Transactions of the ASME

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 47 • NUMBER 4 • DECEMBER 1980

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

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

PRICES: To members, $30.00, annually; to nonmembers, $60.00. Single copies $20.00 each. Add $5.00 for postage to countries outside the United States and Canada.

INDEXED by the Engineering Index, Inc.

TECHNICAL PAPERS

(Contents continued on Page 984)

CONTENTS (CONTINUED)

BRIEF NOTES

(Contents continued on Inside Back Cover)

CONTENTS (CONTINUED)

DISCUSSIONS

BOOK REVIEWS

W. S. Yeung[1]

Department of Mechanical Engineering,
University of California,
Berkeley, Calif. 94720

Laminar Boundary-Layer Flow Near the Entry of a Curved Circular Pipe

The fluid mechanics of a viscous, incompressible fluid entering a circular curved pipe at large Reynolds number is investigated numerically. The flow field is divided into two regions, the boundary-layer region and the inviscid core region. The boundary layer is assumed to be laminar and the method of integral relations is used to solve the governing equations. The core region, on the other hand, is assumed irrotational and is solved by a modified version of Telenin's method. The coupling of the two regions is accounted for through the imposition of the outer edge normal velocity as a boundary condition for the core region. Results are presented for a Reynolds number of 10^4 and a curvature ratio of 0.1. It has been shown that the cross flow in the core region is initially directed from the outer to the inner bend and reverses its direction downstream for the entry condition of uniform axial motion. The core velocity profiles are consistent with a recent experimental investigation, while the boundary-layer results are qualitatively similar to those of Yao and Berger, although a direct quantitative comparison is not applicable, due to the different range of Reynolds number considered.

Introduction

Curved pipe flow has long been a subject of interest in fluid mechanics. Most of the earlier literature deals with the fully developed region of the flow [1–5]. Earlier analyses in the entry region were mainly made using an inviscid rotational model developed by Hawthrone [6]. The first complete analyses of a viscous fluid flowing into a curved pipe were, to the author's knowledge, those of Yao and Berger [7] and Singh [8]. Most recently, Agrawal, Talbot, and Gong [9] have measured experimentally the velocity profiles in the entry region. They found that the initial uniform velocity profile changes to a potential vortex profile immediately downstream of the entry section.

In the present study, we investigate the entry flow problem at large Reynolds number of the order of 10^4 for a uniform entry condition. The physical situation corresponds to a gas leaving a large reservoir and entering a 90° elbow. We shall only discuss the solution in the region where the boundary layer is laminar and attached. In view of the large Reynolds number considered, the boundary layer will eventually become turbulent and the present method must be modified to take account of turbulent effects. For the core region, the flow is always treated as inviscid regardless of the nature of the boundary layer. Furthermore, we shall assume that the core flow is irrotational.

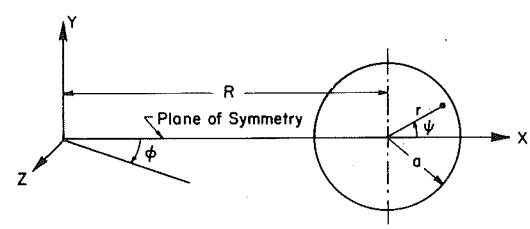


Fig. 1 Toroidal coordinate system

Governing Equations

For the sake of completeness, we give in this section the full Navier-Stokes equation for the motion in the curved pipe. The coordinate system is shown in Fig. 1. R is the radius of curvature of the pipe axis, a is the radius of the cross section, and u, v, w are the velocity components in the direction of increasing r, ψ, ϕ, respectively. Thus, for the r-momentum,

[1] Present address: Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire 03755.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N. Y. 10017, and will be accepted until March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, May, 1980.

$$u\frac{\partial u}{\partial r}+\frac{v}{r}\frac{\partial u}{\partial \psi}+\frac{w}{R+r\cos\psi}\frac{\partial u}{\partial \phi}-\frac{v^2}{r}-\frac{w^2\cos\psi}{R+r\cos\psi}=-\frac{1}{\rho}\frac{\partial p}{\partial r}$$
$$+\nu\left\{\frac{1}{(R+r\cos\psi)^2}\frac{\partial}{\partial\phi}\left[\frac{\partial u}{\partial\phi}-w\cos\psi-(R+r\cos\psi)\frac{\partial w}{\partial r}\right]\right.$$
$$\left.-\left(\frac{1}{r}\frac{\partial}{\partial\psi}-\frac{\sin\psi}{R+r\cos\psi}\right)\left(\frac{\partial v}{\partial r}+\frac{v}{r}-\frac{1}{r}\frac{\partial u}{\partial\psi}\right)\right\}, \quad (1)$$

for the ψ-momentum

$$u\frac{\partial v}{\partial r}+\frac{v}{r}\frac{\partial v}{\partial\psi}+\frac{w}{R+r\cos\psi}\frac{\partial v}{\partial\phi}+\frac{uv}{r}+\frac{w^2\sin\psi}{R+r\cos\psi}$$
$$=-\frac{1}{r\rho}\frac{\partial p}{\partial\psi}+\nu\left\{\left(\frac{\partial}{\partial r}+\frac{\cos\psi}{R+r\cos\psi}\right)\left(\frac{\partial v}{\partial r}+\frac{v}{r}-\frac{1}{r}\frac{\partial u}{\partial\psi}\right)\right.$$
$$\left.-\frac{1}{(R+r\cos\psi)^2}\frac{\partial}{\partial\phi}\left[\frac{R+r\cos\psi}{r}\frac{\partial w}{\partial\psi}-w\sin\psi-\frac{\partial v}{\partial\phi}\right]\right\}, \quad (2)$$

and for the ϕ-momentum

$$u\frac{\partial w}{\partial r}+\frac{v}{r}\frac{\partial w}{\partial\psi}+\frac{w}{R+r\cos\psi}\frac{\partial w}{\partial\phi}+\frac{uw\cos\psi}{R+r\cos\psi}-\frac{vw\sin\psi}{R+r\cos\psi}$$
$$=-\frac{1}{R+r\cos\psi}\frac{1}{\rho}\frac{\partial p}{\partial\phi}+\nu\left\{\left(\frac{\partial}{\partial r}+\frac{1}{r}\right)\right.$$
$$\times\left[\frac{\partial w}{\partial r}-\frac{1}{R+r\cos\psi}\left(\frac{\partial u}{\partial\phi}-w\cos\psi\right)\right]$$
$$\left.+\frac{1}{r}\frac{\partial}{\partial\psi}\left[\frac{1}{r}\frac{\partial w}{\partial\psi}\frac{1}{R+r\cos\psi}\left(\frac{\partial v}{\partial\phi}+w\sin\psi\right)\right]\right\}. \quad (3)$$

The continuity equation is given by

$$\frac{\partial(ur)}{\partial r}+\frac{\partial v}{\partial\psi}+\frac{r}{R+r\cos\psi}\frac{\partial w}{\partial\phi}+\frac{ur\cos\psi-vr\sin\psi}{R+r\cos\psi}=0. \quad (4)$$

In equations (1)–(3), p denotes the pressure, ρ the density, and ν the kinematic viscosity of the gas as usual. The domain of interest is

$$0<r<1,\quad 0<\psi<\pi,\quad 0<\phi<\pi/2, \quad (5)$$

with the planes $\psi=0$ and $\psi=\pi$ being the planes of symmetry. The boundary conditions consist of the no-slip conditions at the wall, uniform entry condition at $\phi=0$ and required exit condition at $\phi=\pi/2$.

Boundary-Layer Region

By means of an order-of-magnitude analysis, one can readily obtain the boundary-layer equations from (1)–(4). We introduce the following dimensionless quantities:

$$U=\frac{uD^{1/2}}{w_e},\quad V=\frac{v}{w_e},\quad W=\frac{w}{w_e},\quad P=\frac{p}{\rho W_i{}^2},\quad W_e=\frac{w_e}{W_i},$$
$$\eta=(a-r)D^{1/2}/a,\quad \psi=\psi,\quad s=\frac{\phi}{\alpha}, \quad (6)$$

where D is the Dean number given by

$$D=\mathrm{Re}\sqrt{\alpha},\quad \mathrm{Re}=\frac{W_i a}{\nu} \quad (7)$$

and α is the curvature ratio

$$\alpha=a/R. \quad (8)$$

Subscript e refers to the outer edge of the boundary layer and W_i the constant entry axial velocity. The appropriate boundary-layer equations can then be expressed as follows:

$$\frac{\partial P}{\partial\eta}=0. \quad (9)$$

$$-U\frac{\partial V}{\partial\eta}+V\frac{\partial V}{\partial\psi}+\frac{W}{J}\frac{\partial V}{\partial s}=V_e\frac{\partial V_e}{\partial\psi}+\frac{1}{J}\frac{\partial V_e}{\partial s}$$
$$+\frac{\alpha}{J}(1-W^2)\sin\psi+\frac{1}{W_e}\left\{\frac{\partial W_e}{\partial\psi}(V_e{}^2-V^2)\right.$$
$$\left.+\frac{1}{J}\frac{\partial W_e}{\partial s}(V_e-WV)\right\}+\frac{\sqrt{\alpha}}{W_e}\frac{\partial^2 V}{\partial\eta^2}, \quad (10)$$

(Cont.)

$$-U\frac{\partial W}{\partial\eta}+V\frac{\partial W}{\partial\psi}+\frac{W}{J}\frac{\partial W}{\partial s}=\frac{\alpha\sin\psi}{J}(VW-V_e)$$
$$+\frac{1}{W_e}\left\{\frac{\partial W_e}{\partial\psi}(V_e-VW)+\frac{1}{J}\frac{\partial W_e}{\partial s}(1-W^2)\right\}+\frac{\sqrt{\alpha}}{W_e}\frac{\partial^2 W}{\partial\eta^2} \quad (11)$$

and the continuity equation

$$-\frac{\partial U}{\partial\eta}+\frac{\partial V}{\partial\psi}+\frac{1}{J}\frac{\partial W}{\partial s}-\frac{\alpha V\sin\psi}{J}$$
$$+\frac{1}{W_e}\left\{V\frac{\partial W_e}{\partial\psi}+\frac{W}{J}\frac{\partial W_e}{\partial s}\right\}=0 \quad (12)$$

where J is defined as

$$J\equiv 1+\alpha\cos\psi. \quad (13)$$

In deriving equations (10)–(12), the radial coordinate r has been approximated by the radius of the pipe, a, since the boundary layer is assumed thin compared with the pipe radius. The boundary conditions are

$$U=V=W=0\quad\text{at}\quad\eta=0,$$
$$U=U_e,\quad V=V_e,\quad W=1,\quad \frac{\partial U}{\partial\eta}=\frac{\partial V}{\partial\eta}=\frac{\partial W}{\partial\eta}=0,\quad \eta\to\infty \quad (14)$$

and U, V, W given at $s=s_i$, an initial station.

Basic Integral Relations

To solve equations (9)–(12) subject to (14), we employ the Method of Integral Relations (M.I.R.), which has been applied by numerous investigators to a wide variety of current problems in fluid dynamics. The interested reader should consult the book by Holt [10] for a full description of M.I.R. and its related applications.

We shall now derive the basic integral relations from equations (10)–(12). Introducing a complete set of linearly independent functions $\{g_k(W)\}$ the elements of which satisfy the condition

$$\lim_{W\to 1} g_k(W)=0,\quad \text{for all } k \quad (15)$$

one obtains the first integral by multiplying (12) and (11) by $g_k(W)$ and $g_k'(W)$, respectively, adding and integrating the result from $\eta=0$ to $\eta\to\infty$ and finally changing the variable of integration to W by defining Z to be

$$Z=\left(\frac{\partial W}{\partial\eta}\right)^{-1}. \quad (16)$$

The result is

$$\frac{\partial}{\partial\psi}\int_0^1 VZg_k dW+\frac{1}{1+\alpha\cos\psi}\frac{\partial}{\partial s}\int_0^1 WZg_k dW$$
$$=\left(\frac{\alpha\sin\psi}{1+\alpha\cos\psi}-\frac{1}{W_e}\frac{\partial W_e}{\partial\psi}\right)\int_0^1[(VW-V_e)g_k'+Vg_k]ZdW$$
$$+\frac{1}{1+\alpha\cos\psi}\frac{1}{W_e}\frac{\partial W_e}{\partial s}\int_0^1[(1-W^2)g_k'-Wg_k]ZdW$$
$$-\frac{\sqrt{\alpha}}{W_e}\frac{g_k'(0)}{Z_0}-\frac{\sqrt{\alpha}}{W_e}\int_0^1\frac{g_k''}{Z}dW, \quad (17)$$

where

$$g_k'(W)\equiv\frac{dg_k(W)}{dW}$$

and Z_0 is the value of Z at the wall.

The second integral relation is derived by first transforming (η,ψ,s) to (W,ψ,s) in equation (10), then multiplying the transformed equation by a weighting function $h_k(W)$, an element of another complete linear set $\{h_k(W)\}$. The result is integrated from $W=0$ to $W=1$, and we have

$$\frac{\partial}{\partial\psi}\int_0^1 \frac{V^2}{2} h_k dW + \frac{1}{1+\alpha\cos\psi}\frac{\partial}{\partial s}\int_0^1 WVh_k dW = \frac{1}{W_e}\frac{\partial W_e}{\partial\psi}$$
$$\times \int_0^1 \left\{(V_e{}^2 - V^2) - \frac{\partial V}{\partial W}(V_e - VW)\right\} h_k(W)dW$$
$$+ \frac{1}{1+\alpha\cos\psi}\frac{1}{W_e}\frac{\partial W_e}{\partial s}\int_0^1 \left\{V_e - WV - \frac{\partial V}{\partial W}(1-W^2)\right\} h_k dW$$
$$+ \frac{\alpha\sin\psi}{1+\alpha\cos\psi}\int_0^1 \left\{(1-W^2) - \frac{\partial V}{\partial W}(VW - V_e)\right\} h_k(W)dW$$
$$+ \left(V_e\frac{\partial V_e}{\partial\psi} + \frac{1}{1+\alpha\cos\psi}\frac{\partial V_e}{\partial s}\right)\int_0^1 h_k(W)dW$$
$$+ \frac{\sqrt{\alpha}}{W_e}\int_0^1 \frac{\partial^2 V}{\partial W^2}\frac{1}{Z^2} h_k(W)dW. \quad (18)$$

On the plane of symmetry, $\psi = 0$, we obtain two auxiliary integral relations from (17) and from (18) after differentiation with respect to ψ:

$$\frac{1}{1+\alpha\cos\psi}\frac{d}{ds}\int_0^1 WZdW = -\int_0^1 ZSg_k dW$$
$$+ \frac{1}{1+\alpha\cos\psi}\frac{1}{W_e}\frac{dW_e}{ds}\int_0^1 [(1-W^2)g_k{}' - Wg_k]ZdW$$
$$- \frac{\sqrt{\alpha}g_k{}'(0)}{W_e Z_0} - \frac{\sqrt{\alpha}}{W_e}\int_0^1 \frac{g_k{}''}{Z}dW, \quad (19)$$

$$\frac{1}{1+\alpha\cos\psi}\frac{d}{ds}\int_0^1 WSh_k dW = -\int_0^1 S^2 h_k dW$$
$$+ \frac{1}{1+\alpha\cos\psi}\frac{1}{W_e}\frac{dW_e}{ds}\int_0^1 \left[(S_e - WS) - \frac{\partial S}{\partial W}(1-W^2)\right] h_k dW$$
$$+ \frac{\alpha\cos\psi}{1+\alpha\cos\psi}\int_0^1 (1-W^2)h_k dW$$
$$+ \left(S_e{}^2 + \frac{1}{1+\alpha\cos\psi}\frac{dS_e}{ds}\right)\int_0^1 h_k dW$$
$$+ \frac{\sqrt{\alpha}}{W_e}\int_0^1 \frac{\partial^2 S}{\partial W^2}\frac{1}{Z^2} h_k dW, \quad (20)$$

respectively. The new variable S is given by

$$S = \frac{\partial V}{\partial\psi}. \quad (21)$$

The boundary condition on S can be readily deduced as

$$S = 0 \quad \text{at} \quad W = 0 \quad (22)$$

and

$$S = S_e \quad \text{at} \quad W = 1.$$

The unknowns Z, V, and S in equations (17)–(20) are represented as

$$Z = \left\{b_{01} + \sum_{j=1}^{N-1} b_{j1}g_j(W)\right\} \Big/ (1-W), \quad (23)$$

$$V = \left\{V_e + \sum_{j=1}^{N} b_{j2}h_j(W)\right\} W, \quad (24)$$

$$S = \left[S_e + \sum_{j=1}^{N} e_j h_j(W)\right] W. \quad (25)$$

Furthermore, for reasons of efficiency [11], $\{g_k(W)\}$ and $\{h_k(W)\}$ are chosen as orthonormal sets derived from the functions $(1-W)^k$, $k = 1, 2, \ldots$ as follows:

$$g_k(W) = \sum_{j=1}^{k} a_{kj}(1-W)^j, \quad (26)$$

$$h_k(W) = \sum_{j=1}^{k} c_{kj}(1-W)^j \quad (27)$$

such that

$$\int_0^1 g_i(W)g_j(W)\frac{W}{1-W}dW = \delta_{ij} \quad (28)$$

and

$$\int_0^1 W^2 h_i(W)h_j(W)dW = \delta_{ij}, \quad (28)$$

where δ_{ij} is the Kronecker delta. The existence and uniqueness of g_k and h_k are the content of the Gram-Schmidt process [12].

