Between an Edge Dislocation and Circular Inclusion," *Journal of the Mechanics and Physics of Solids*, Vol. 12, pp. 177-189.

Dundurs, J., and Sendeckyj, G. P., 1965, "Edge Dislocation Inside a Circular Inclusion," *Journal of the Mechanics and Physics of Solids*, Vol. 13, pp. 141-147.

Dundurs, J., 1968, Discussion, ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 650-652.

Hetenyi, M., and Dundurs, J., 1962, "The Elastic Plane With a Circular Insert, Loaded by a Tangentially Directed Force," ASME JOURNAL OF APPLIED MECHANICS, Vol. 84, pp. 1-7.

Muskhelishvili, N. I., 1954, *Some Basic Problems in the Theory of Elasticity*, Noordhoff, Leyden, The Netherlands.

Rubinstein, A. A., 1986, "Macrocrack-Microcrack Interaction," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 505-510.

Lévy Type Solutions for Symmetrically Laminated Rectangular Plates Using First-Order Shear Deformation Theory

J. N. Reddy,[16,19] A. A. Khdeir,[17,19] and L. Librescu[18,19]

Introduction

Closed-form solutions for the bending case of the classical laminated plate theory and the first-order shear deformation theory were developed by Whitney and Leissa (1969), Whitney and Pagano (1970), Bert and Chen (1978), and Reddy and Chao (1981) for two types of simply supported boundary conditions and certain lamination schemes (Reddy, 1984). These served as excellent references for comparison by numerical analysis (Reddy, 1980). Such closed-form solutions of the first-order shear deformation theory for composite laminates with other types of boundary conditions are not reported in the literature. The present study deals with the development of Lévy-type solutions originated by Lévy (1899) and approached here in conjunction with the state-space concept. The analysis concerns the solution of the first-order transverse shear deformation theory of symmetrically laminated rectangular plates with two opposite edges simply supported and the remaining edges subjected to a combination of free, simply supported, and clamped boundary conditions.

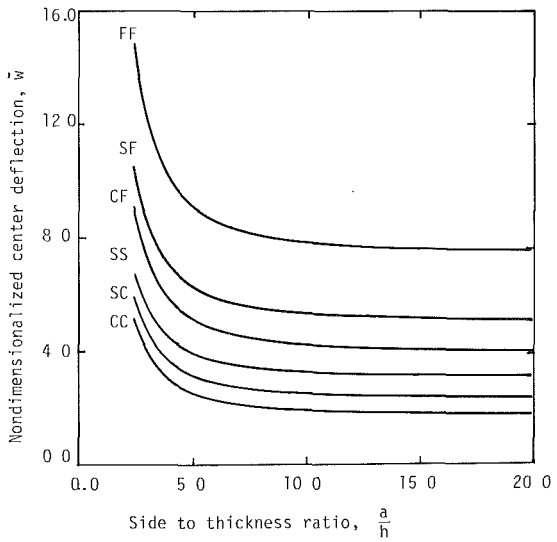


Fig. 1 Nondimensionalized center deflection versus side to thickness ratio of orthotropic plates under uniformly distributed, q_o

Governing Equations

Consider a rectangular $(a \times b)$ laminated plate composed of N orthotropic layers, symmetrically located with respect to the midplane of the laminate. For such symmetric composite flat plates the stretching and bending states of stress are uncoupled. The equations governing the transverse bending are given by (Reddy, 1984)

$$A_{55}\frac{\partial}{\partial x}\left(\psi_x+\frac{\partial w}{\partial x}\right)+A_{44}\frac{\partial}{\partial y}\left(\psi_y+\frac{\partial w}{\partial y}\right)+q=0$$

$$\frac{\partial}{\partial x}\left(D_{11}\frac{\partial \psi_x}{\partial x}+D_{12}\frac{\partial \psi_y}{\partial y}\right)+D_{66}\frac{\partial}{\partial y}\left(\frac{\partial \psi_x}{\partial y}+\frac{\partial \psi_y}{\partial x}\right) - A_{55}\left(\psi_x+\frac{\partial w}{\partial x}\right)=0$$

$$D_{66}\frac{\partial}{\partial x}\left(\frac{\partial \psi_x}{\partial y}+\frac{\partial \psi_y}{\partial x}\right)+\frac{\partial}{\partial y}\left(D_{12}\frac{\partial \psi_x}{\partial x}+D_{22}\frac{\partial \psi_y}{\partial y}\right) - A_{44}\left(\psi_y+\frac{\partial w}{\partial y}\right)=0 \quad (1)$$

Here w denotes the transverse displacement, ψ_x and ψ_y denote the rotations of a transverse normal about the y and x axes, respectively, q is the distributed transverse load, while A_{ij} $(i, j = 4, 5)$ and D_{ij} $(i, j = 1, 2, 6)$ denote the transverse shear and bending rigidities, respectively.

The Solution Procedure

The Lévy method of solution is applicable to rectangular plates with two opposite edges simply supported and the remaining ones having arbitrary boundary conditions. The method reduces the partial differential equations to ordinary ones. The plate edges $x = 0, a$ are considered simply supported (SS) while the other two edges $(y = \pm b/2)$ have arbitrary edge conditions.

[16]Clifton C. Garvin Professor, Mem. ASME.

[17]Assistant Professor.

[18]Professor.

[19]Department of Engineering Science and Mechanics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

Manuscript received by ASME Applied Mechanics Division, May 16, 1986; final revision January 22, 1987.