Substitution of (23)–(25) into the basic integral relations yield a system of first-order partial differential equations of hyperbolic type. In the Nth approximation, there are $2N$ such equations in $2N$ unknowns generated from equations (17)–(18) or (19)–(20). Before discussing the numerical procedure of solving the basic integral relations, let us first discuss the flow in the core region.

Flow in the Core Region

We first define the following outer variables for use in the core region:

$$U_0 = \frac{u}{W_i}, \quad V_0 = \frac{v}{W_i}, \quad W_0 = \frac{w}{W_i}, \quad P_0 = \frac{p}{\rho W_i{}^2},$$
$$r_0 = \frac{r}{a}, \quad \psi_0 = \psi, \quad s_0 = \phi. \quad (30)$$

Since we have assumed an irrotational core, we can define a potential function Ω such that

$$U_0 = \frac{1}{\alpha}\frac{\partial\Omega}{\partial r_0}, \quad V_0 = \frac{1}{\alpha r_0}\frac{\partial\Omega}{\partial\psi_0}, \quad W_0 = \frac{1}{1+\alpha r_0\cos\psi}\frac{\partial\Omega}{\partial s_0} \quad (31)$$

and Ω satisfies the Laplace equations

$$\frac{\partial^2\Omega}{\partial r_0{}^2} + \frac{\partial\Omega}{\partial r_0}\left[\frac{1}{r_0} + \frac{\alpha\cos\psi_0}{1+\alpha r_0\cos\psi_0}\right] + \frac{1}{r_0{}^2}\frac{\partial^2\Omega}{\partial\psi_0{}^2}$$
$$- \frac{\alpha\sin\psi_0}{1+\alpha r_0\cos\psi_0}\frac{\partial\Omega}{\partial\psi_0} + \frac{\alpha^2}{(1+\alpha r_0\cos\psi_0)^2}\frac{\partial^2\Omega}{\partial s_0{}^2} = 0. \quad (32)$$

The boundary conditions used in the present study are the following:

(*i*) Uniform entry:

$$\partial\Omega/\partial s_0 = 1 + \alpha r_0\cos\psi_0 \text{ at } s_0 = 0.$$

(*ii*) Flow in or out of the boundary-layer region:

$$\frac{\partial\Omega}{\partial r_0} = f(\psi_0, s_0) \quad \text{at} \quad r_0 = 1. \quad (33)$$

(*iii*) Exit condition:

$$\Omega = 0 \quad \text{at} \quad s_0 = \pi/2.$$

Some comments are in order. The second boundary condition represents the coupling between the core and the boundary layer region, i.e., the function $f(\psi_0, s_0)$ is supplied from the solution of the boundary-layer equations. The crucial assumption is that the boundary layer must be thin enough so that this condition can be applied at the pipe wall $r_0 = 1$. Since the boundary-layer thickness is of order $O(D^{-1/2})$, this assumption is reasonable for large Dean number. The third condition is imposed mainly to close the boundary-value problem. Some authors impose the fully developed condition at the exit. However, it is very unlikely that this is valid for flow into a 90° elbow at high Reynolds number. By means of a simple perturbation analysis on (32) [13], it can be shown that the flow adjusts itself to whatever exit conditions applied in a region of order α.[2] Since α, the curvature ratio, for most practical situations is small, the use of (33*iii*) does not invalidate our results for the flow development over most parts of the

[2] Since the perturbation analysis given in [13] is for the potential equation, this conclusion is valid only if the irrotationality assumption is valid itself. In actual situations, the exit conditions may influence a larger region. Experimental investigation should be the ultimate means of justifying this conclusion.

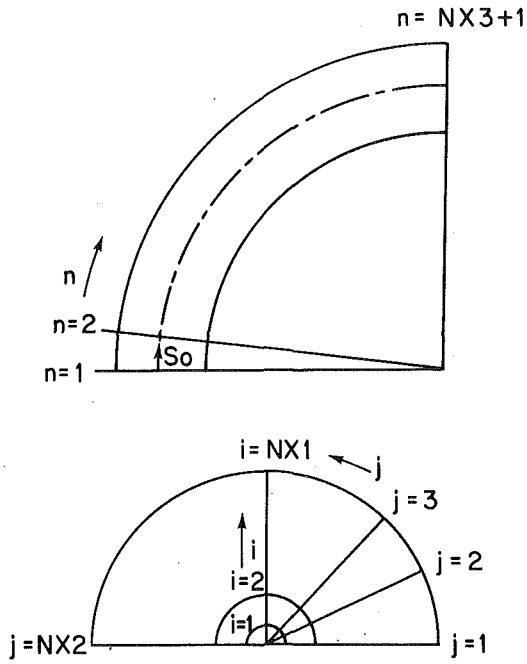


Fig. 2 Grid size representation for the core region

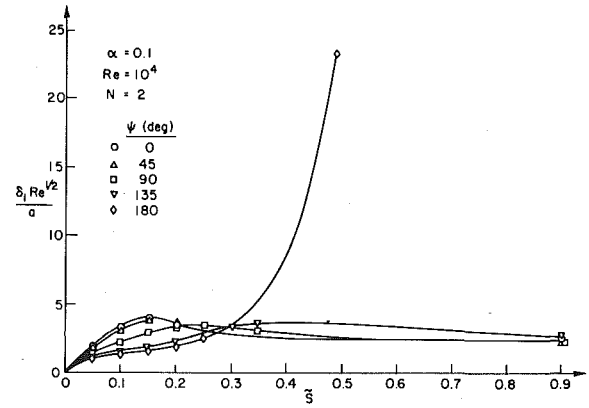


Fig. 3 Displacement thickness along the pipe

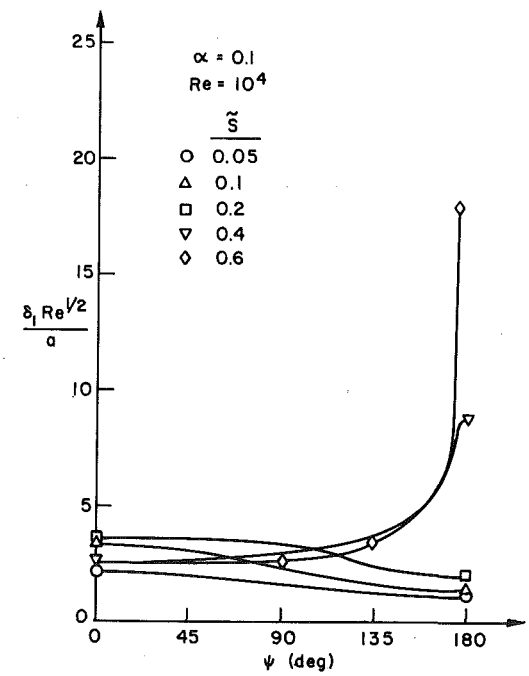


Fig. 4 Variation of displacement thickness with azimuthal angle ψ

curved pipe. Should an experimental exit profile be available, it might be of use in place of (33*iii*).

For the solution of equation (32) subject to (33), we have chosen Telenin's method [14], modified to overcome certain difficulties involved in the present problem. Fig. 2 shows the grid system for the core region. The indices i, k, n identify the r_0, ψ_0, s_0 coordinates, respectively. Thus, for example, $\mathcal{F}_{ik}{}^n$ represents the value of $\mathcal{F}$ at $r_0 = r_{0,i}$, $\psi = \psi_{0,k}$ and $s_0 = s_{0,n}$. Also, owing to the artificial singularity at $r_0 = 0$, we exclude the center of the semicircular region in our calculation. Following the method of Telenin, on planes of constant ψ_0, Ω is approximated by a Lagrange polynomial, determined from values at $NX1$ points and the boundary condition at $r_0 = 1$. On planes of constant r_0, Ω is approximated by a finite cosine series, determined from values at $NX2$ points. These approximations are introduced into (32) to yield a system of $NX1 \cdot NX2$ second-order ordinary differential equations in Ω_{ik}, $i = 1, \ldots, NX1$ and $k = 1, \ldots, NX2$, together with the entry and exit conditions. An implicit, second-order finite-difference scheme is employed to solve the resulting two point boundary-value problem. The final system consists of $NX1 \cdot NX2 \cdot (NX3 + 1)$ algebraic equations for $\Omega_{ik}{}^n$ at each node, the coefficient matrix of which is of block tridiagonal type. We shall now discuss the numerical procedure for the whole region of interest.

Numerical Procedure

The outer flow is initially represented by a two-dimensional point vortex profile to provide the values of V_e, W_e and their derivatives in the basic integral relations. The cross derivatives, i.e., $\partial Z/\partial\psi$ and $\partial V/\partial\psi$ in (17)–(20) are approximated by a backward first-order difference scheme and the resulting ordinary differential equations are integrated by means of a variable order Adams-Moulton method. The necessary initial conditions are approximated by the Blasius solution for flow over a flat plate. The outer edge normal velocity is then evaluated from this boundary-layer solution to provide a current estimate of f in (33) and Ω is recalculated. This gives new values of V_e, W_e and their derivatives for the next integration of the integral relations. The process is repeated until there is no appreciable change in the results between successive iterations. For a more detailed description of the numerical procedure, see [13].

Results and Discussion

Numerical results have been obtained for the whole flow field inside a 90° elbow for $\alpha = 0.1$ and $\mathrm{Re} = 10^4$. Convergence is achieved after 10 iterations for $N = 1$ and about 40 iterations for $N = 2$.

Fig. 3 shows the displacement thickness δ_1 defined as

$$\frac{\delta_1}{a} = \frac{1}{D^{1/2}} \int_0^1 (1 - W) Z dW. \tag{34}$$

along the curved pipe at different azimuthal locations. The streamwise variable is modified as $\tilde{s} = \phi/\pi/2$. It is relatively thin compared with the radius of the pipe, except at the inner bend, $\psi = 180°$, where it increases continuously as the flow moves downstream, as one would expect physically. The variation of δ_1 azimuthally is shown in Fig. 4. In the early stages of the flow, the boundary layer is thinner at the inner bend than it is at the outer bend. This is because initially the outer flow is faster at the inner bend than at the outer bend and the curvature effect is small. As the flow develops, curvature effect be-

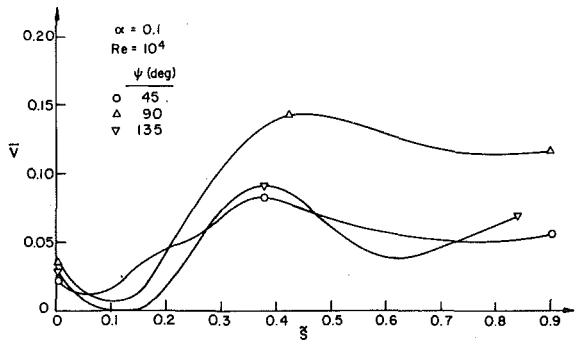


Fig. 5 Average nondimensional velocity in the boundary layer along the pipe

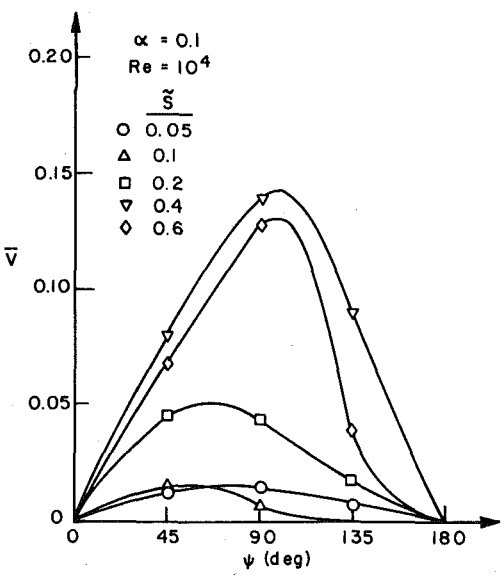
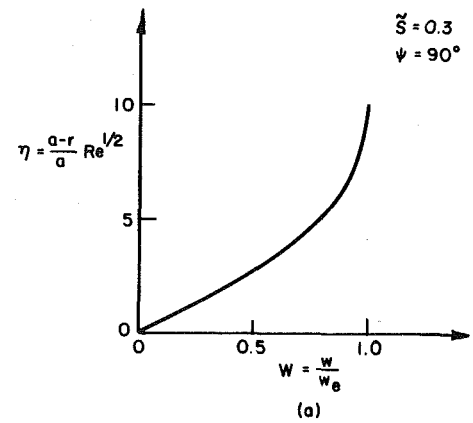


Fig. 6 Variation of $\overline{V}$ with ψ at different streamwise stations



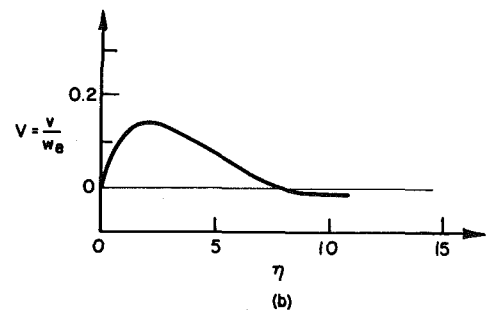


Fig. 7 Velocity profiles across the boundary layer

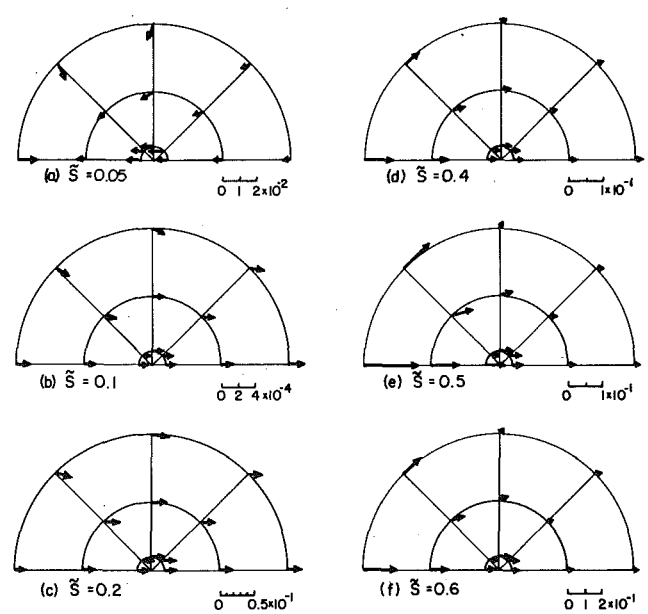


Fig. 8 Crossflow in the core region: (*a*) $\tilde{s} = 0.05$, (*b*) $\tilde{s} = 0.1$, (*c*) $\tilde{s} = 0.2$, (*d*) $\tilde{s} = 0.4$, (*e*) $\tilde{s} = 0.5$, (*f*) $\tilde{s} = 0.6$

comes dominant and the boundary layer is thicker at the inner bend than at the outer bend. Further downstream, the boundary layer stays quite uniform around the pipe and increases abruptly as ψ approaches 180°. This indicates secondary flow separation somewhere ahead of $\psi = 180°$.

Next, the variations of the average azimuthal velocity, $\overline{V}$, defined as

$$\overline{V} = \int_0^1 V dW \tag{35}$$

along and around the pipe are shown in Figs. 5 and 6, respectively. The absolute magnitude of the azimuthal velocity inside the boundary layer is small everywhere, as indicated in the figures. Finally, typical profiles of the axial velocity and azimuthal velocity across the boundary layer are shown in Fig. 7.

We now turn to the flow in the core region. The development of the secondary flow at different streamwise sections is shown in Fig. 8. The crossflow velocity vector at each node is drawn. The direction of the flow is represented by the direction of the arrow and the magnitude by the arrow's length base on the scale shown on the bottom right of each figure. The predicted secondary motion is directed from the outer bend ($\psi < 90°$) toward the inner bend ($\psi > 90°$) very near the entry, as shown in Fig. 8(*a*). Physically, since the flow develops from a uniform profile to the potential vortex profile immediately downstream of the entry section, the fluid must move from the outer bend to the inner bend for reasons of mass conservation requirement. As the flow develops, the predicted secondary motion begins to reverse its direction and, in Fig. 8(*c*), the fluid has already moved from the inner bend toward the outer bend. This is due to the subsequent bound-

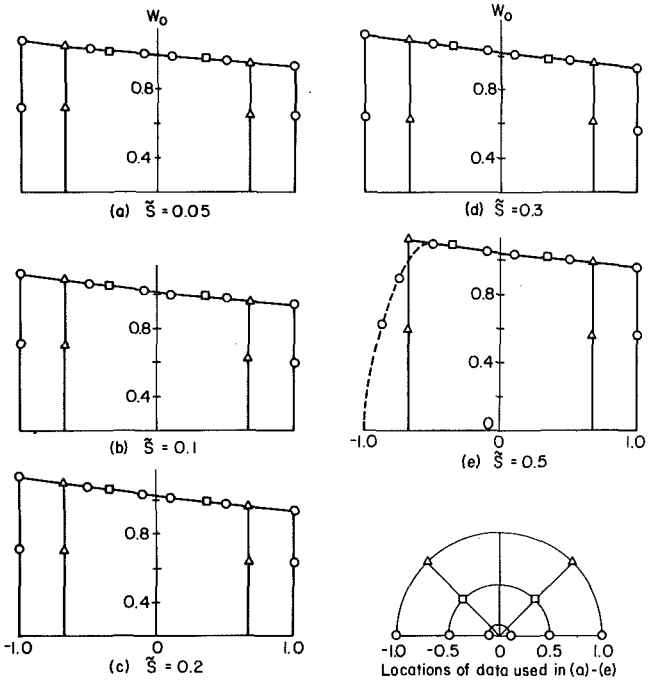


Fig. 9 Axial velocity profile in the core region: (*a*) $\tilde{s} = 0.05$, (*b*) $\tilde{s} = 0.1$, (*c*) $\tilde{s} = 0.2$, (*d*) $\tilde{s} = 0.3$, (*e*) $\tilde{s} = 0.5$, (*f*) legend

ary-layer growth around the pipe. It is interesting to note that the direction of the crossflow is essentially parallel to the plane of symmetry at all the stations shown in Fig. 8, an assumption used by many investigators of curved pipe flow.

The axial velocity profiles are shown in Fig. 9. The uniform entry profile develops into a potential vortex type profile immediately downstream as shown in Fig. 9(*a*). Since the boundary layer in the present study is very thin, except near $\psi = 180°$ when the flow develops further downstream, the axial velocity profile does not change much over the entire elbow. Thus the characteristic of the potential vortex profile is seen at every station shown in Fig. 9. Furthermore, the velocity results are almost independent of the vertical distance measured from the symmetry plane. In Fig. 9(*c*) the boundary-layer profile is joined to the outer profile at the inner symmetry plane because the boundary layer there is relatively thick. The present model does not predict the shift of the maximum axial velocity toward the outer bend which occurs eventually when the viscous region in the inner bend extends further into the core region. This is not surprising because we have assumed that the viscous region is confined near the pipe wall everywhere. Moreover, the assumption of irrotationality will be invalid.

Conclusion

A numerical investigation has been carried out concerning the development of steady, laminar, incompressible flow in the entry region of a curved pipe at high Reynolds number for uniform entry conditions. The flow field is divided into an irrotational core region and a viscous boundary-layer region. An efficient scheme based on Telenin's method was used to solve the Laplace equation in three dimensions and the orthonormal method of integral relations was applied to obtain numerical solution to the present three-dimensional boundary-layer equations. The irrotational assumption causes the uniform entry profile to develop into a potential vortex profile very near the entry. Due to the effect of the boundary layer, it has been shown that there is a novel reversal of the direction of the secondary motion in the core region, a result which has not been reported before. For the boundary-layer region, the present results are qualitatively the same as those obtained by Yao and Berger [7], but any direct quantitative comparison is not applicable because of the different range of Reynolds number considered and different assumptions concerning the core region. It must be emphasized that it is not appropriate to apply the present model to describe the fully developed region since viscous effect is no longer confined to the pipe wall. Finally, any results presented here must be eventually checked against reliable experimental data. Such data are now being assembled in experiments performed at the Lawrence Berkeley Laboratory.

Acknowledgment

The author is grateful to Dr. Alan Levy and Prof. M. Holt for their continuous encouragement during the course of this work. Special thanks are due to Prof. S. A. Berger for his valuable comments and to Dr. J. J. Chattot, who is now in France, for his useful discussion on the numerical schemes. This work is funded by the Basic Energy Science Division of the Department of Energy.

References

1 Dean, W. R., "Note on the Motion of Fluid in a Curved Pipe," *Philosophical Magazine*, Vol. 4, 1927, pp. 208–223.

2 Dean, W. R., "The Streamline Motion of Fluid in a Curved Pipe," *Philosophical Magazine*, Vol. 5, 1928, pp. 673–695.

3 Ito, H., "Laminar Flow in Curved Pipes," *ZAMM*, Vol. 49, 1969, pp. 653–663.

4 Collins, W. M., and Dennis, S. C. R., "The Steady Motion of a Viscous Fluid in a Curved Tube," *Quarterly Journal of Mechanics and Applied Mathematics*, Vol. 28, 1975, pp. 133–156.

5 Smith, F. T., "Steady Motion Within a Curved Pipe," *Proceedings of the Royal Society*, London, Series A, Vol. 347, 1976, pp. 345–370.

6 Hawthrone, W. R., "Secondary Circulation in Fluid Flow," *Proceedings of the Royal Society*, London, Series A, Vol. 206, 1951, pp. 374–387.

7 Yao, L. S., and Berger, S. A., "Entry Flow in a Curved Pipe," *Journal of Fluid Mechanics*, Vol. 67, 1975, pp. 177–196.

8 Singh, M. P., "Entry Flow in a Curved Pipe," *Journal of Fluid Mechanics*, Vol. 65, 1974, pp. 517–539.

9 Agrawal, Y., Talbot, L., and Gong, K., "Laser Anemometry Study of Entry Curved Pipe Flow," *Journal of Fluid Mechanics*, Vol. 85, 1978, pp. 497–518.

10 Holt, M., *Numerical Methods in Fluid Dynamics*, Springer-Verlag, Berlin, Heidelberg, New York, 1977.

11 Fletcher, C. A. J., and Holt, M., "Modified Version of the Method of Integral Relations," *Journal of Computational Physics*, Vol. 18, 1975, pp. 154–164.

12 Issacson, E., and Keller, H. B., *Analysis of Numerical Methods*, Wiley, 1966, p. 199.

13 Yeung, W. S., "Gas-Particle Flow in the Entry Region of a Curved Pipe," PhD Dissertation, Lawrence Berkeley Laboratory, Report LBL-9905, University of California, Berkeley, Apr. 1979.

14 Gilinskii, S. M., Telenin, G. F., and Tinyakov, G. P., *Izv. Akad. Nauk. SSSR Mekh. Mash.*, Vol. 4, 1964, (translated as NASA TTF 297).

C. A. Felippa

Staff Scientist,
Applied Mechanics Laboratory,
Lockheed Palo Alto Research Laboratory,
Palo Alto, Calif. 94304

A Family of Early-Time Approximations for Fluid-Structure Interaction

A hierarchical family of early-time, high-frequency asymptotic, surface interaction approximations is derived for a structure submerged in an infinite acoustic fluid. Kirchhoff's retarded-potential integro-differential formulation is used as exact source formula. The well-known plane-wave and curved-wave approximations result as the first two members of the hierarchy. Acoustic impedance characteristics of the first four members are exhibited for several sample geometries.

Introduction

A structure is submerged in an infinite acoustic fluid satisfying the three-dimensional scalar wave equation. The structure is subjected to impulsive loading; e.g., a compressive shock wave propagates through the fluid and impinges on the structure. If the fluid is computationally viewed as a "magic membrane" that envelopes the structure, the ensuing interaction between fluid and structure can be exactly expressed through Kirchhoff's retarded potential formula (RPF), which is the boundary-integral formulation of Huygens' principle [1].

For structures of arbitrary geometry, the direct application of the RPF to transient response calculations is computationally cumbersome. A key drawback is the need for preserving a vast amount of historic information as the coupled equations of motion are numerically integrated; and this amount grows linearly with the number of time steps taken. In addition, tracking of wavefront discontinuities at each time step demands nontrivial logic [2]. Because of these difficulties, a variety of simpler interaction approximations of current-time form (i.e., memoryless) have been proposed. Geers [3] categorizes these approximations into three classes:

1 Early-time approximations (ETA), which are valid for short acoustic wavelengths (high frequencies) and are useful for modeling early stages of the transient interaction;

2 Late-time approximations (LTA), which are valid for long acoustic wavelengths (low frequencies) and are useful for modeling late stages of the response; and

3 Doubly asymptotic approximations (DAA), which are valid for both short and long acoustic wavelengths [4, 5] and thus are applicable to complete response calculations.

Doubly asymptotic approximations have been particularly successful in large-scale calculations involving complex structures [6, 7]. They can be constructed by pairing early and late-time approximations through the method of matched asymptotic expansions in the frequency domain [8].

From the preceding considerations it follows that the derivation of ETA families is not only useful in its own right, but also provides one of the two required ingredients for building DAA families. In the present paper an ETA family is precipitated from the RPF through a simple local source analysis.

Retarded Potential Formulation

Consider an infinite acoustic fluid governed by the isotropic, scalar wave equation

$$c^2\Delta\phi = \ddot{\phi}, \tag{1}$$

where c is the acoustic wavespeed, superposed dots denote temporal differentiation, and the *velocity potential* ϕ is related to the *acoustic excess pressure* p and *fluid-particle velocity* $\mathbf{v}$ as

$$p = \rho\dot{\phi}, \qquad \mathbf{v} = \nabla\phi. \tag{2}$$

We assume that the potential ϕ has no singularities outside a closed interaction surface S (the "wet surface" of the submerged structure) for all values of the time t from $-\infty$ up to the instance under consideration, and decays as R^{-1} at a large distance R from S. Then Kirchhoff's retarded potential formula (RPF) can be expressed as [1, 9]

$$4\pi\epsilon\phi = -\int_S \left\{R^{-1}\left[\frac{\partial\phi}{\partial n}\right] + \beta R^{-2}([\phi] + \frac{R}{c}[\dot{\phi}])\right\} dS + 4\pi\epsilon\phi_w, \tag{3}$$

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, March 1980; final revision, June, 1980.

where $\phi = \phi(\mathbf{X}, t)$ is the velocity potential at a *field point* F of coordinates $\mathbf{X} = (X, Y, Z)$, and

$R = ||\mathbf{R}||_2 = |\mathbf{X} - \mathbf{X}_S|$, the distance from the field point to a source point P of coordinates $\mathbf{X}_S$ on S

$\mathbf{n}$ = outward normal to S at the source point (positive going into the fluid)

$\beta = \partial R/\partial n$, cosine of the angle formed by vectors $\mathbf{R}$ and $\mathbf{n}$

$[f]$ = the *retarded value*, $f(\mathbf{X}, t - R/c)$, of a field function $f(\mathbf{X}, t)$

ϕ_w = incident wave potential

$\epsilon = 1, \frac{1}{2}, 0$ depending on whether the field point is within the fluid, on the surface S (assumed to be smooth), or inside S, respectively. (At a sharp corner of S with enclosed solid angle α, $\epsilon = \alpha/4\pi$)

For deriving surface interaction approximations, the field point F is considered to be on $S(\epsilon = \frac{1}{2})$ and the incident-wave potential ϕ_w is dropped; i.e., only the scattered component of the pressure field will be retained. The resulting formula is

$$2\pi\phi = \int_S \left\{ R^{-1}[u] - \beta R^{-2}([\phi] + \frac{R}{c}[\dot{\phi}]) \right\} dS \tag{4}$$

in which

$$u = -\left.\frac{\partial \phi}{\partial n}\right|_S \tag{5}$$

denotes the *normal fluid-particle velocity* at S. Equation (4) is a homogeneous "retarded" integro-differential equation in which all variables pertain to S. This is complemented by the initial rest conditions

$$\phi(t) = u(t) = 0 \qquad t \le 0 \tag{6}$$

for every field point on S. Note that $\dot{\phi}(0+)$ may be nonzero, however; that is, occurrence of a step-pressure at $t = 0$ (or at other positive values of t) is not precluded.

Local Geometry Analysis

The Contributing Surface. Early-time approximations (ETA) are interaction expressions asymptotically valid for

$$ct \ll a \tag{7}$$

where a is a characteristic dimension of S. These approximations are *local* in nature, i.e., depend only on geometric information collected in the neighborhood of an individual field point, F. The "contributing cap" $S_F(t)$ for $t > 0$ is constituted by surface source points P located at a distance $R \le ct$ from F, as illustrated in Fig. 1.

A local orthogonal Cartesian system (x, y, z) is defined with (x, y) in the tangent plane and z along the external normal at F, respectively. (The approximations will turn out to be invariant with respect to the orientation of x and y, as could be expected.) A cylindrical coordinate system (r, θ, z) with $r^2 = x^2 + y^2$ and θ = angle (r, x) is also used (cf. Fig. 1). The distance R between P and F is then expressable as $R^2 = r^2 + z^2$. Now the Taylor expansion of z about a regular point F can be written as

$$z = Ar^2 + Br^3 + Cr^4 + Dr^5 + \ldots = AR^2 + BR^3 + (C - A^3)R^4 + \left(D - \frac{7}{2}A^2B\right)R^5 + \ldots, \tag{8}$$

where $A, B, \ldots$ are functions of the angle θ only:

$$\begin{aligned} A &= (z_{xx}\cos^2\theta + 2z_{xy}\cos\theta\sin\theta + z_{yy}\sin^2\theta)/2!, \\ B &= (z_{xxx}\cos^3\theta + 3z_{xxy}\cos^2\theta\sin\theta + \ldots + z_{yyy}\sin^3\theta)/3!, \\ &\ldots \end{aligned} \tag{9}$$

in which $z_{xx} = \partial^2 z/\partial x^2, \ldots$ etc., are assumed to exist at F.

The expansion (8) permits us to express the "cap" S_F in the parametric series form

$$x = r(R, \theta)\cos\theta, \quad y = r(R, \theta)\sin\theta, \quad z = z(R, \theta), \tag{10}$$

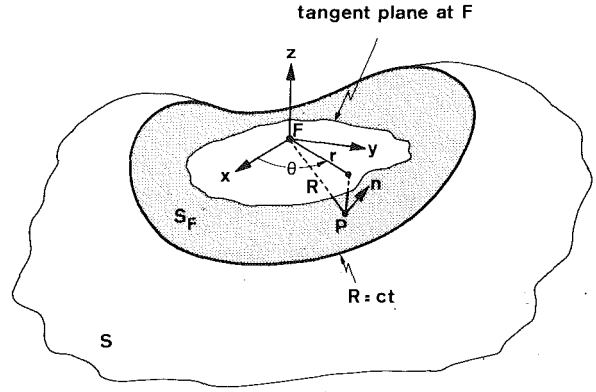


Fig. 1 Coordinate systems for local source analysis

in which $r(R, \theta) = +\{R^2 - z^2(R, \theta)\}^{1/2}$. Hence R and θ will play the role of surface coordinates.

Surface Metric Coefficients. The two geometric quantities needed for (4) are the element of area dS and the combination $\beta R^{-2}dS$. Now

$$dS = W_1 dR d\theta = \{r^2(r_R^2 + z_R^2) + (r_R z_\theta - z_R r_\theta)^2\}^{1/2} dR d\theta, \tag{11}$$

$$\beta R^{-2} dS = W_2 dR d\theta = -(r/R)^3 (z/r)_R dR d\theta, \tag{12}$$

where subscripts denote partial derivatives with respect to the surface coordinates R, θ. Inserting (8) into (11) and (12) yields

$$W_1/R = 1 + \frac{1}{2}A_\theta^2 R^2 + (AB + A_\theta B_\theta)R^3 + \frac{1}{8}(16AC - 16A^4 + 12B^2 + 4B_\theta^2 - 20A^2A_\theta^2 - A_\theta^4)R^4 + \ldots, \tag{13}$$

$$-W_2 = A + 2BR + 3(C - A^3)R^2 + (4D - 14A^2B)R^3 + 5(E + 2A^5 - 4AB^2 - 4A^2C)R^4 + \ldots. \tag{14}$$

For a spherical surface, $W_1 = R$ and $W_2 = -A$ exactly, which provides some checks on the coefficients of the series (13) and (14). For a circular cylinder of radius a, $W_1/R = \{1 - \sin^2\theta\cos^2\theta\,(R/a)^2\}^{-1/2}$ and $W_2 = (1 - W_1/R)/(R^2\sin^2\theta)$, which provides further verification.