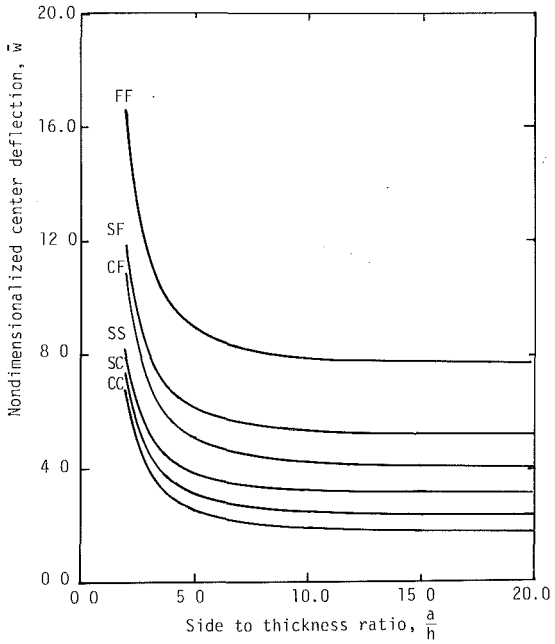


Fig. 2 Nondimensionalized center deflection versus side to thickness ratio of cross-ply (0/90/0) laminates (*a/b* = 1) under uniformly distributed transverse load, q_o

An adequate form of solution to equations (1) that satisfies the boundary conditions $w = 0; M_x = 0; \psi_y = 0$ on edges $x = 0, a$ is:

$$w(x,y) = \sum_{m=1}^{\infty} W_m(y)\sin\frac{m\pi x}{a}$$

$$\psi_x(x,y) = \sum_{m=1}^{\infty} X_m(y)\cos\frac{m\pi x}{a}$$

$$\psi_y(x,y) = \sum_{m=1}^{\infty} Y_m(y)\sin\frac{m\pi x}{a}$$

$$q(x,y) = \sum_{m=1}^{\infty} Q_m(y)\sin\frac{m\pi x}{a} \tag{2}$$

Substituting equations (2) into equations (1), we obtain

$$W_m'' = c_1 W_m + c_2 X_m + c_3 Y_m' + c_o Q_m$$
$$X_m'' = c_4 W_m + c_5 X_m + c_6 Y_m'$$
$$Y_m'' = c_7 W_m' + c_8 X_m' + c_9 Y_m \tag{3}$$

where

$$c_o = -\frac{1}{A_{44}},\ c_1 = \alpha^2\frac{A_{55}}{A_{44}},\ c_2 = c_1/\alpha,$$

$$c_3 = -1,\ c_4 = \alpha\frac{A_{55}}{D_{66}},$$

$$c_5 = \frac{A_{55}+\alpha^2 D_{11}}{D_{66}},\ c_6 = -\alpha\frac{(D_{12}+D_{66})}{D_{66}}$$

$$c_7 = \frac{A_{44}}{D_{22}},\ c_8 = -\frac{D_{66}}{D_{22}}c_6,\ c_9 = \frac{A_{44}+\alpha^2 D_{66}}{D_{22}} \tag{4}$$

and $\alpha \equiv m\pi/a$ while the primes on the variables indicate differentiation with respect to y.

The ordinary differential equations in equation (3) will be solved by using the "state-space concept" (Franklin, 1968). As per this method, we define the (state) variables,

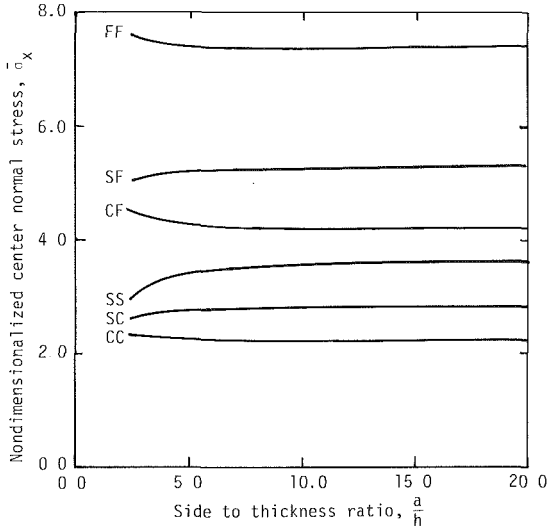


Fig. 3 Nondimensionalized center normal stress versus side to thickness ratio for orthotropic plates (*a/b* = 1) under uniformly distributed transverse load, q_o

$$Z_1 = W_m,\ Z_2 = W_m',\ Z_3 = X_m,\ Z_4 = X_m',$$
$$Z_5 = Y_m,\ Z_6 = Y_m' \tag{5}$$

and rewrite equations (3) in the form,

$$\mathbf{Z}' = A\mathbf{Z} + \mathbf{r} \tag{6}$$

where

$$\mathbf{Z} = \begin{Bmatrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ \cdot \\ Z_6 \end{Bmatrix},\ \mathbf{r} = \begin{Bmatrix} 0 \\ c_o Q_m \\ 0 \\ \\ 0 \\ 0 \end{Bmatrix},\ A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ c_1 & 0 & c_2 & 0 & 0 & c_3 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ c_4 & 0 & c_5 & 0 & 0 & c_6 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & c_7 & 0 & c_8 & c_9 & 0 \end{bmatrix} \tag{7}$$

The solution of equation (6) is given by

$$\mathbf{Z}(y) = e^{Ay}\mathbf{K} + e^{Ay}\int e^{-A\eta}\mathbf{r}\,d\eta \tag{8}$$

where $\mathbf{K}$ is the vector of integration constants, determined from the boundary conditions. The operator e^{Ay} can be expressed in terms of the matrix of eigenvectors $[C]$ and distinct eigenvalues λ_i $(i = 1, 2, \ldots, 6)$ associated with the matrix A:

$$e^{Ay} = [C]\begin{bmatrix} e^{\lambda_1 y} & 0 & & \mathbf{0} \\ & e^{\lambda_2 y} & & \\ & & \cdot & \\ & & & \cdot \\ \mathbf{0} & & & \cdot \\ & & & e^{\lambda_6 y} \end{bmatrix}[C]^{-1} \tag{9}$$

At this point we consider various boundary conditions on the remaining two edges (i.e., edges parallel to the x axis). For free, simply supported and clamped edges (on edges $y = \pm b/2$), we have the following boundary conditions:

Free (F):	$M_y = Q_y = M_{xy} = 0$	
Simply Supported (S):	$w = \psi_x = M_y = 0$	(10)
Clamped (C):	$w = \psi_x = \psi_y = 0$	

Numerical Solutions

A numerical illustration of the solution procedure for two

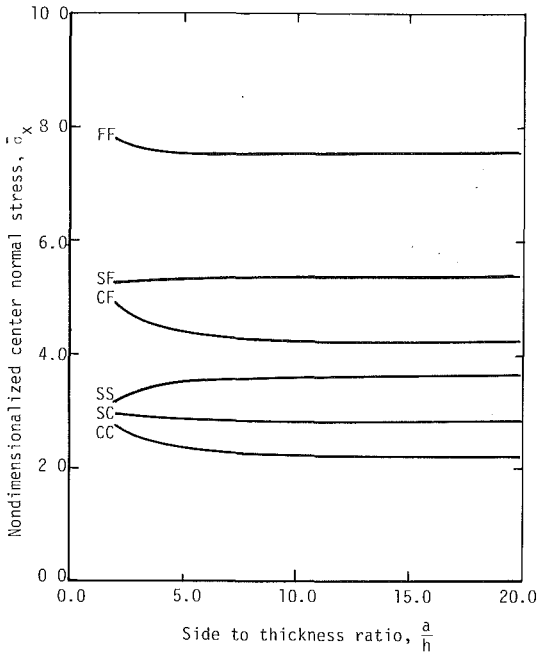


Fig. 4 Nondimensionalized center normal stress versus side to thickness ratio of cross-ply (0/90/0) square (a/b = 1) laminates under uniformly distributed transverse load, q_o

cases, (a) a single layered orthotropic plate and (b) a three layered cross-ply composite plate, is considered. It is assumed that each layer has the following engineering constants used in the literature (Craig and Dawe, 1986):

$$E_1 = 20.83 \times 10^6 \text{ psi}(143.6 \text{ GPa}),\ E_2 = 10.94 \times 10^6 \text{ psi}(75.4 \text{ GPa}),$$

$$G_{12} = 6.10 \times 10^6 \text{ psi}(42.1 \text{ GPa}),\ G_{13} = 3.71 \times 10^6 \text{ psi}(25.6 \text{ GPa}),$$

$$G_{23} = 6.19 \times 10^6 \text{ psi}(42.7 \text{ GPa}),\ \nu_{12} = 0.44$$

The designations SS, CC, FF, SC, SF, and CF will be referred to the edge conditions associated to the edges $y = \pm b/2$ only.

The numerical results presented in Figs. 1-4, indicate the effect of transverse shear deformation on the nondimensionalized transverse deflection, $\bar{w}$ [$\equiv (w(a/2, 0)h^3E_2/q_oa^4) \times 10^2$] and central normal stress, $\bar{\sigma}_{11}$ [$\equiv (\sigma_{11}(a/2, 0, h/2)h^2/q_oa) \times 10$].

From these figures it is clear that the effect of transverse shear deformation is more pronounced on deflections than on stresses and that this effect is increasingly higher for boundary conditions SC, SS, CF, SF, and FF. It is also interesting to note that for certain boundary conditions (SC, SS, and SF) the shear deformation has the effect of decreasing the stresses for thick orthotropic and cross-ply laminates.

Summary

Lévy type solutions of the first-order transverse shear deformation theory of laminated orthotropic composite plates have been developed. The approach is based on the state-space concept which allows one to treat, in a unified manner, various boundary conditions. The numerical examples illustrate the effect of transverse shear deformations on the transverse deflection and central normal stress as well as the influence of the various boundary conditions on the deflections and stresses.

Acknowledgment

The research reported herein was carried under the sponsorship of the Mathematics Division of the Army Research Office through Grant DAAG-29-85-K-0007. The support is gratefully acknowledged.

References

Bert, C. W., and Chen, T. L. C., 1978, "Effect of Shear Deformation on Vibration of Antisymmetric Angle-Ply Laminated Rectangular Plates," *International Journal of Solids and Structures*, Vol. 14, pp. 465–473.

Craig, T. J., and Dawe, D. J., 1986, "Flexural Vibration of Symmetrically Laminated Composite Rectangular Plates Including Transverse Shear Effects," *International Journal of Solids and Structures*, Vol. 22, No. 2, pp. 155–169.

Franklin, J. N., 1968, *Matrix Theory*, Prentice-Hall, Englewood Cliffs, New Jersey.

Lévy, M., 1989, "Sur l'Equilibre Elastique d'une Plaque Rectangulaire," *Comptes Rendus De L'Academie Des Sciences*, Vol. 129, pp. 535–539.

Librescu, L., 1975, *Elastostatics and Kinetics of Anisotropic and Heterogeneous Shell-Type Structures*, Noordhoff, Leyden, The Netherlands.

Reddy, J. N., 1980, "A Penalty Plate-Bending Element for the Analysis of Laminated Anisotropic Plates," *International Journal Numerical Methods in Engineering*, Vol. 15, pp. 1187–1206.

Reddy, J. N., 1984, *Energy and Variational Methods in Applied Mechanics*, Wiley, New York.

Reddy, J. N., and Chao, W. C., 1981, "A Comparison of Closed Form and Finite Element Solutions of Thick Laminated Anisotropic Rectangular Plates," *Nuclear Engineering and Design*, Vol. 64, pp. 153–167.

Whitney, J. M., and Leissa, A. W., 1969, "Analysis of Heterogeneous Anisotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 261–266.

Whitney, J. M., and Pagano, N. J., 1970, "Shear Deformation in Heterogeneous Anisotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 1031–1036.

Diffusion in Hydromagnetic Oscillatory Flow Through a Porous Channel

A. Ramachandra Rao[20] and K. S. Deshikachar[21]

Introduction

It has been established that the axial dispersion of a contaminant through tubes of circular cross section under both steady (Taylor, 1953) and oscillatory (Watson, 1983) laminar flow conditions is considerably larger than in the absence of flow. Further, the dispersion in an oscillatory flow has been found to be comparable in magnitude to values found in Taylor diffusion under steady state conditions when the flow oscillation period is comparable with the radial diffusion of the contaminent but becomes much less efficient when the Womersley number is either very large or small. An analogous problem in heat conduction for an oscillatory flow in a channel has been studied by Kurzweg (1985) and an enhancement of effective diffusivity has been observed in the presence of oscillations.