Retarded Field Integrals. In the following section we shall encounter integrals of the form:

$$\bar{g} = \frac{1}{2\pi}\int_0^{2\pi} g(\theta)d\theta, \tag{15}$$

$$\left.\begin{aligned} J_m &= \int_0^{ct} R^m[u]dR, \\ K_m &= \int_0^{ct} R^m([\phi] + \frac{R}{c}[\dot{\phi}])dR. \end{aligned}\right\} m = 0, 1, \ldots \tag{16}$$

The change of variable $t' = t - R/c$, $R = c(t - t')$ transforms the "retarded" integrals (16) into ordinary momentum integrals. These are easily evaluated through repeated integration by parts:

$$\begin{aligned} &J_0 = cu^*, \quad J_1 = c^2u^{**}, \quad J_2 = 2c^3u^{***}, \ldots \\ &K_0 = 2c\phi^*, \quad K_1 = 2c^2\phi^{**}, \quad K_2 = 8c^3\phi^{***}, \ldots \end{aligned} \tag{17}$$

in which each superscript asterisk denotes temporal integration from $t = 0$ [at which both u and ϕ vanish on account of the initial conditions (6)] through t.

Early-Time Approximation Family

Derivation. The formal procedure for obtaining an ETA family is now straightforward: insert (11) through (14) into (4); integrate in θ (from 0 to 2π) and divide through by 2π; integrate in R (from 0 to ct); and insert the values (17) for the momentum integrals. Finally, time-differentiate once and use the first of (2) to pass from velocity-potential to pressure variables on the left-hand side. The resulting expression is

$$p + P_1 p^* + P_2 p^{**} + \ldots = \rho c(u + U_1 u^* + U_2 u^{**} + \ldots), \quad (18)$$

where

$$P_1 = -2c\overline{A},$$
$$P_2 = -6c^2\overline{B} = 0,$$
$$P_3 = -24c^3(\overline{C} - \overline{A}^3),$$
$$P_4 = -120\, c^4(\overline{D} - 3.5\, \overline{A^2 B}) = 0,$$
$$P_5 = -720\, c^5(\overline{E} + 2\overline{A}^5 - 4\overline{AB}^2 - 4\overline{A^2 C}), \ldots$$
$$U_1 = 0,$$
$$U_2 = \overline{c^2 A_\theta^2},$$
$$U_3 = c^3(\overline{AB} + \overline{A_\theta B_\theta}) = 0,$$
$$U_4 = 3c^4(16\overline{AC} - 16\,\overline{A}^4 + 8\overline{A_\theta C_\theta} - 20\,\overline{A^2 A_\theta^2} - \overline{A}_\theta^4 + 12\overline{B}^2 + 4\overline{B}_\theta^2), \ldots \quad (19)$$

The reader is reminded that a bar over a θ-dependent expression is used to denote its circumferential average, as defined by equation (15).

For a regular field point [i.e., one at which the Taylor-series expansion (8) holds], all coefficients P_{2j}, U_{2j-1} ($j = 1, 2 \ldots$) vanish. The nonvanishing ones can be expressed in terms of surface metric invariants such as

$$\kappa = (z_{xx} + z_{yy})/2 = \text{mean curvature},$$
$$\tau = z_{xx} z_{yy} - z_{xy}^2 = \text{total curvature}. \quad (20)$$

For example,

$$2\overline{A} = \kappa,$$
$$16\overline{A}^3 = \kappa(5\kappa^2 - 3\tau),$$
$$2\overline{A}_\theta^2 = \kappa^2 - \tau,$$
$$64\overline{C} = z_{xxxx} + 2z_{xxyy} + z_{yyyy}. \quad (21)$$

ETA Family Members. The integro-differential series (18) can be used to generate specific members by truncating it to a finite number of terms taken from both sides. The lowest order member, ETA_1, is the well-known *plane wave approximation* (PWA) of Mindlin and Bleich [10].

$$p = \rho c u. \quad (22)$$

The second member, ETA_2, is the *curved wave approximation* (CWA)

$$p - \kappa c p^* = \rho c u. \quad (23)$$

The CWA was first used by Haywood [11] but specialized to a cylindrical geometry, and later generalized by Bedrosian and DiMaggio [12].

The next two members, ETA_3 and ETA_4, are

$$p - \kappa c p^* + P_3 p^{***} = \rho c(u + U_2 u^{**}), \quad (24)$$

$$p - \kappa c p^* + P_3 p^{***} + P_5 p^{*****} = \rho c(u + U_2 u^{**} + U_4 u^{****}). \quad (25)$$

Any of (23)–(25) can be readily transformed to an ordinary differential equation by repeated temporal differentiation. Potential users should beware, however, that such an operation introduces *spurious initial conditions* at $t = 0$ (or, in general, at any time where the pressure is not sufficiently t-differentiable). Ensuing numerical difficulties should be taken care of appropriately in any computational implementation.

Specific Geometries. Consider a torus obtained by rotating a circle of radius a about a coplanar axis displaced $a + b$ from the circle center. If $b = -a$ and ∞, the torus reduces to a sphere and cylinder, respectively. We focus our attention on points of the *inner generated circle*. These points will be *hyperbolic, parabolic,* or *elliptic* according to whether $|a + b| > a$, $b = \infty$, or $|a + b| < a$, respectively.

Table 1 ETA coefficients for some geometries

	Sphere* (b = -a)	Cylinder* (b = ∞)	Torus† (b = a)	Torus† (b = a/2)
aP_1/c	1	1/2	0	-1/2
a^3P_3/c^3	0	3/16	3/4	9/16
a^5P_5/c^5	0	135/256	0	-3375/256
a^2U_2/c^2	0	1/8	1/2	9/8
a^4U_4/c^4	0	27/128	3/8	315/128

* All points
† At points of inner generated circle

If the local axis x is taken along the meridional direction, then the only nonzero partials contributing to the first four ETA members are: $z_{xx} = -1/a$, $z_{yy} = 1/b$, $z_{xxxx} = -3/a^3$, $z_{xxyy} = -1/ab^2$, $z_{yyyy} = 3/b^3$, $z_{xxxxxx} = -45/a^5$, $z_{xxxxyy} = -3/a^3b^2 + 2/a^3b^2$, $z_{xxyyyy} = -9/ab^4$, and $z_{yyyyyy} = 45/b^5$. The coefficients $P_1 \ldots U_4$ are listed in Table 1 for the four specific cases $b = -a$ (sphere), $b = \infty$ (cylinder), $b = a$, and $b = a/2$. Note that for the sphere all coefficients but P_1 vanish, as noted in the "Surface Metric Coefficients" section.

Computer Implementation

Suppose that the submerged structure is discretized as an assembly of finite elements. Because of the local nature of ETA formulas, their implementation in terms of "boundary fluid elements" is straightforward. One could attach, for instance, such elements to "wet face" centroids of structural finite elements [7]. In the case of low-order ETA members, interpretation in terms of mechanical components is immediate; e.g., the "PWA fluid element" is nothing more than a ρc dashpot.

If refined ETA formulas are used, high-order geometry information is required to calculate the coefficients (19) at control points. For simple geometries, such as the circular cylinder, this may be fed *a priori* as part of the input data. For arbitrary shapes, a logical procedure is to resort to local surface interpolation using coordinates of a sufficient number of adjacent surface node points.

The resulting fluid-structure semidiscrete equations of motion may be either solved as a full coupled system (in which case an explicit time integration method would normally be used), or through a partitioned integration approach such as the staggered solution procedure [13].

Spectral Characteristics

Acoustic Impedance. Insight into the properties of ETA family members can be gained through examination of their impedance as a function of driving frequency. Following Junger and Feit [14] and Geers [3], the *specific acoustic impedance* (impedance per unit of surface area) of a memoryless pressure-velocity expression such as (18) can be defined as the ratio

$$\zeta(\omega) = \frac{\hat{p}(-j\omega)}{\hat{u}(-j\omega)} = \xi(\omega) - j\omega\mu(\omega), \quad (26)$$

where ω is driving circular frequency, $\hat{p}(s)$ and $\hat{u}(s)$ are the one-sided Laplace transforms of $p(t)$ and $u(t)$, respectively, and $j^2 = -1$. The *specific acoustic resistance* $\xi = \text{Re}\,(\zeta)$ characterizes the radiation damping power; while $\mu = -\text{Im}\,(\zeta)/\omega$ is either *specific acoustic inertia* (also called *specific accession to inertia*) if positive (pressure in phase with acceleration) or effective *acoustic spring stiffness* if negative (pressure in phase with displacement).

Impedance Diagrams for Sample Geometries. Figs. 2–5 show dimensionless ratios $\xi/\rho c$ and $\mu/\rho a$ as functions of the dimensionless

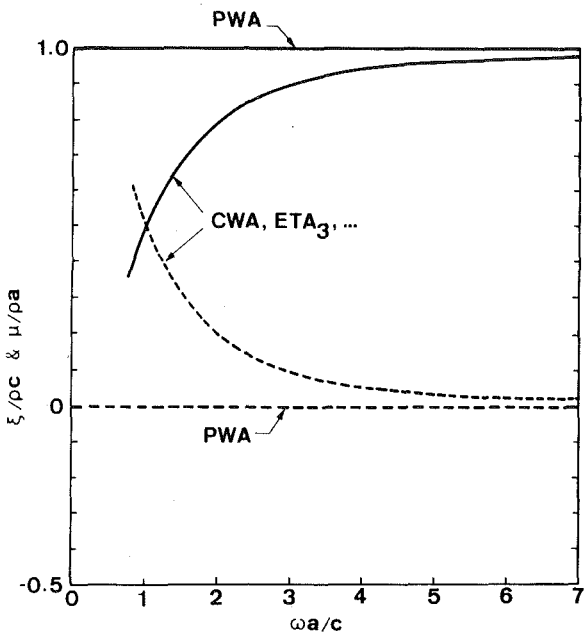


Fig. 2 Specific acoustic resistance ξ (solid lines) and inertia μ (dashed lines) exerted by an external acoustic medium on points of the surface of a sphere of radius a, for the first four members of the ETA family (18)

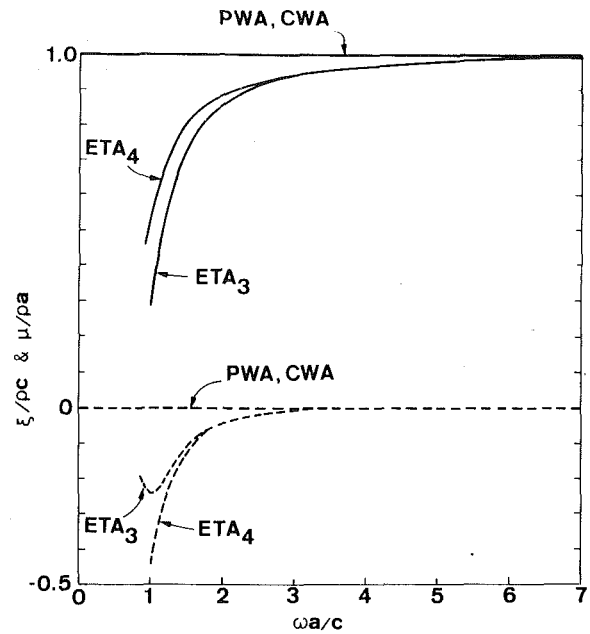


Fig. 4 Specific acoustic resistance ξ (solid lines) and effective spring stiffness μ (dashed lines) exerted by an external acoustic medium on points of the inner generator of a torus of meridional radius a and inner generation radius $b = a$, for the first four members of the ETA family (18)

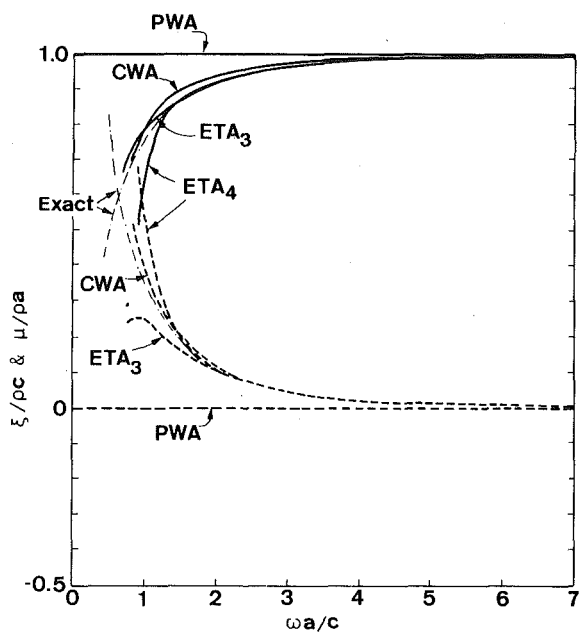


Fig. 3 Specific acoustic resistance ξ (solid lines) and inertia μ (dashed lines) exerted by an external acoustic medium on points of the surface of a cylinder of radius a, for the first four members of the ETA family (18), and for the exact breathing-mode impedance (cf. Appendix A)

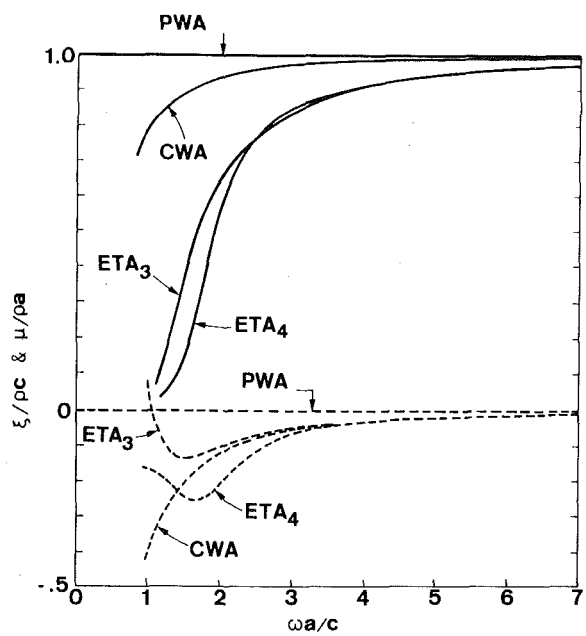


Fig. 5 Specific acoustic resistance ξ (solid lines) and effective spring stiffness μ (dashed lines) exerted by an external acoustic medium on points of inner generation of a torus of meridional radius a and inner generation radius $b = a/2$, for the first four members of the ETA family (18)

driving frequency $\omega c/a$, for the four specific geometries of Table 1 and the first four members of the ETA family. These are physical coordinate "local" impedances rather than the more usual modal impedances [3, 14]. Inasmuch as $\hat{p}$ and $\hat{u}$ are effectively assumed to be constant over the cap of Fig. 1, (26) can be interpreted in global sense only for uniformly pulsating bodies; e.g., breathing motions of spheres and cylinders. Another "globalization" restriction is related to the use of the series (8) to describe nonlocal geometry.

The sphere diagrams (Fig. 2) offer no surprises, as it is well known that the CWA expresses the exact interaction for uniformly pulsating spheres.

The cylinder diagrams (Fig. 3) are contrasted with the exact ξ and μ for breathing motions [14]. The curves clearly display the high-frequency asymptotic character of the ETA family; e.g., for $\omega a/c < 1.5$, ETA_4 is inferior to ETA_3, which in turn collapses for $\omega a/c < 1$. As shown in Appendix A, the low-frequency performance of higher-order ETAs can be improved by transforming to a Padé-quotient representation.

For the $b = a$ torus (Fig. 4) the surface points under consideration are hyperbolic and of *zero mean curvature.* As $P_1 = -\kappa c$ vanishes, the PWA and CWA coalesce, and it is necessary to proceed to ETA_3 to obtain some frequency dependence. Note that for both ETA_3 and

ETA_4, $\mu(\omega)$ is negative, i.e., the external acoustic medium acts as a distributed spring stiffness. This is caused by the near-field "entrainment" of fluid brought about by the *concavity* of sectors of the surface; the effect resembles that of a fluid partly enclosed in a vibrating container [14, Section 2.8]. No exact impedance solutions for the torus are known.

For the $b = a/2$ torus (Fig. 5) the surface points are hyperbolic with heavy dominance of concavity, i.e., *positive mean curvature.* The CWA does not coincide with the PWA as in Fig. 4, but provides a poor (and unconservative) approximation of the resistance ξ, which characterizes radiation damping.

Usage Recommendations. A tentative conclusion from this preliminary study is that for *convex* interaction surfaces (those containing only elliptic or parabolic points) there seems little point in going beyond the CWA. Application of more refined ETA formulas deserves consideration, however, if the surface contains hyperbolic points with zero or positive mean curvature. In the latter case, use of the CWA may incur significant radiation damping overestimation, which can in turn translate into unconservative predictions of early time peak responses.

Concluding Remarks

In view of the elementary mathematical tools used, the derivation of the ETA family "spawner" (18) seems surprisingly straightforward. The two underlaying physical assumptions are: the external fluid is an acoustic medium, and the early-time interaction is governed by local effects. A similar derivation procedure appears feasible to construct ETA families for coupled-field problems where a retarded boundary-integral formulation is available; for example, structure-rock interaction.

As noted in the Introduction, the application range over which stand-alone ETA formulas are useful is restricted to high-frequency dominated interactions such as impulsive acoustic transients. But this range can be significantly enlarged by utilizing such formulas as *outer expansions* in doubly asymptotic approximations valid over the entire frequency spectrum. The range-expanding reformulation described in Appendix *A*, although promising, cannot be expected to replace doubly asymptotic forms because it does not account for added mass effects and thus remains local in nature.

The approximations presented here may be contrasted with the short-wavelength approximations of optics and geometric acoustics. The main objective of the latter is to provide a simplified description of the scattered near-field. Here the structural response of the scatterer is of main concern. Hence, what happens at the interaction surface is of primary importance, and the resulting approximations are expressed only in terms of surface values.

Acknowledgment

The author expresses his appreciation to Dr. T. L. Geers for valuable criticism of the first draft of the paper. This work was performed under the Coupled Systems Technology Project of the Independent Research Program of Lockheed Missiles & Space Company, Inc.

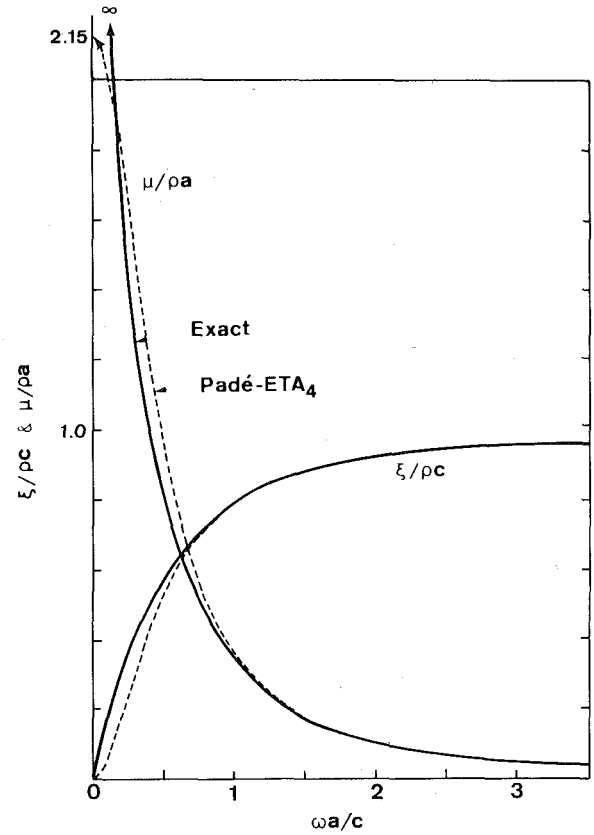


Fig. 6 Specific acoustic impedance components for circular cylinder: exact (solid lines) versus Padé approximant (28) (dashed lines)

References

1 Baker, B. B., and Copson, E. T., *The Mathematical Theory of Huygens' Principle,* Clarendon Press, Oxford, 1949.

2 Huang, H., Everstine, G. C., and Wang, Y. F., "Retarded Potential Techniques for the Analysis of Submerged Structures Impinged by Weak Shock Waves," in *Computational Methods for Fluid-Structure Interaction Problems,* eds., Belytschko, T., and Geers, T. L., AMD, Vol. 26, ASME, New York, 1977, pp. 83–94.

3 Geers, T. L., "Transient Response Analysis of Submerged Structures," in *Finite Element Analysis of Transient Nonlinear Structural Behavior,* eds., Belytschko, T., et al., AMD Vol. 14, ASME, New York, 1975, pp. 59–84.

4 Geers, T. L., "Residual Potential and Approximate Methods for Three-Dimensional Fluid-Structure Interaction Problems," *Journal of the Acoustical Society of America,* Vol. 49, 1971, pp. 1505–1510.

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

6 Ranlet, D., DiMaggio, F. L., Bleich, H. H., and Baron, M. L., "Elastic Response of Submerged Shells With Internally Attached Structures to Shock Loading," *Computers and Structures,* Vol. 7, 1977, pp. 355–364.

7 DeRuntz, J. A., Geers, T. L., and Felippa, C. A., "The Underwater Shock Analysis (USA) Code: A Reference Manual," Report DNA-4524F to Defense Nuclear Agency, Lockheed Palo Alto Research Laboratory, Palo Alto, Calif., 1978.

8 Felippa, C. A., "Top Down Derivation of Doubly Asymptotic Approximations for Structure-Fluid Interaction Analysis," *Proceedings of the Second International Symposium on Innovative Numerical Analysis in Applied Engineering Science,* Montreal, Canada, June 1980, pp. 79–88.

9 Sobolev, S. L., *Partial Differential Equations of Mathematical Physics,* Pergamon Press, London, 1964.

10 Mindlin, R. D., and Bleich, H. H., "Response of an Elastic Cylindrical Shell to a Transverse Step Shock Wave," ASME JOURNAL OF APPLIED MECHANICS, Vol. 20, 1953, pp. 189–195.

11 Haywood, J. H., "Response of an Elastic Cylindrical Shell to a Transverse Shock Wave," *Quarterly Journal of Mechanics and Applied Mathematics,* Vol. 11, 1958, pp. 129–141.

12 Bedrosian, B., and DiMaggio, F. L., "Acoustic Approximations in Fluid-Shell Interactions," *Journal of the Engineering Mechanical Division, ASCE,* Vol. 98, 1972, pp. 731–742.

13 Felippa, C. A., and Park, K. C., "Staggered Solution Procedures for Coupled Mechanical Systems: Formulation," *Comp. Meth. Appl. Mech. Eng.,* 1980 (to be published).

14 Junger, M. C., and Feit, D., *Sound, Structures and Their Interaction,* The MIT Press, Cambridge, Mass., 1972.

15 Baker, G. A., *Essentials of Padé Approximants,* Academic Press, New York, 1975.

16 Struik, D. J., *Lectures on Classical Differential Geometry,* Addison-Wesley, Reading, Mass., 1961.

APPENDIX *A*

Padé Early-Time Approximations

A powerful technique for extending the range of applicability of a one-sided asymptotic power-series expansion is transformation to Padé approximants [15], which are instances of the so-called rational-function or continued-fraction approximations. The technique will be illustrated here for the circular cylinder. The Laplace transform of ETA_4, with s normalized to c/a, is (cf. Table 1):

$$\hat{p} + \frac{1}{2}\hat{p}s^{-1}\left(1 + \frac{3}{8}s^{-2} + \frac{135}{128}s^{-4}\right) = \rho c\hat{u}\left(1 + \frac{1}{8}s^{-2} + \frac{27}{128}s^{-4}\right) \quad (27)$$

Express the terms in parenthesis as quotients of quadratic polynomials, and match their descending expansions in s with (27). The result is the Padé (2, 2) approximant:

$$\hat{p}\left(1 + \frac{1}{2}s^{-1}\frac{s^2 - 39/16}{s^2 - 45/16}\right) = \rho c\hat{u}\left(\frac{s^2 - 25/16}{s^2 - 27/16}\right) \tag{28}$$

Fig. 6 shows components of the specific acoustic impedance of (28) and of the exact impedance $\zeta = jH_0(\omega a/c)/H_0'(\omega a/c)$, $\omega = jcs/a$; $H_n(x)$ being the Hankel function of first kind and order n. Comparing to Fig. 3 it is seen that (28) does not collapse at low frequencies as (27) does; furthermore, it retains good accuracy (~3 digits) for $\omega a/c > 1.5$. One computational drawback of the Padé reformulation is that all pressure-integral terms survive on rationalizing (28) and transforming back to the time domain; whereas about half such terms vanish on the left-hand side of equation (25).

The effectiveness of this technique for more complex geometries such as the torus is presently difficult to assess, as exact impedance solutions are unknown.

APPENDIX *B*

Derivation of Surface Metric Coefficients

The expression for the element of area of an (x, y, z) space surface described in terms of surface coordinates (R, θ) is [16, Chapter 2]:

$$dS = \{(x_R^2 + y_R^2 + z_R^2)(x_\theta^2 + y_\theta^2 + z_\theta^2) - (x_R x_\theta + y_R y_\theta + z_R z_\theta)^2\}^{1/2} dR d\theta. \tag{29}$$

Replacement of x and y by $r \cos\theta$ and $r \sin\theta$, respectively, yields equation (11). Derivation of the next expression starts from the cosine formula,

$$\beta = \cos(\mathbf{n}, \mathbf{R}) = \{x(y_R z_\theta - y_\theta z_R) + y(z_R x_\theta - z_\theta y_R) + z(x_R y_\theta - x_\theta y_R)\}/(W_1 R)$$

$$= -r(z_R r - z r_R)/(W_1 R) = -r^3 \frac{\partial}{\partial R}(z/r)/(W_1 R), \tag{30}$$

from which (12) follows. (The author has been unable to find this remarkably compact expression in differential geometry textbooks.)

M. P. Paidoussis
Professor and Chairman.
Mem. ASME.

N. T. Issid
Research Assistant.

M. Tsui
Research Assistant.

Department of Mechanical Engineering,
McGill University,
817 ouest, rue Sherbrooke,
Montreal, Québec, H3A 2K6,
Canada

Parametric Resonance Oscillations of Flexible Slender Cylinders in Harmonically Perturbed Axial Flow

Part 1: Theory

This paper studies theoretically the dynamical behavior of a flexible slender cylinder in pulsating axial flow. The dynamics of the system in steady, unperturbed flow are examined first. For various sets of boundary conditions the eigenfrequencies of the system at any given flow velocity are determined, and the critical flow velocities are established, beyond which buckling (divergence) would occur. The behavior of the system in pulsating flow is examined next, establishing the existence of parametric resonances. The effects of the mean flow velocity, boundary conditions, dissipative forces, and virtual (hydrodynamic) mass on the extent of the parametric instability zones are then discussed.

Introduction

The study of flow-induced vibrations and fluid-elastic instabilities of structural components has been intensified in recent years. This is partly due to the increased need for reliability, especially in the power generating industry where repeated equipment failures have evidenced the inadequate state of the art [1]. Thus it has now become imperative to try to understand and to be able to predict the dynamical behavior of slender flexible structures in flow, such as might be found in boilers, nuclear reactors, heat exchangers, and steam generators [2]. Although most failures are associated with conditions of crossflow, cases of axial flow have also been shown to be of importance. However, quite apart from practical considerations, these problems are of sufficient intrinsic interest, in the realm of dynamics of systems subjected to circulatory or gyroscopic forces, to merit study for their own sake. This paper deals with an idealized model of a system, not directly related to any particular problem.

The study of the dynamics of slender cylinders in *steady* axial flow began seriously in the 1960's [3, 4], and has progressed impressively since then [5, 6]. This is not the case when the flow is unsteady, with only one previous attempt to study the problem theoretically, when the flow is harmonically perturbed [7].

Harmonic perturbations to the mean flow may arise both in internal (contained) flows and in the atmospheric or oceanic environment. Thus, in circulating systems, such perturbations may occur, e.g., through the action of a pump, as a result of thermohydraulic instabilities in two-phase flows, or by periodic vortex shedding somewhere upstream of the structure of interest. This last mechanism may also be at work in the case of atmospheric flows. Finally, harmonic perturbations may arise about ocean structures, as a result of wave action.

Whenever such periodic perturbations are present in the axial flow about cylindrical structures, there exists the distinct possibility that they may cause parametric resonances, otherwise known as parametric instabilities. It has been shown [7] that such resonances may occur if the circular frequency of the periodic flow component, Ω, lies in the vicinity of a fractional multiple of one of the natural frequencies of the cylinder, Ω_n, i.e., if $\Omega \simeq 2\Omega_n/k$, where $k = 1, 2, 3 \ldots$. The most important of these resonances, the so-called principal primary parametric resonance, occurs when $k = 1$, so that $\Omega \simeq 2\Omega_n$—a well-established result from the analogous problem of a column subjected to a harmonically perturbed axial load [8–10].

This study is presented in two parts. This paper, Part 1 of the study, presents the theoretical model and its predictions of the dynamical behavior of the system. Part 2 [11] describes the complementary experimental program.

The Equation of Motion

Consider a flexible slender cylinder immersed in an incompressible

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, June, 1980. Paper No. 80-WA/APM-30.

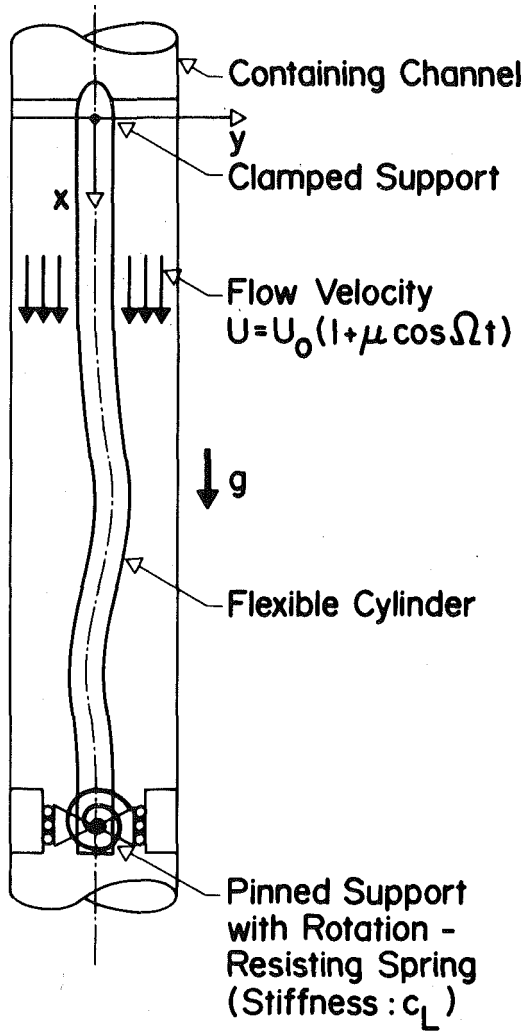


Fig. 1 Definition diagram of the system under consideration

fluid, flowing with velocity U parallel to the x-axis, which coincides with the center line of the cylinder at rest (Fig. 1). The flow velocity is composed of a steady component, U_0, and of a small oscillating component, $\mu U_0 \cos \Omega t$, where μ and Ω are sufficiently small for the flow to be essentially uniform within the flow channel at any given instant. The cylinder is considered to be clamped at its upstream end. The downstream end is pinned, but rotation at that point is resisted by a spring; hence, through the use of a very large or a vanishingly small spring constant, one may retrieve an ideally clamped or pinned downstream end, respectively.

Consider next, the forces to which the cylinder is subjected, as a result of small lateral displacements, $y(x, t)$. In addition to the usual forces associated with motions of an Euler-Bernoulli beam, which models the cylinder itself, gravity and hydrostatic forces will also be considered, as well as the hydrodynamic forces; the latter will be distinguished as inviscid and viscous, as the case may be, for ease of formulation.

For an element δx of the cylinder, according to inviscid slender-body theory [3], there will be a normal force

$$F_A \delta x = [(\partial/\partial t) + U(\partial/\partial x)]\{\chi \rho A[(\partial y/\partial t) + U(\partial y/\partial x)]\}\delta x, \quad (1)$$

where $\chi \rho A$ is the virtual, or hydrodynamic, mass of the fluid per unit length, ρ being the fluid density, and A the cylinder cross-sectional area. For an unconfined fluid, χ is equal to unity, and the virtual mass takes on its classical value of ρA; for confined fluid, larger fluid accelerations associated with cylinder motions result in $\chi > 1$ [12].