The studies of longitudinal dispersion in the flow in a parallel plate channel with permeable walls for hydrodynamic and hydromagnetic cases have been carried out by Eroshenko and Zaichik (1980) and Annapurna et al. (1985), respectively, highlighting the importance of these studies in practical situations. Here, we use Watson's (1983) exact analysis to study the dispersion of a solute matter in an oscillatory flow of an electrically conducting fluid in a channel with porous walls in the presence of a transverse magnetic field.

Formulation

Consider the laminar oscillatory flow of an electrically conducting, incompressible viscous fluid in a channel with porous walls situated at $y = \pm h$ with an x axis along the center line of the channel. The plates $y = \pm h$ are subjected to uniform injection and suction, respectively, with the same velocity v_0 and a uniform magnetic field B_0 is applied perpendicular to the walls. The equations governing the unsteady flow under

[20]Department of Applied Mathematics, Indian Institute of Science, Bangalore 560012, India.

[21]Department of Mathematics, K.R.E.C., Surathkal 574157, India.

Manuscript received by ASME Applied Mechanics Division, September 8, 1986; final revision, January 20, 1987.

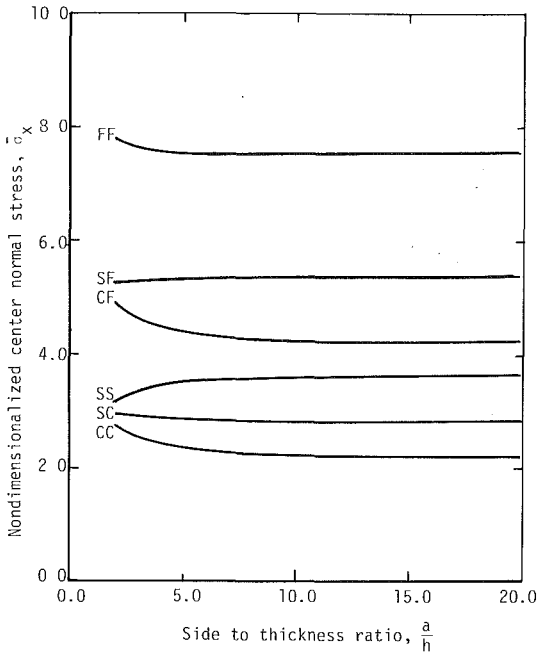


Fig. 4 Nondimensionalized center normal stress versus side to thickness ratio of cross-ply (0/90/0) square (a/b = 1) laminates under uniformly distributed transverse load, q_o

cases, (a) a single layered orthotropic plate and (b) a three layered cross-ply composite plate, is considered. It is assumed that each layer has the following engineering constants used in the literature (Craig and Dawe, 1986):

$$E_1 = 20.83 \times 10^6 \text{ psi}(143.6 \text{ GPa}),\ E_2 = 10.94 \times 10^6 \text{ psi}(75.4 \text{ GPa}),$$

$$G_{12} = 6.10 \times 10^6 \text{ psi}(42.1 \text{ GPa}),\ G_{13} = 3.71 \times 10^6 \text{ psi}(25.6 \text{ GPa}),$$

$$G_{23} = 6.19 \times 10^6 \text{ psi}(42.7 \text{ GPa}),\ \nu_{12} = 0.44$$

The designations SS, CC, FF, SC, SF, and CF will be referred to the edge conditions associated to the edges $y = \pm b/2$ only.

The numerical results presented in Figs. 1-4, indicate the effect of transverse shear deformation on the nondimensionalized transverse deflection, $\bar{w}$ [$\equiv (w(a/2, 0)h^3E_2/q_oa^4) \times 10^2$] and central normal stress, $\bar{\sigma}_{11}$ [$\equiv (\sigma_{11}(a/2, 0, h/2)h^2/q_oa) \times 10$].

From these figures it is clear that the effect of transverse shear deformation is more pronounced on deflections than on stresses and that this effect is increasingly higher for boundary conditions SC, SS, CF, SF, and FF. It is also interesting to note that for certain boundary conditions (SC, SS, and SF) the shear deformation has the effect of decreasing the stresses for thick orthotropic and cross-ply laminates.

Summary

Lévy type solutions of the first-order transverse shear deformation theory of laminated orthotropic composite plates have been developed. The approach is based on the state-space concept which allows one to treat, in a unified manner, various boundary conditions. The numerical examples illustrate the effect of transverse shear deformations on the transverse deflection and central normal stress as well as the influence of the various boundary conditions on the deflections and stresses.

Acknowledgment

The research reported herein was carried under the sponsorship of the Mathematics Division of the Army Research Office through Grant DAAG-29-85-K-0007. The support is gratefully acknowledged.

References

Bert, C. W., and Chen, T. L. C., 1978, "Effect of Shear Deformation on Vibration of Antisymmetric Angle-Ply Laminated Rectangular Plates," *International Journal of Solids and Structures*, Vol. 14, pp. 465–473.

Craig, T. J., and Dawe, D. J., 1986, "Flexural Vibration of Symmetrically Laminated Composite Rectangular Plates Including Transverse Shear Effects," *International Journal of Solids and Structures*, Vol. 22, No. 2, pp. 155–169.

Franklin, J. N., 1968, *Matrix Theory*, Prentice-Hall, Englewood Cliffs, New Jersey.

Lévy, M., 1989, "Sur l'Equilibre Elastique d'une Plaque Rectangulaire," *Comptes Rendus De L'Academie Des Sciences*, Vol. 129, pp. 535–539.

Librescu, L., 1975, *Elastostatics and Kinetics of Anisotropic and Heterogeneous Shell-Type Structures*, Noordhoff, Leyden, The Netherlands.

Reddy, J. N., 1980, "A Penalty Plate-Bending Element for the Analysis of Laminated Anisotropic Plates," *International Journal Numerical Methods in Engineering*, Vol. 15, pp. 1187–1206.

Reddy, J. N., 1984, *Energy and Variational Methods in Applied Mechanics*, Wiley, New York.

Reddy, J. N., and Chao, W. C., 1981, "A Comparison of Closed Form and Finite Element Solutions of Thick Laminated Anisotropic Rectangular Plates," *Nuclear Engineering and Design*, Vol. 64, pp. 153–167.