Assuming turbulent flow, the viscous hydrodynamic forces may be approximated [3] by

$$F_L \delta x = \tfrac{1}{2}\rho D U^2 C_f \delta x \quad \text{and}$$
$$F_N \delta x = F_L\{[(\partial y/\partial t) + U(\partial y/\partial x)]/U\}\delta x \quad (2)$$

in the longitudinal and normal directions, respectively, D being the cylinder diameter and C_f the frictional coefficient.

Finally, the hydrostatic forces in the axial and lateral directions may be expressed by

$$F_p^x \delta x = 0$$

and

$$F_p^y \delta x = (\partial/\partial x)[pA(\partial y/\partial x)]\delta x, \quad (3)$$

where p is the static pressure. Furthermore, assuming that the pressure varies linearly along the channel, the pressure drop may be expressed as

$$-A(\partial p/\partial x) \simeq F_L(D/D_h) - \rho g A + \rho A(dU/dt), \quad (4)$$

where D_h is the hydraulic diameter of the annular flow and g is the acceleration due to gravity [12].

The problem is simplified here by assuming that there is no constraint to axial extension of the cylinder at the downstream support (i.e., axial sliding is permitted there), so that external tensioning of the cylinder is not possible, and effects due to the mean external pressure do not arise (cf. [12]). Denoting next the flexural rigidity of the cylinder by EI, its viscoelastic Kelvin-Voigt dissipative coefficient by E^*, its length by L and mass per unit length by m, and the form drag coefficient at the downstream end by C_b, the force and moment balance equations may be written. Utilizing then equations (1)–(4), the equation of small lateral motions may be obtained, i.e.,

$$\begin{aligned} E^*I \frac{\partial^5 y}{\partial t \partial x^4} + EI\frac{\partial^4 y}{\partial x^4} + \chi\rho A\left\{\frac{\partial^2 y}{\partial t^2} + 2U\frac{\partial^2 y}{\partial x \partial t} + U^2\frac{\partial^2 y}{\partial x^2} + \frac{dU}{dt}\frac{\partial y}{\partial x}\right\} \\ - \left\{\left[\frac{1}{2}\rho D U^2 C_f\left(1 + \frac{D}{D_h}\right) + (m - \rho A)g + \rho A\frac{dU}{dt}\right](L - x) \right. \\ \left. + \frac{1}{2}\rho D^2 U^2 C_b\right\}\frac{\partial^2 y}{\partial x^2} + \frac{1}{2}\rho D U C_f\left\{\frac{\partial y}{\partial t} + U\frac{\partial y}{\partial x}\right\} \\ + \left\{(m - \rho A)g + \rho A\frac{dU}{dt} + \frac{1}{2}\rho D U^2 C_f\frac{D}{D_h}\right\}\frac{\partial y}{\partial x} \\ + m\frac{\partial^2 y}{\partial t^2} = 0. \quad (5) \end{aligned}$$

The cylinder is subject to the boundary conditions

$$y(0, t) = y(L, t) = 0$$

$$\frac{\partial y}{\partial x} = 0 \quad \text{at} \quad x = 0, \quad \text{and} \quad EI\frac{\partial^2 y}{\partial x^2} + c_L\frac{\partial y}{\partial x} = 0 \quad \text{at} \quad x = L, \quad (6)$$

where c_L is the spring constant of the spring at the lower support (Fig. 1).

Defining next the dimensionless parameters

$$\xi = x/L, \quad \eta = y/L, \quad \tau = \{EI/(m + \rho A)\}^{1/2}t/L^2,$$
$$\beta = \rho A/(\rho A + m), \quad u = (\rho A/EI)^{1/2}UL,$$
$$\alpha = \{I/[E(\rho A + m)]\}^{1/2}E^*/L^2,$$
$$\gamma = (m - \rho A)gL^3/EI, \quad \epsilon = L/D, \quad h = D/D_h, \quad c_f = (4/\pi)C_f$$
$$c_b = (4/\pi)C_b, \quad \kappa = c_L L/EI, \quad (7)$$

the equations of motion and boundary conditions may be written in dimensionless form, as follows:

$$\begin{aligned} \alpha\dot{\eta}^{\text{iv}} + \eta^{\text{iv}} + \{\chi u^2 - [\tfrac{1}{2}\epsilon c_f u^2(1 + h) + \gamma + \beta^{1/2}\dot{u}][1 - \xi] - \tfrac{1}{2}c_b u^2\}\eta'' \\ + 2\chi\beta^{1/2}u\dot{\eta}' + [\tfrac{1}{2}\epsilon c_f u^2(1 + h) + \gamma + (1 + \chi)\beta^{1/2}\dot{u}]\eta' \\ + \tfrac{1}{2}\epsilon c_f \beta^{1/2}u\dot{\eta} + [1 + (\chi - 1)\beta]\ddot{\eta} = 0, \quad (8) \end{aligned}$$

and

$$\eta(0, \tau) = \eta'(0, \tau) = \eta(1, \tau) = \eta''(1, \tau) + \kappa\eta'(1, \tau) = 0, \quad (9)$$

where dots and primes denote differentiation with respect to τ and ξ, respectively. These equations, together with the form of u considered,

$$u = u_0(1 + \mu \cos \omega\tau), \tag{10}$$

fully define the problem. Here ω is the dimensionless frequency, related to Ω by

$$\omega = \{(\rho A + m)/EI\}^{1/2}\Omega L^2. \tag{11}$$

After substituting (10) into (8), the equation of motion may be written in symbolic form

$$\alpha\dot{\eta}^{\text{iv}} + \eta^{\text{iv}} + (A_0 + A_1\xi)\eta'' + A_2\dot{\eta}' + A_3\eta' + A_4\dot{\eta} + A_5\ddot{\eta} = 0, \tag{12}$$

where A_0 to A_5 are obvious groupings of terms; e.g., $A_1 = -\beta^{1/2}u_0\mu\omega \times \sin \omega\tau + \gamma + \frac{1}{2}\epsilon c_f(1 + h)u_0^2(1 + \mu \cos \omega\tau)^2$.

Methods of Analysis

The continuous system is first rendered discrete by the use of the Ritz-Galerkin method, where $\eta(\xi, \tau)$ is expressed as a series involving the eigenfunctions $\phi_r(\xi)$ of a beam with the same boundary conditions as the cylinder under consideration, i.e.,

$$\eta(\xi, \tau) = \sum_{r=1}^{\infty} \phi_r(\xi)q_r(\tau). \tag{13}$$

The $\phi_r(\xi)$ and the corresponding eigenvalues, λ_r, both of which are functions of κ, are fairly easy to determine. Substituting into the equation of motion, applying the Ritz-Galerkin method, and truncating the series at a suitable $r = N$, one obtains the matrix equation

$$A_5\mathbf{I}\ddot{\mathbf{q}} + [\alpha\mathbf{\Lambda} + A_4\mathbf{I} + A_2\mathbf{B}]\dot{\mathbf{q}} + [\mathbf{\Lambda} + A_0\mathbf{C} + A_1\mathbf{D} + A_3\mathbf{B}]\mathbf{q} = \mathbf{0}, \tag{14}$$

where $\mathbf{I}$ is the unit matrix, $\mathbf{\Lambda}$ is the diagonal matrix with elements λ_r^4, $\mathbf{q} = \{(q_1, q_2, \ldots q_N\}^T$, and $\mathbf{B}, \mathbf{C}, \mathbf{D}$ are square matrices whose elements are given, respectively, by

$$b_{sr} = \int_0^1 \phi_s\phi_r' d\xi, \quad c_{sr} = \int_0^1 \phi_s\phi_r'' d\xi, \quad d_{sr} = \int_0^1 \phi_s\xi\phi_r'' d\xi. \tag{15}$$

If the flow is steady, i.e., if $\mu = 0$, then the coefficients A_0 to A_5 are scalar constants, and equation (14) may easily be transformed to a standard eigenvalue problem, from which the dimensionless eigenfrequencies of the system, ω_n, may be determined. These will generally be complex, so that

$$\omega_n = \text{Re}\,(\omega_n) + \text{Im}\,(\omega_n)i. \tag{16}$$

If the flow is harmonically perturbed, $\mu \neq 0$, then the parametric resonance regions may be determined by several different methods [8–10]. Here the method described by Bolotin [8] will be utilized. One may categorize the resonances as primary and secondary, the former occurring in the vicinity of $\omega/\omega_n = 2, \frac{2}{3}, \frac{2}{5}, \ldots$, while the latter in the vicinity of $\omega/\omega_n = 1, \frac{1}{2}, \frac{1}{3}, \ldots$, for an undamped system; for a damped system, involving complex ω_n's, the aforementioned ratios are associated with $\omega/\text{Re}\,(\omega_n)$. As mentioned in the Introduction, the most important parametric resonance region lies in the neighbourhood of $\omega/\text{Re}\,(\omega_n) = 2$ and is called the principal primary region. In addition to the foregoing, there is another class of parametric instabilities, the so-called combination resonances, which occur at ω equal to fractional combinations of two or more ω_n's; these cannot be determined by Bolotin's method and will not be considered here.

To determine the regions of parametric resonance, Bolotin's method proceeds as follows:

(*i*) $\mathbf{q}$ is expressed as

$$\mathbf{q} = \sum_{k=1,3,5\ldots}^{\infty} \{\mathbf{a}_k \sin (\tfrac{1}{2}k\omega\tau) + \mathbf{b}_k \cos (\tfrac{1}{2}k\omega\tau)\}; \tag{17}$$

(*ii*) Equation (17), truncated at a suitable $k = K$, is substituted into (14) and terms in $\sin [\frac{1}{2}(k + l)\omega\tau]$ and $\cos [\frac{1}{2}(k + l)\omega\tau]$ are collected, yielding

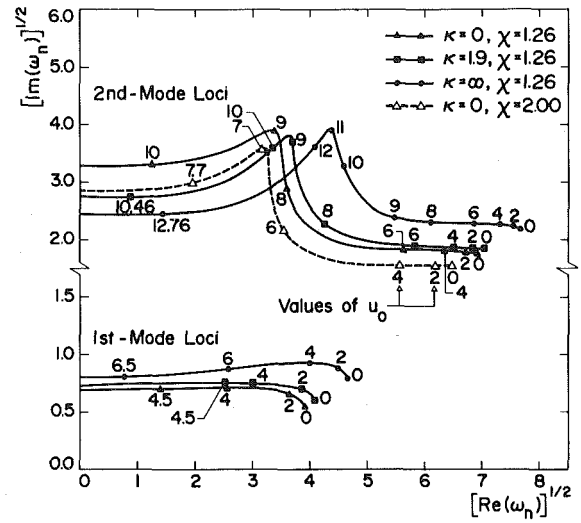


Fig. 2 Argand diagrams of the loci of the first and second-mode eigenfrequencies of a slender cylinder with different boundary conditions (different κ), for system parameters $\beta = 0.467$, $\gamma = 3.98$, $\alpha = 0.0028$, $\chi = 1.26$ or $\chi = 2.00$, $h = 0.519$, $\epsilon c_f = 0.5$, $c_b = 0.2$

$$\sum_l \sum_{k=1,3,\ldots}^{K} \{[\mathbf{H}_{lk}^1\mathbf{a}_k + \mathbf{H}_{lk}^2\mathbf{b}_k] \sin [\tfrac{1}{2}(k + l)\omega\tau] + [\mathbf{J}_{lk}^1\mathbf{a}_k + \mathbf{J}_{lk}^2\mathbf{b}_k] \cos [\tfrac{1}{2}(k + l)\omega\tau]\} = \mathbf{0}, \tag{18}$$

where, in this case, $l = -4, -2, 0, 2, 4$; $\mathbf{H}_{lk}^1$, $\mathbf{H}_{lk}^2$, $\mathbf{J}_{lk}^1$, $\mathbf{J}_{lk}^2$ are square matrices which are functions of A_0 to A_5 and $\mathbf{I}, \mathbf{\Lambda}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$;

(*iii*) Expanding equation (18) for k, and collecting terms in $\sin (\frac{1}{2}p\omega\tau)$ and $\cos (\frac{1}{2}p\omega\tau)$, $p = 1, 3, \ldots, K + 4$, one obtains the matrix equation

$$[\mathbf{G}] \begin{Bmatrix} \mathbf{a}_k \\ \mathbf{b}_k \end{Bmatrix} = \mathbf{0}; \tag{19}$$

(*iv*) Setting det $[\mathbf{G}] = 0$, the boundaries of the regions of parametric resonance are defined by the ω's satisfying this equation.

The secondary resonance regions may be determined in a similar manner, except that equation [17] then involves an even trigonometric series.

The Behavior of the System in Steady Flow

It is essential to consider the behavior of the system in steady flow, in order to be able to interpret its behavior in pulsatile flow.

Fig. 2 shows the loci of the first and second-mode eigenfrequencies, as functions of the dimensionless flow velocity u_0, for a typical cylinder in axial flow; the calculations were done for three values of κ, $\kappa = 0$, 1.9[1] and ∞, the first and last cases corresponding to a classical pinned and clamped support, respectively. The loci are plotted as Argand diagrams. The choice of the square roots of Re (ω_n) and Im (ω_n) as coordinates is simply for clarity of graphical presentation of the results; the conversion may readily be made, e.g., for $\kappa = 1.9$ and $u_0 = 4$, one has $\omega_1 = 8.96 + 0.562i$ and $\omega_2 = 42.1 + 3.43i$. These results were obtained with eight terms ($N = 8$) in the Galerkin expansion of equation (13).

Let us first consider the three sets of curves with $\chi = 1.26$. For $u_0 = 0$, the Re (ω_n) associated with $\kappa = 0$ and $\kappa = \infty$ are close to those of a classical beam ($\alpha = \gamma = 0, \chi = 1$) with the same boundary conditions; e.g., here we have Re $(\omega_1) = 15.05$ and 21.60 for $\kappa = 0$ and ∞, respectively, as compared to Re $(\omega_1) = \omega_1 = 15.42$ for a classical clamped-

[1] The value $\kappa = 1.9$, as well as the other system parameters may appear odd values to choose; however, they correspond to values of direct interest in the experiments of Part 2[11].

pinned beam and 22.37 for a clamped-clamped one. The differences are due to the effect of gravity on the cylinder ($\gamma \neq 0$), on the one hand, which tends to raise the frequency, and the effect of dissipation ($\alpha \neq 0$) and confinement ($\chi > 1$), on the other, tending to lower the frequency. The fact that Im $(\omega_n) > 0$ at $u_0 = 0$ reflects the existence of dissipation ($\alpha \neq 0$), which according to the Kelvin-Voigt model is proportional to frequency, so that Im $(\omega_2) >$ Im (ω_1).

With increasing u_0, following effects are noted:

(*i*) The oscillation frequencies, Re (ω_n), are reduced.

(*ii*) The damping ratio, which is proportional to Im (ω_n)/Re (ω_n), is increased.

At flow velocities larger than shown, the loci bifurcate on the Im (ω)-axis, and one branch eventually becomes associated with Im $(\omega_n) < 0$, which indicates the onset of buckling (divergence), first in the first mode and then in the second; at yet higher u_0, coupled mode flutter is also possible [12]. It is of interest that the behavior of the system with an intermediate κ is qualitatively similar to that of an ideally clamped or pinned downstream end.

Also shown in Fig. 2 is the second mode of a clamped-pinned system with an appreciably higher value of χ, $\chi = 2$. It is seen that an increased virtual mass results in lowering both Re (ω_2), and hence, Im (ω_2). Moreover, the system loses stability in this case at considerably lower flow velocity, an effect not shown in this figure, but one readily perceived upon noting that for a given u_0 the locus is much closer to the origin for $\chi = 2$ than for $\chi = 1.26$.

The eigenfrequencies obtained here will be used in the next section in two ways: first, in identifying the regions of parametric resonance to be found, by comparing the Re (ω_n) at a given u_0 to the pulsation frequency ω; second, in assessing the effective damping at a given u_0 on the extent of these resonance regions, where a measure of this damping is the logarithmic decrement δ_n, defined by $\delta_n = 2\pi$ Im (ω_n)/Re (ω_n).

The Behavior of the System in Pulsatile Flow

The principal primary and secondary regions of parametric resonance, associated with the first and second modes of a typical system, are shown in Fig. 3, for two values of the dimensionless mean flow velocity, $u_0 = 2.30$ and $u_0 = 3.61$; indeed these results are for one of the systems, the behavior of which in steady flow is shown in Fig. 2 ($\kappa = 0$, $\chi = 1.26$). These calculations were conducted with $N = 5$ and $K = 3$ or $K = 2$.