Whitney, J. M., and Leissa, A. W., 1969, "Analysis of Heterogeneous Anisotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 261–266.

Whitney, J. M., and Pagano, N. J., 1970, "Shear Deformation in Heterogeneous Anisotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 1031–1036.

Diffusion in Hydromagnetic Oscillatory Flow Through a Porous Channel

A. Ramachandra Rao[20] and K. S. Deshikachar[21]

Introduction

It has been established that the axial dispersion of a contaminant through tubes of circular cross section under both steady (Taylor, 1953) and oscillatory (Watson, 1983) laminar flow conditions is considerably larger than in the absence of flow. Further, the dispersion in an oscillatory flow has been found to be comparable in magnitude to values found in Taylor diffusion under steady state conditions when the flow oscillation period is comparable with the radial diffusion of the contaminent but becomes much less efficient when the Womersley number is either very large or small. An analogous problem in heat conduction for an oscillatory flow in a channel has been studied by Kurzweg (1985) and an enhancement of effective diffusivity has been observed in the presence of oscillations.

The studies of longitudinal dispersion in the flow in a parallel plate channel with permeable walls for hydrodynamic and hydromagnetic cases have been carried out by Eroshenko and Zaichik (1980) and Annapurna et al. (1985), respectively, highlighting the importance of these studies in practical situations. Here, we use Watson's (1983) exact analysis to study the dispersion of a solute matter in an oscillatory flow of an electrically conducting fluid in a channel with porous walls in the presence of a transverse magnetic field.

Formulation

Consider the laminar oscillatory flow of an electrically conducting, incompressible viscous fluid in a channel with porous walls situated at $y = \pm h$ with an x axis along the center line of the channel. The plates $y = \pm h$ are subjected to uniform injection and suction, respectively, with the same velocity v_0 and a uniform magnetic field B_0 is applied perpendicular to the walls. The equations governing the unsteady flow under

[20]Department of Applied Mathematics, Indian Institute of Science, Bangalore 560012, India.

[21]Department of Mathematics, K.R.E.C., Surathkal 574157, India.

Manuscript received by ASME Applied Mechanics Division, September 8, 1986; final revision, January 20, 1987.

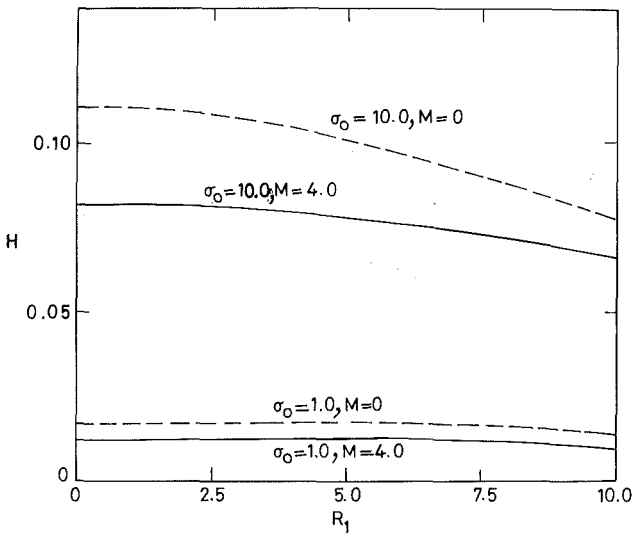


Fig.1 Variation of H with R_1 for a fixed W_0 and for different values of σ_0 and M

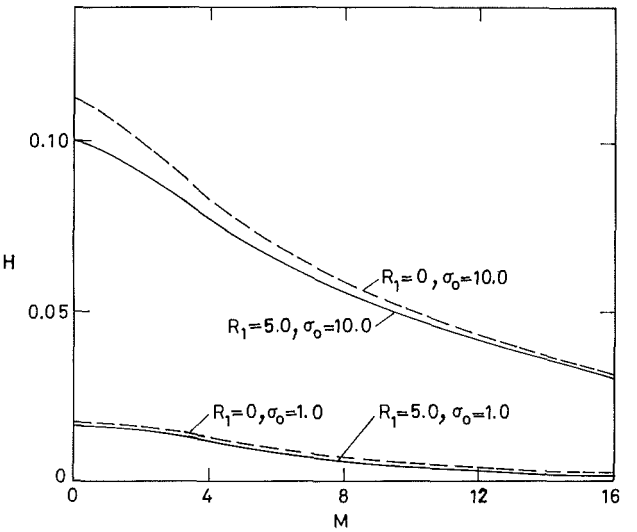


Fig. 2 Variation of H with M, for a fixed W_0 and for different values of σ_0 and R_1

the usual magnetohydrodynamic approximations reduce to a single equation given by

$$u_t + v_0 u_y = -p_x/\rho + \nu u_{xx} - \sigma B_0^2 u/\rho, \tag{1}$$

where $u(x,t)$ is the x component of velocity, ρ is the density, p is the pressure, ν is the kinematic coefficient of viscosity, σ is the electrical conductivity, t is the time, and subscripts denote partial differentiation with respect to the variable. Here we have assumed that the magnetic Reynolds number is small.

By taking $p_x = -P\cos\omega t$ and $u = \mathrm{Re}(f(y)e^{i\omega t})$, in equation (1), we get

$$\alpha^2 f/\rho + v_0 f_y = P/\rho + \nu f_{yy}, \quad \alpha^2 = i\omega\rho + \sigma B_0^2. \tag{2}$$

The no-slip boundary condition at the walls give $f = 0$. It is assumed that the diffusing substance is a passive contaminant for which the concentration is very small so that the physical properties of the fluid and the diffusivity k of the contaminant are taken as constants. Further, the walls of the channel are impermeable for the contaminant. The govening equation and boundary conditions for the concentration $\theta(x, y, t)$ of the contaminant are given by

$$\theta_t + u\theta_x = k(\theta_{xx} + \theta_{yy}), \quad \theta_y = 0 \text{ at } y = \pm h. \tag{3}$$

Equation (3) has an exact solution of the form, Taylor (1953), $\theta = -\gamma x + \gamma \mathrm{Re}(g(y)e^{i\omega t})$, where γ is a constant and this on substitution in equation (3) gives

$$i\omega g - f = k g_{yy}, \quad g_y = 0 \text{ at } y = \pm h. \tag{4}$$

Introducing nondimensional quantities:

$$f = -P\alpha^{-2}[F(\eta) - 1], \quad g = iP\alpha^{-2}[G(\eta) - 1],$$

$$y = \eta h, \tag{5}$$

in equations (2) and (4), we get

$$F_{\eta\eta} + R_1 F_\eta - \alpha_1^2 F = 0, \quad F = 1 \text{ at } \eta = \pm 1. \tag{6}$$

$$G_{\eta\eta} = \sigma_1^2 (G - F), \quad G_\eta = 0 \text{ at } \eta = \pm 1. \tag{7}$$

where $R_1 = v_0 h/\nu$ (suction Reynolds number), $\alpha_1^2 = M^2 + iW_0^2$, $M^2 = \sigma B_0^2 h^2/\rho\nu$ (Hartmann number), $W_0 = \omega h^2/\nu$ (Womersley number), and $\sigma_1^2 = i\sigma_0 W_0^2$, $\sigma_0 = \nu/k$ (Schmidt number).

Exact Solution

The solutions of equations (6) and (7) are given by

$$F(\eta) = [\sinh \zeta_2 \cdot e^{\zeta_1 \eta} + \sinh \zeta_1 \cdot e^{-\zeta_2 \eta}] \tag{8}$$

$$G(\eta) = A_1 e^{\sigma_1 \eta} + B_1 e^{-\sigma_1 \eta} + A_2 e^{\zeta_1 \eta} + B_2 e^{-\zeta_2 \eta}. \tag{9}$$

where

$$\zeta_1 = \xi_0 + \eta_0, \quad \zeta_2 = \xi_0 - \eta_0,$$

$$\xi_0 = (R_1^2 + 4\alpha_1^2)^{1/2}/2, \quad \eta_0 = R_1/2, \tag{10}$$

A_1, A_2, B_1, and B_2 are known constants, the expressions of which are omitted for briefness. From equation (9), we observe that G = constant on the boundary walls. The effective diffusivity K in an oscillatory flow is (Watson, 1983) $K = k(1 + R)$, where the relative increase in the flux R is given by

$$8khR = \int_{-h}^{h} (f\bar{g} + \bar{f}g)\, dy, \tag{11}$$

with bars denoting complex conjugates. Using equation (5) in equation (11), we obtain

$$R = iP^2\sigma_0 (8\nu\omega\alpha^2\alpha^{-2})^{-1} \int_{-1}^{1} (F\bar{G} - \bar{F}G + G - \bar{G} - F + \bar{F})\, d\eta. \tag{12}$$

Substituting the values of F and G from equations (8) and (9) in equation (12), we have

$$R = iP^2\sigma_0^2 h^6 (2\alpha_1 \rho\nu^2 W_0)^{-2} (P_1 - \bar{P}_1), \tag{13}$$

where P_1 is a known complex constant, the expression of which is omitted for briefness. Following Watson (1983) the expression for tidal volume is

$$V_1^2 = 16P^2 h^2 Q_1 (\omega^2 |\alpha^2|)^{-2}, \tag{14}$$

where

$$Q_1 = |(2\xi_0 \sinh \zeta_1 \sinh \zeta_2 - \alpha_1^2 \sinh 2\xi_0)/\alpha_1^2 \sinh 2\xi_0|^2.$$

Now using equation (14) in equation (13), we get

$$R = h^{-4} V_1^2 H(W_0, M, R, \sigma_0),$$

where

$$H = Rh^4 V_1^{-2} = \sigma_1^2 (P_1 - \bar{P}_1)/64 Q_1. \tag{15}$$

Taking the limit $R_1 \to 0$ and $W_1 \to 0$, we observe that the above results coincide with the results presented by Watson (1983) for a plane channel.

Discussion

The value of H, which gives qualitatively the effective diffusivity in an oscillatory flow, is computed from equation (15) for various values of the parameters R_1, M, W_0, and σ_0. Figure 1 shows the variation of H with R_1 for $W_0 = 0.5$ for different values of M and σ_0. It is observed that the value of H decreases with an increase in M or R_1, whereas it increases with an increase in σ_0 for fixed values of W_0 M, and R_1. H as a function of M for $W_0 = 0.5$ and for various values of R_1 and σ_0 is presented in Fig. 2. It is interesting to observe that H decreases rapidly with an increase in M for large values of σ_0. This is in accordance with the physical expectation as an increase in Hartmann number flattens the velocity profile.

Acknowledgment

The authors would like to thank the referees and Professor Z. Warhaft for their useful suggestions.

References

Annapurna, N., Gupta, A. S., and Dandapat, B. S., 1985, "Hydromagnetic Convective Diffusion Between Parallel Plates With Suction," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 213–215.

Eroshenko, V. M., And Zaichik, L. I., 1980, "Longitudinal Diffusion of an Impurity in a Pipe With Permeable Walls," Translated from *Inzhenesno Fizicheskie Zhurnal*, Vol. 38, pp. 314–317.

Kurzweg, U. H., 1985, "Enhanced Heat Conduction in Oscillating Viscous Flows Within Parallel Plate Channels," *J. Fluid Mech.*, Vol. 156, pp. 291–300.

Taylor, G. I., 1953, "Dispersion of Soluble Matter in Solvent Flowng Slowly Through a Tube," *Proc. Roy. Soc. London*, Sec. A, Vol. 219, pp. 176–203.

Watson, E,. J., 1983, "Diffusion in Oscillatory Pipe Flows," *J. Fluid Mech.*, Vol. 133, pp. 233–244.

Journal of Applied Mechanics

Discussion

Frequency Response of a Nonlinear Pneumatic System[1]

H. M. Paynter[2] and E. P. Fahrenthold.[3] The writers note with great interest this work by the authors on an important gas compression process. However, we were also dismayed to find no citation of research on very similar systems conducted over a period of nearly two decades by David Otis and colleagues at the University of Wisconsin (Otis, 1970, 1974; Otis and Pourmovahed, 1985). Also the recent paper of Kagawa (1985) is of value to readers and investigators in this area.