Considering the resonance regions in the case of $u_0 = 2.30$, it is noted that the upper one is centered about $\omega = 92$; this, compared to Re $(\omega_1) = 12.49$ and Re $(\omega_2) = 44.34$, indicates that this region is indeed the principal primary region associated with the second mode, as $\omega \simeq 2$ Re (ω_2); (the discrepnacy, of about 4 percent, is due to the fact that these calculations were done with $N = 5$, while those of Fig. 2 with $N = 8$). Similarly, the larger lower region, centered about $\omega = 25 \simeq 2$ Re (ω_1), may be identified as the principal primary region associated with the first mode.

Finally, the small lowermost region is centered about $\omega \simeq 12$ and is clearly the principal *secondary* region associated with the first mode. No such region corresponding to the second mode exists for the range of μ considered. The main reason for this is that the effective damping associated with the second mode is higher than for the first one; i.e., the logarithmic decrement of the first mode is $\delta_1 = 0.23$, while that of the second is twice as large, or $\delta_2 = 0.46$. This, in turn, is a function of the dissipative model adopted in this theory, whereby dissipation is proportional to frequency.

Considering next, the resonance regions of Fig. 3 associated with the higher flow velocity, $u_0 = 3.61$, the various regions may be identified in the same way as before. Comparing the two sets of regions, it is noted that for the higher u_0 the regions are larger. (Yet, there is still no second-mode secondary region for $\mu < 0.7$.) The reason for this is that the amplitude of the pulsatile force is not μ, but μu_0; hence, everything else being equal, one would expect larger regions, the larger the value of u_0. Of course, everything else is not quite equal: the effective damping for $u_0 = 3.61$ is higher than for $u_0 = 2.30$ (i.e., $\delta_1 = 0.37$

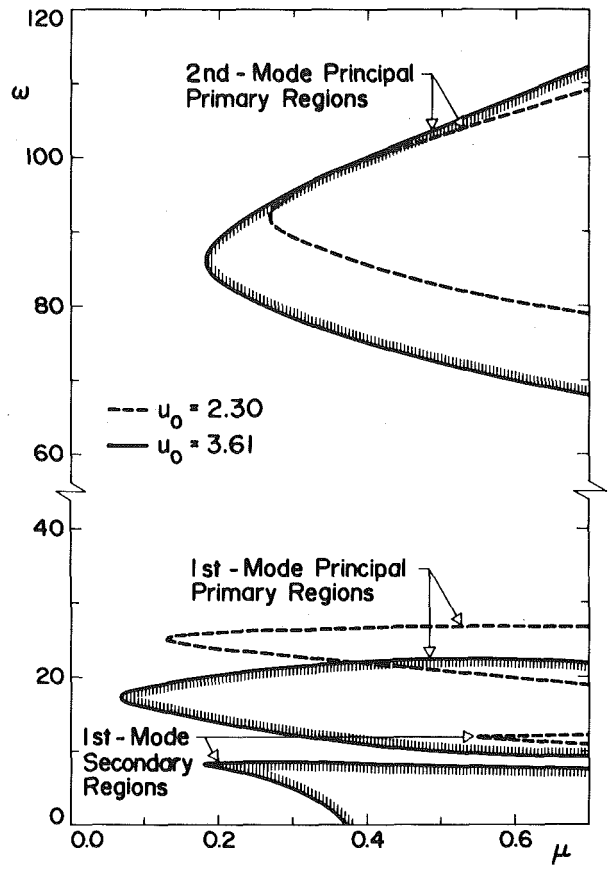


Fig. 3 The effect of the dimensionless mean flow velocity, u_0, on the principal primary and secondary parametric resonance regions, associated with the first and second modes of a clamped-pinned cylinder ($\kappa = 0$, $\chi = 1.26$; other parameters as in Fig. 2)

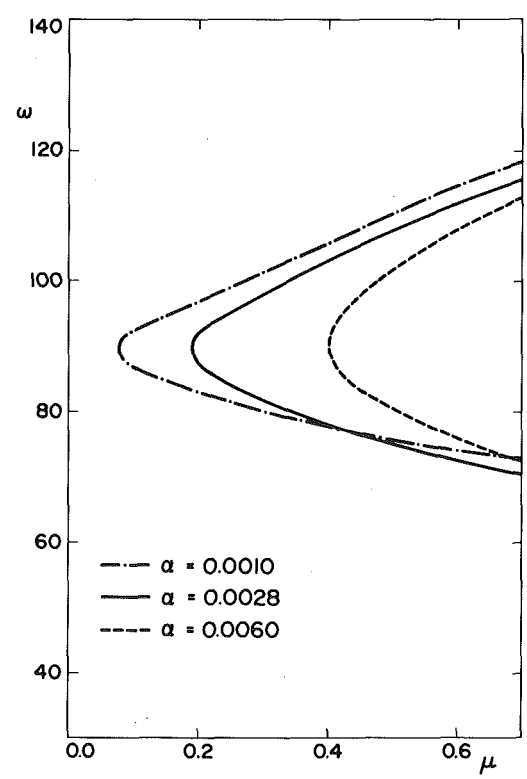


Fig. 4 The effect of viscoelastic damping on the principal primary resonance regions associated with the second mode of a system with $\kappa = 1.9$, $\chi = 1.26$, $u_0 = 3.61$, α as shown, and the other parameters as in Fig. 2

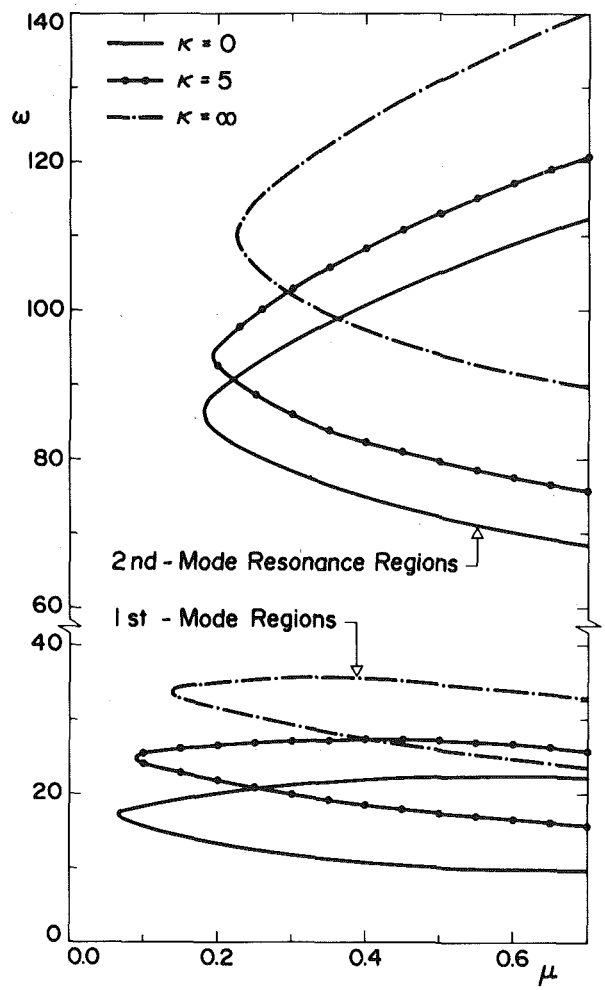


Fig. 5 The principal resonance regions associated with the first and second modes of a cylinder with different boundary conditions (different κ), for a system with $\beta = 0.467$, $\gamma = 3.98$, $\alpha = 0.0028$, $\chi = 1.26$, $h = 0.519$, $\epsilon c_f = 0.5$, $c_b = 0.2$, and $u_0 = 3.61$

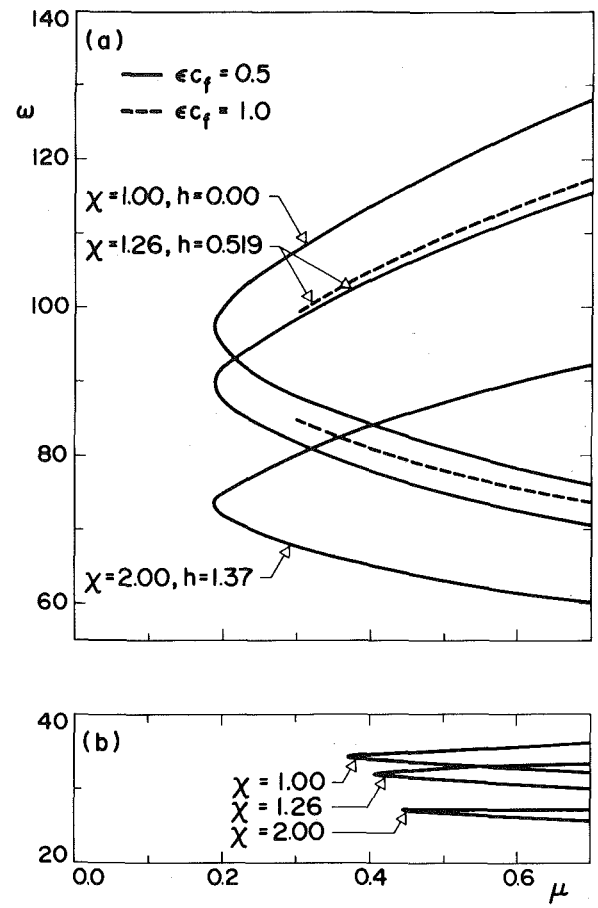


Fig. 6 The effect of confinement and frictional forces, characterized by the parameters χ, h, and ϵc_f; (*a*) second-mode principal primary resonance regions for a system with $\kappa = 1.9$, χ, h, and ϵc_f as shown, and the other parameters as in Fig 5, (*b*) first-mode principal primary resonance regions for the same system, except $u_0 = 1.00$; the values of h corresponding to each χ are the same as in the upper figure

versus 0.23), which moderates the effect of higher u_0. Nevertheless, this effect of u_0 on the size of resonance regions apparently holds true in general, and is not confined to the results shown here; thus, for $u_0 = 4.5$, the second-mode principal primary region is quite enormous, extending from $\omega = 48$ to 108, approximately, for $\mu = 0.7$; in contrast, for $u_0 = 0.5$ this second-mode region disappears altogether, and the first-mode principal primary region becomes very small, extending from $\omega = 29$ to 30, approximately, at $\mu = 0.7$.

In all the foregoing it is seen that a minimum value of μ, μ_{min}, is required for parametric resonance oscillations to be generated. Indeed, the results of Fig. 3 show that μ_{min} may be quite considerable; e.g., for $u_0 = 2.30$, one has $\mu_{min} > 0.1$, i.e., the harmonically pulsating component must be greater than 10 percent of the mean flow velocity. As is shown in Fig. 3, μ_{min} decreases with increasing u_0, which is another manifestation of the effect of u_0 in increasing the overall size of the instability regions.

The values of μ_{min} depend critically on dissipation, as may be seen in Fig. 4. Here it is seen that not only is μ_{min} increased with increasing α, but the overall extent of the resonance region in the (μ,ω)-plane is substantially reduced. To this extent these results are quite similar to the well-known behavior of clamped column beams subjected to a pulsatile end load. (In this connection it should be remarked that this analogy cannot be carried too far, because the system here under consideration is not conventionally conservative, but gyroscopic conservative, and such systems have been shown in the past to possess some unusual characteristics [13, 14].)

It should further be noted here that the values of α in Figs. 3 and 4 are atypically high, at least for metallic cylinders with not excessively dissipative supports (e.g., welded supports). In such cases the dissipative forces, normally expressed in terms of a hysteretic damping coefficient, would typically be one order of magnitude less, resulting in larger instability regions and smaller μ_{min}—i.e., much more like the region for $\alpha = 0.001$ in Fig. 4 than the others. The value of $\alpha = 0.0028$ used in Fig. 3 and elsewhere is typical for rubber-like materials, such as those used in the experiments of Part 2 [11].

The effect of boundary conditions is shown in Fig. 5. Apart from the downward shift of the regions as κ is decreased, reflecting the corresponding reductions in Re (ω_n) (cf. Fig. 2), it is seen that the parametric resonance regions are qualitatively quite similar. The progressively smaller values of μ_{min} as κ is reduced simply reflect the corresponding diminution of δ.

Finally, Fig. 6 shows the effect of the parameters χ, h, and ϵc_f. It is recalled that a higher χ implies a larger virtual mass through increased confinement of the flow about the cylinder. Physically, one way of doing this is by decreasing the size of the annular flow passage (Fig. 1), which results also in an increased h, since the hydraulic diameter becomes smaller. Denoting the channel diameter by D_{ch}, one may easily find that

$$\chi = \frac{1 + (D/D_{ch})^2}{1 - (D/D_{ch})^2}, \quad h = \frac{D/D_{ch}}{1 - D/D_{ch}},$$

where the expression for χ was obtained from reference [15].

The effect of increased confinement is to reduce the extent of the parametric resonance regions. In the case of the principal primary

regions associated with the second mode and high flow velocity ($u_0 = 3.61$) shown in Fig. 6(a), this effect is not very strong. However, in other cases it is, as is shown for instance in Fig. 6(b) for the corresponding regions associated with the first mode of the system at low flow velocity ($u_0 = 1.00$).

The effect of ϵc_f, as well as that of changing h while keeping χ constant (not shown here), is not very large. As expected, however, increasing ϵc_f decreases the size of the resonance regions; the reason why this effect is not as pronounced as might have been expected is that ϵc_f on the one hand damps motions, and on the other increases the flow-induced tension of the cylinder, making it effectively stiffer.

Discussion

It has been established theoretically that a cylinder subjected to harmonically perturbed axial flow may develop parametric resonances (instabilities), over certain ranges of perturbation frequencies and amplitudes. Furthermore, the effect of some of the system parameters was investigated, to discover how importantly they affect the extent and location of the parametric resonance regions in the (μ,ω)-plane, and to obtain their effect on the minimum value of μ necessary for resonance oscillations, $\mu_{\min}$. The most important of these were found to be the mean flow velocity u_0 and the dissipative constant α (*vide* Figs. 3 and 4).

Concerning flow velocity, it was found that the higher u_0, the larger are the resonance regions. This simply reflects the fact that the amplitude of the pulsating flow, i.e., the excitation amplitude, is actually μu_0, and not μ. It is of interest to consider the form of the resonance regions of Fig. 3 when plotted as a (μu_0, ω)-diagram. In such a diagram the minimum pulsation amplitude, $\mu_{\min} u_0$ necessary for resonance oscillations would be similar in the two cases of $u_0 = 3.61$ and $u_0 = 2.30$ (e.g., for the second-mode regions $\mu_{\min} u \simeq 0.6$ in both cases), the remaining difference being associated with the different δ_2 involved. However, the width of the regions in the direction of the ordinate continues being larger for the larger u_0.

Concerning the width of the resonance regions referred to previously it is of interest that the second-mode regions are much larger than the first-mode ones (*vide* Figs. 3 and 5). This again is partly due to the form of presentation of the results. If the ordinate is changed from ω to ω/Re (ω_n)—Re (ω_n) being the real part of the eigenfrequency of the mode concerned at the pertinent u_0—clearly the width of the second-mode regions would diminish *vis-à-vis* that of the first-mode regions, since Re (ω_2) > Re (ω_1); indeed, generally but not always, the second-mode regions would then be slimmer than the first-mode ones.[2] From the physical point of view, presentation of the results in such "normalized" form, in the [μ, ω/Re (ω_n)]-plane, has some advantages, as it shows how much ω may fractionally differ from 2 Re (ω_n) and still give rise to parametric resonance oscillations; however, the simpler presentation in the (μ, ω)-plane has been preferred in this paper.

The second parameter which greatly affects the extent of the parametric resonance regions in the value of the dissipation constant α, an effect illustrated in Fig. 4. Another manifestation of this effect is that $\mu_{\min}$ for the first-mode regions is smaller than for the second-mode ones (Figs. 3 and 5). This is characteristic of the Kelvin-Voigt viscoelastic model utilized here, according to which damping is proportional to the frequency of oscillation. If a hysteretic (structural) dissipative model had been used, then $\mu_{\min}$ for first and second-mode regions would be similar.

The conclusions of this study, as well as the Acknowledgments, will be presented after the experiments of Part 2 [11] have been considered.

[2] This is exaggerated by the effect of dissipation, as already discussed.

References

1 *"Practical Experiences With Flow-Induced Vibrations—IAHR/IUTAM Symposium; Karlsruhe 1979,* eds., Naudascher, E., and Rockwell, D., Springer-Verlag, Berlin, 1980.