Unfortunately the authors' brief reference to *Fluid Power Control* (edited by Blackburn, Reethof, and Shearer, 1960) omitted mentioning that nearly thirty-five years ago Dr. Lowen Shearer and others at the M.I.T. Dynamic Analysis and Control Laboratory (DACL) employed shunt tank-resistor mass transfer dampers to stabilize pneumatic ram positioning servos. The diagram on page 550 of this last cited reference is in fact strikingly similar to Fig. 1 of the authors' paper. This earlier system was built and tested, as well as analyzed from both linear and nonlinear viewpoints. The fact that Dr. Richard Booton, while at the DACL, extended Kochenberger's sinusoidal describing function linearization techniques to complete harmonic balance methods and to random inputs as well, is a strong indication that the DACL employed nonlinear methods.

Our own continuing concern with nonlinear pneumatic dynamics is indicated by some very recent papers (Paynter et al., 1987a, 1987b; Paynter and Fahrenthold, 1987). Based on the earlier efforts of Otis and coworkers, we have established that it is thermodynamically inconsistent to assume a polytropic exponent "n" without at the same time incorporating a conjugated dissipative, phase-shifted process. If the authors had in fact constructed and tested their system, this flaw in their model would have become quite apparent.

References

Blackburn, J. F., Reethof, G., and Shearer, J. L., eds., 1960, *Fluid Power Control*, M.I.T. Press, Cambridge, MA.

Kagawa, T., 1985, "Heat Transfer Effects on the Frequency Response of a Pneumatic Flapper Nozzle," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 107, pp. 332–336.

Otis, D. R., 1970, "Thermal Damping in Gas-Filled Composite Materials During Impact Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 38–43.

Otis, D. R., 1974, "New Developments in Predicting and Modifying Performance of Hydraulic Accumulators," *Proceedings of the National Conference on Fluid Power*, Vol. 28, pp. 473–489.

Otis, D. R., and Pourmovahed, A., 1985, "An Algorithm for Computing Nonflow Gas Processes in Gas Springs and Hydropneumatic Accumulators," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 107, pp. 93–96.

Paynter, H. M., and Fahrenthold, E. P., 1987, "On the Nonexistence of Simple Polytropes and Other Thermodynamic Consequences of the Dispersion Relation," submitted to the session on Thermodynamics, Heat and Mass Transfer in Biotechnology, 1987 ASME Winter Annual Meeting, Boston, MA.

Paynter, H. M., Fahrenthold, E. P., and Rotz, C. A., 1987a, "Linearizing Transformations for the Thermal Dynamics of Gas Compression Processes," *Proceedings of the 1987 American Control Conference*, Minneapolis, MN.

Paynter, H. M., Fahrenthold, E. P., and Rotz, C. A., 1987b, "Wall Heat Transfer and Storage Effects on the Thermal Dynamics of Gas Compression Processes," in preparation.

Authors' Closure

The authors are familiar with the literature cited by the discussers; nevertheless, these references, along with some other related publications on pneumatic, hydraulic, and electromechanical systems, had to be excluded. The literature review included only those references most relevant to the main focus of the paper, which was on the construction of an approximate solution to the nonlinear orifice problem. In an earlier publication (Wang and Singh, ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 956–958, Dec. 1986), a pneumatic chamber with nonlinearities due to compressibility and friction was analyzed using again the harmonic balance technique.

Figure 1 was chosen as a generic example case, which could represent many physical systems including Shearer's research work on continuous motion control. To the best of the authors' knowledge, no one, including the researchers cited by the discussers, has analyzed the same fundamental problem using the same analytical tool. Hence, the authors' work is believed to supplement the available literature.

The polytropic index employed in this paper essentially indicates that the thermodynamic process is somewhere in the isothermal to isentropic regime. In many practical systems and components, it is closer to the isentropic index. The issue of using the polytropic model has also been addressed by Singh and his coworkers with application to a reciprocating compressor (see, for instance, *ASHRAE Transactions*, Vol. 92, part 1, pp. 250–258, 1986).

A Rate-Independent Constitutive Theory for Finite Inelastic Deformation[4]

R. Hill[5] and J. R. Rice.[6] Hill (1968) showed that the

[1] By Y.-T. Wang and R. Singh and published in the March, 1987, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 209–214.

[2] Emeritus Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

[3] Assistant Professor, Department of Mechanical Engineering, University of Texas, Austin, TX 78712.

[4] By M. M. Carroll and published in the March, 1987, issue of ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, No. 1, pp. 15–21.

[5] Professor (emeritus), Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, CB39EW England.

[6] Professor, Division of Applied Sciences, Harvard University, Cambridge, MA 02138. Fellow ASME.

$$\delta\Gamma(\mathbf{S}) = \delta\mathbf{S} \cdot [\partial\Gamma(\mathbf{S})/\partial\mathbf{S}]_{\mathbf{S}=\mathbf{S}_*} < 0 \qquad (2')$$

remains valid with the same interpretation of $\delta\mathbf{S}$, as pointing locally into the elastic domain from $\mathbf{S}_*$, as earlier. Also, since $d^p\mathbf{E} = \partial(d^p\Psi)/\delta\mathbf{S}$ at $\mathbf{S} = \mathbf{S}_*$, one has

$$\text{limit}[d^p\mathbf{E}/\,|d\mathbf{K}\,|] = [\partial\Gamma(\mathbf{S})/\partial\mathbf{S}]_{\mathbf{S}=\mathbf{S}_*}$$

This shows that inequality (2′) requires

$$\delta\mathbf{S} \cdot \{\text{limit}[d^p\mathbf{E}/\,|d\mathbf{K}\,|]\} < 0 \qquad (3')$$

which clearly expresses normality.

References

Hill, R., 1968, "On Constitutive Inequalities for Simple Materials – II," *J. Mech. Phys. Solids*, Vol. 16, pp. 315–322.

Hill, R., and Rice, J. R., 1973, "Elastic Potentials and the Structure of Inelastic Constitutive Laws," *SIAM J. Appl. Math.*, Vol. 25, pp. 448–461.

Il'yushin, A. A., 1961, "On a Postulate of Plasticity," *Prikl. Mat. Mech.*, Vol. 25, pp. 503–507.

Naghdi, P. M., and Trapp, J. A., 1975, "Restrictions on Constitutive Equations of Finitely Deformed Elastic-Plastic Materials," *Q. J. Mech. Appl. Math.*, Vol. 28, pp. 25–46.

Author's Closure

The substance of the Discussion is my statement (Carroll, 1987) that Hill and Rice (1973) did not prove that Il'yushin's postulate implies normality and that they did not claim to have done so. The Discussion clearly disposes of the latter claim, so that we can concentrate on the former.

Apart from the issue of the starting point, discussed below, the debate is not about the validity of the result, rather it pertains to the validity of the proof by Hill and Rice. On p. 454 of their 1973 paper they established the inequality (1) in their Discussion, and its strain space counterpart, by considering finite cycles that involve a first order inelastic history $d\mathbf{K}$, which is "accumulated at a state $\mathbf{E}_*$, say, on a yield surface, whereas the path starts from some *finitely distant state E* within the yield surface." I have changed the notation to conform with that in may 1987 paper and the Discussion, and the italics are mine. I argued that the inequality should not then be used in the limit $\mathbf{E}_O \to \mathbf{E}_*$ to obtain normality, because to do so would reduce the expression in (1) to second order and that second order terms were ignored in the derivation of (1). Rice and Hill correctly state that I tacitly assumed that the differentials $d\mathbf{K}$ and $\partial\mathbf{E} = \mathbf{E}_* - \mathbf{E}_O$ are of the same order. In view of the above quotation from their 1973 paper, and the fact that different orders of differentials are not mentioned there, I do not agree that my unhappiness was "self-generated."

I have never been comfortable with differentials, except in the context of limiting values and linear relations. Hill and Rice (1973, p. 451) appear to share this view, to the extent that even though they state that $d^p\varphi$ is defined as the change in the function ϕ (**E**, **K**) when **K** is changed to $\mathbf{K} + d\mathbf{K}$, so that

$$d^p\varphi = \varphi(\mathbf{E}, \mathbf{K} + d\mathbf{K}) - \varphi(\mathbf{E}, \mathbf{K}),$$

it is clear that from the context (cf their equation (9)) that this equality holds only to first order, i.e., that $d^p\phi$ is really defined as the first order change in φ. (The same is true of the various relations indicated as identities in the Discussion.) For this reason, I am very uncomfortable with powers and products of differentials, and especially so with the notion of differentials which are arbitrarily large compared with other differentials.

The Discussion by Rice and Hill certainly clarifies the issue. My bias against this use of differentials, which the reader may or may not share, makes me still prefer the normality proof of Naghdi and Trapp (1975) and the shorter proof of Casey (1984), which are based on the evaluation of work integrals for limits of sequences of finite deformation cycles.

With regard to the comment in the Discussion that the work assumption of Naghdi and Trapp (1975) is apparently the same as the Il'yushin postulate, I should mention that the differences between the two are discussed in their 1975 paper and need not be repeated here. In essence, Naghdi and Trapp point out that the Il'yushin postulate is a condition on the integral of the stress power in a cycle of infinitesimal deformation, since this is the context of Ilyushin's (1961) paper, while the work assumption is a condition on the external work on a body in a cycle of finite homogeneous deformation, which is not necessarily quasistatic.

I take this opportunity to comment on a remark in the Introduction of my 1987 paper to the effect that the theory presented there "is stable in the sense that the external work needed to carry a body through any closed cycle of deformation is nonnegative." This remark is misleading since it suggests that the normality condition is *sufficient* to ensure nonnegative work. In fact it is not sufficient; further necessary conditions are set down by Naghdi and Trapp (1975) and by Casey (1984), for example, and I am not aware that sufficient conditions have been established.

References

Il'yushin, A. A., 1961, "On a Postulate of Plasticity," *Prikl. Mat Mech.*, Vol. 25, pp. 503–507.

Hill, R., and Rice, J. R., 1973 "Elastic Potentials and the Structure of Inelastic Constitutive Laws," *SIAM. J. Appl. Math.*, Vol. 25, pp. 448–461.

Naghdi, P. M., and Trapp, J. A., 1975, "Restrictions on Constitutive Equations of Finitely Deformed Elastic-Plastic Materials," *Q.J. Mech. Appl. Math.*, Vol. 28, pp. 25–46.

Casey, J., 1984, "A Simple Proof of a Result in Finite Plasticity," *Q. Appl. Math.*, Vol. 42, pp. 61–71.

Carroll, M. M., 1987, "A Rate-Independent Constitutive Theory for Finite Inelastic Deformation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 15–21.

A Composite Plate Theory for Arbitrary Laminate Configurations[7]

Y. Y. Yu.[8] The authors should be congratulated for having made another contribution to the theory of laminated plates, by using Reissner's new mixed variational principle and by adopting simultaneously the assumptions of piecewise linear continuous displacements and quadratic transverse shear stress distributions. In doing so, they were rather modest in stating that their proposed theory possesses two main drawbacks: first, the number of equilibrium equations and edge boundary conditions increases with the number of layers; and second, no clear physical meaning seems to be associated with the coupled natural edge boundary conditions. These are actually characteristics of a layered-plate theory derived by assuming piecewise linear continuous displacements and, as such, are inherent with the mathematical modeling of the physical system. In this case the laminated plate is a physical system whose components are the individual laminates. Although the system is recognized as being intrinsically more complex than its components, we would still like to formulate mathematical theories for treating laminated plates in the same manner as we have always treated single-layered plates. Of course this is not always possible, unless we are willing to compromise and accept oversimplified theories.

Indeed, the above characteristics of the laminated plate theory need not be drawbacks. In vibration analysis, as the number of equations of motion (and edge boundary conditions) increases with the number of layers, the theory may just provide the proper ingredients that are needed to analyze the modes and to cover the frequency range adequately, as has been demonstrated in the past.

[7] By A. Toledano and H. Murakami and published in the March, 1987, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 181–189.

[8] Professor of Mechanical Engineering, New Jersey Institute of Technology, Newark, NJ 07102. Fellow ASME.