2 Paidoussis, M. P., "Flow-Induced Vibrations in Nuclear Reactors and Heat Exchangers: Practical Experiences and State of Knowledge," in reference [1], pp. 1–81.

3 Paidoussis, M. P., "Dynamics of Flexible Slender Cylinders in Axial Flow. Part 1: Theory," *Journal of Fluid Mechanics*, Vol. 26, 1966, pp. 717–736.

4 Paidoussis, M. P., "Dynamics of Flexible Slender Cylinders in Axial Flow. Part 2: Experiments," *Journal of Fluid Mechanics,* Vol. 26, 1966, pp. 737–751.

5 Chen, S.-S., "Vibration of Nuclear Fuel Bundles," *Nuclear Engineering and Design,* Vol. 35, 1975, pp. 399–422.

6 Paidoussis, M. P., "The Dynamics of Clusters of Flexible Cylinders in Axial Flow: Theory and Experiments," *Journal of Sound and Vibration,* Vol. 65, 1979 pp. 391–417.

7 Paidoussis, M. P., "Stability of Flexible Slender Cylinders in Pulsatile Axial Flow," *Journal of Sound and Vibration,* Vol. 42, 1975, pp. 1–11.

8 Bolotin, V. V., *The Dynamic Stability of Elastic Systems,* Holden-Day, Inc., San Francisco, 1964.

9 Evan-Iwanowski, R. M., *Resonance Oscillations in Mechanical Systems,* Elsevier Scientific Publ. Co., New York, 1976.

10 Nayfeh, A. H., and Mook, D. T., *Nonlinear Oscillations,* Wiley, New York, 1979.

11 Paidoussis, M. P., Issid, N. T., and Tsui, M., "Parametric Resonance Oscillations of Flexible Slender Cylinders in Harmonically Perturbed Axial Flow. Part 2: Experiments," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, Dec. 1980, pp. 715–719.

12 Paidoussis, M. P., "Dynamics of Cylindrical Structures Subjected to Axial Flow," *Journal of Sound and Vibration,* Vol. 29, 1973, pp. 365–385.

13 Shieh, R. C., "Energy and Variational Principles for Generalized (Gyroscopic) Conservative Problems," *International Journal of Nonlinear Mechanics,* Vol. 5, 1971, pp. 495–509.

14 Paidoussis, M. P., "Flutter of Conservative Systems of Pipes Conveying Incompressible Fluid," *Journal of the Mechanical Engineering Science,* Vol. 17, 1975, pp. 19–25.

15 Fritz, R. J., "The Effect of Liquids on the Dynamic Motions of Immersed Solids," ASME *Journal of Engineering for Industry,* Vol. 94, 1972, pp. 167–173.

M. P. Paidoussis
Professor and Chairman.
Mem. ASME.

N. T. Issid
Research Assistant.

M. Tsui
Research Assistant.

Department of Mechanical Engineering,
McGill University,
817 ouest, rue Sherbrooke,
Montreal, Québec, H3A 2K6,
Canada

Parametric Resonance Oscillations of Flexible Slender Cylinders in Harmonically Perturbed Axial Flow

Part 2: Experiments

This paper examines experimentally the dynamical behavior of a flexible slender cylinder in axial flow, perturbed harmonically in time. Parametric resonance oscillations were found to exist over certain ranges of frequencies and amplitudes of flow-velocity perturbations. The most prominent of the resonances, in these experiments, were associated with the second-mode principal primary resonance, and were studied extensively. Agreement with theory was found to be quite good.

Introduction

This paper describes the experiments conducted in order to study the dynamical behavior of flexible slender cylinders in axial flow, the velocity of which is perturbed harmonically. It is a companion to the theoretical paper, which is Part 1 of this study [1]. It was shown by theory that, over specified ranges of frequency and amplitude of the pulsating component of flow, parametric resonances may occur.

The main aim of the experiments was (*i*) to observe the dynamical behaviour of the system, and (*ii*) to compare the observed behavior with that predicted by theory. This is believed to be the first experimental study of the subject.

Description of the Apparatus

The experiments were conducted with silicone rubber cylinders, 10 to 14 mm in diameter, 30 to 50 cm long, mounted centrally in a transparent cylindrical pipe, either 32 or 40 mm in diameter, in which a harmonically perturbed axial flow of water was maintained, as shown in Fig. 1.

To the straight-through steady flow was added a harmonically perturbed component provided by a "plunger pump," which was driven by a reciprocating mechanism connected to a variable-speed drive. Thus, both frequency and amplitude of the harmonic perturbation could be varied; the usable frequency range was 1 to 16 Hz. The flow velocity was measured just upstream of the test section by a hot-film anemometer. Traces of the periodically perturbed flow velocity showed that, provided the perturbations were not unduly large, the apparatus provided effectively truly sinusoidal perturbations to the flow, so that the flow velocity, U, could be represented by

$$U = U_0(1 + \mu \cos \Omega t), \tag{1}$$

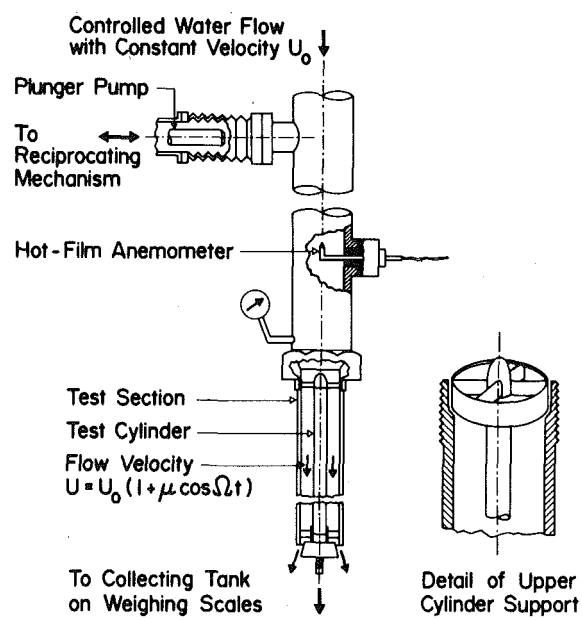


Fig. 1 Schematic diagram of the apparatus

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, June, 1980. Paper No. 80-WA/APM-31.

where U_0 is the mean flow velocity, and μU_0 is the amplitude of the perturbation, of circular frequency Ω; t is the time.

In the final form of the apparatus, adequate care was taken to straighten the flow coming from upstream, by means of screens, and to effect a smooth entry of the flow into the test section. (This, as will be discussed later, is very important.) The flexible bellows shown were used to isolate, as much as possible, the test section from vibration transmitted by the reciprocating mechanism.

All experiments were conducted with nominally clamped cylinder ends. The upper end was positively clamped: the cylinder was glued onto a brass "nose cone" (of a hemi-ellipsoidal shape); this, via four streamlined spider legs, was connected to a brass ring of the same internal diameter as the test-section pipe, which was clamped securely to the rest of the test section when the latter was screwed to the piping above. The lower support was a sleeve, into which the end of the test cylinder was inserted, normally to a depth of 7.5 mm, with a sliding fit. Thus the lower end allowed sliding in the axial direction. In practice, so as not to impede sliding, small lateral clearances were permitted in the sleeve, which, as will be discussed in the next section, resulted in a less-than-perfect clamped condition at the lower end.

Measurements, Preliminary Tests, and Procedures

The flexural rigidity, EI, and damping characteristics of the cylinder were determined by simple free-vibration experiments, using a fiber-optics sensor. Then, the cylinder was mounted in the (empty) test section and its first-mode natural frequency measured. It was found that it differed appreciably from that of the theoretical value for a clamped-clamped beam, computed with the previously found EI. Careful study revealed that this was associated with slight clearances, and with rotation at the lower support by bending of the cylinder inside the sleeve. Although in reality the support condition involved is probably complex and nonlinear, it was found adequate to model it as a pinned support with a rotational spring attached, which would resist rotation at the end—as was done in the theory of Part 1 [1]. The spring stiffness, c_L could then be determined in a way such that the measured first-mode natural frequency be equal to the theoretical value for such boundary conditions.

The theoretical circular first-mode natural frequency, Ω_1, was determined by Rayleigh's method. It may be obtained from the following expression:

$$(5\omega_1 + 1512 - 18\gamma)b^2 - (15\omega_1 + 2520 + 48\gamma)b + (12\omega_1 - 42\gamma) = 0, \quad (2)$$

where $\omega_1 = (mL^4/EI)^{1/2}\Omega_1$, $\gamma = mgL^3/EI$, $b = (\kappa + 4)/(\kappa + 6)$, and $\kappa = c_L L/EI$. In the foregoing, m is the mass of the cylinder per unit length and L is the free length of the cylinder. Hence, utilizing the experimental Ω_1 in conjunction with equation (2), one may find b, and hence κ.

Equation (2) was obtained through the use of the polynomial comparison function $y/L = [\xi^2 - (1 + b)\xi^3 + b\xi^4]$, which satisfies the boundary conditions; ξ is the axial distance divided by L. This equation was tested for some limiting cases, e.g., when $\gamma = 0$ and $c_L \to 0$ (clamped-pinned beam) or $c_L \to \infty$ (clamped-clamped beam); agreement for ω_1 was found to be better than 1 percent.

Conducting several experiments with different lengths of one cylinder, while taking due care to simulate the *in situ* experimental conditions (e.g., immersing the lower end in water, to simulate the conditions when the lower support is lubricated by the water flow), an average value of κ was obtained through the use of equation (2), namely, $\kappa = 1.9$. This is the value utilized in the theoretical calculations, when comparing theory with experiments.

In conjunction with the foregoing experiments aimed at determining EI and κ, the damping characteristics of the cylinders were also studied. It was found that, in the frequency range of interest, the damping of the material of the cylinder could adequately be represented by a hysteretic (structural) dissipative model. Since there are certain difficulties in utilizing hysteretic damping in the theory of Part 1, for the purposes of comparing theory with experiment, an equivalent viscoelastic damping coefficient was utilized; for a hysteretic damping coefficient ν, the equivalent viscoelastic coefficient α in the vicinity of the nth-mode dimensionless frequency ω_n, is given by $\alpha = \nu/\omega_n$ (cf. [2, 3]). Here α is *dimensionless* and is defined in [1], and $\omega_n = (mL^4/EI)^{1/2}\Omega_n$, where Ω_n is the circular frequency of the nth mode.

The system for measuring the flow velocity was calibrated under steady flow conditions. This was done by measuring the time necessary to collect a certain weight of water in the collecting tank beneath the test section, and simultaneously noting the voltage output from the electronics associated with the hot-film anemometer. Thus calibration curves of flow velocity in the annular passage versus voltage output could be constructed. The flow velocity in the annulus was determined from that at the measuring station simply by continuity.

When the flow had a pulsating component, its amplitude μU_0 could be determined from the flow-velocity trace on a storage oscilloscope or a fibre-optics-type chart recorder; then, from the calibration charts, the maximum voltage of the sinusoidal curve corresponds to $U_0 + \mu U_0$, while the minimum to $U_0 - \mu U_0$; hence, both U_0 and μ may be determined.

Prior to conducting the experiments, some tests were done to investigate the shape of the velocity profile in the flow annulus. In these tests the hot-film probe was mounted on the test-section itself, while the flexible cylinder was replaced by a rigid one. Traversing the probe across the annular gap, while the flow was steady, it was found that the flow profile was typical of a fully developed turbulent flow. The situation was investigated next with pulsating flow, in order to determine if the flow profile was deformed and if there were any phase differences introduced at different locations within the annular gap. To do this the signal from a wall-mounted pressure transducer was also recorded to give a reference for the phase angle. Over the pulsation frequency range of which the apparatus was capable and the range of pulsation amplitudes used in the experiments, no phase difference could be found; the velocity profile in the presence of pulsating flow retained the same shape as in steady flow.

All the foregoing were necessary to establish correspondence, at least for some key features, between the theoretical model and the experimental setup. In the case of the flow profile the theoretical model was found to be adequate, while in terms of the initially assumed boundary conditions—that both ends were ideally clamped—the theoretical model had to be adapted to take into account the less-than-perfect clamped lower support.

We shall next discuss the main experiments of this study, which were concerned with the dynamical behavior of the cylinder in the harmonically perturbed axial flow. The following procedure was used:

1 The flow-measuring system was calibrated as previously described.

2 With the stroke of the plunger pump and the mean flow velocity fixed, observations of the dynamical behavior of the cylinder were made, as the frequency of pulsation was increased gradually; at any point where there was a change of behavior, or an occurrence of special interest, traces of flow velocity versus time were obtained (yielding U_0, μ, and Ω).

3 The stroke was changed and step 2 repeated.

4 The system was recalibrated at the end of the experiment, to insure that the first calibration had remained valid.

Experiments at different flow velocities were then conducted with the same cylinder. This first set of experiments was usually conducted in the larger diameter test section; sometimes, experiments were then repeated in the narrower test section. Subsequently, in most cases, the length of the cylinder was reduced, by cutting a piece off the downstream end, and a new set of experiments conducted.

General Observations

The behavior of the system in pulsating axial flow will first be described qualitatively, before considering the results quantitatively.

At low frequencies, the system was stable. Very small oscillations could be observed, which were partly due to the response of the cylinder to the turbulent pressure field and partly to transmitted me-

chanical excitation. Increasing the pulsation frequency beyond a threshold value caused large amplitude harmonic oscillations to occur spontaneously; their amplitude increased with pulsation frequency, reached a maximum of up to 1½ cylinder diameters peak-to-peak, and then decreased, until it faded completely. The frequency of pulsation was noted to be twice the frequency of oscillation, and the modal form was similar to that of the second mode of the beam; so, these were likely principal primary resonance oscillations associated with the second mode, as will be shown definitively in the next section.

The ease of pin-pointing the onset of instability decreased as μ was diminished; this was because the amplitude of limit-cycle motion became smaller, so that there was no longer a step change in amplitude of oscillation at the threshold pulsation frequency, but merely a change in character of the vibration: from a random one to one having an appreciable, but difficult to separate, harmonic component.

Within the limits of the apparatus and the system parameters tested, this second-mode principal resonance was the only one that could be pin-pointed with accuracy. Possible reasons for this will be discussed later.

Several differences in the pattern of observations made in the foregoing should be noted. At high strokes of the plunger pump (high μU_0), there were bursts of organized periodic oscillation at pulsation frequencies inferior to the threshold frequency where regular parametric resonance oscillations established themselves. (Similar observations were made in experiments on parametric resonances of pipes conveying fluid [3].) Also, at very high values of μ, after the onset of parametric resonance, the oscillations persisted to the maximum frequency attainable by the apparatus; beyond a certain frequency, however, there sometimes was a change in character of the oscillation into what appeared to be a combination resonance (cf. [3].) At small strokes of the plunger pump, it was found that a certain time was necessary for resonance oscillations to develop; unless one allowed for this, the true threshold of resonance could be missed. Finally, and importantly, if the value of μ was sufficiently small, no parametric resonances could be detected at all.

Other general observations were the following:

(*i*) At high flow velocities, the level of turbulence was higher, and the threshold of instability was more difficult to pin-point.

(*ii*) The general behavior of the system was qualitatively the same in the two test-sections of different diameters, as well as for cylinders of different lengths and diameters.[1]

Let us now turn our attention to first-mode principal resonances, which, although predicted to exist by theory [1], were not observed in the experiments. Reasons for this may be discussed with the aid of Fig. 2, showing theoretical first and second-mode principal resonance regions for two different cylinders, involving different lengths and different dimensionless flow velocities (but the same dimensional flow velocity). Superimposed on these are traces of nondimensional pulsation frequency ω *versus* amplitude μ, as generated by the pulsating apparatus; the three strokes of the reciprocating mechanism shown, in the range investigated may be characterized as "small," "medium," and "very large." It should be mentioned that strokes greater than 2.5 cm were not normally used, one reason for this being that when combined with high frequencies they produced unacceptably high levels of mechanical vibration and deformed sinusoidal perturbations.

Fig. 2(*a*) is for a cylinder of length 39.5 cm, which is in the range normally tested in these experiments. It may be seen that the first-mode resonance region lies below the frequencies and amplitudes normally investigated, which partly explaines why first-mode resonances were not normally observed. Fig. 2(*b*) corresponds to a cylinder shorter than usual (28 cm), which was tested in an attempt to bring first-mode resonances within the (μ, ω)-range which the apparatus was comfortably capable of generating. It is seen that only with very high strokes, which were not normally used, would resonance oscil-

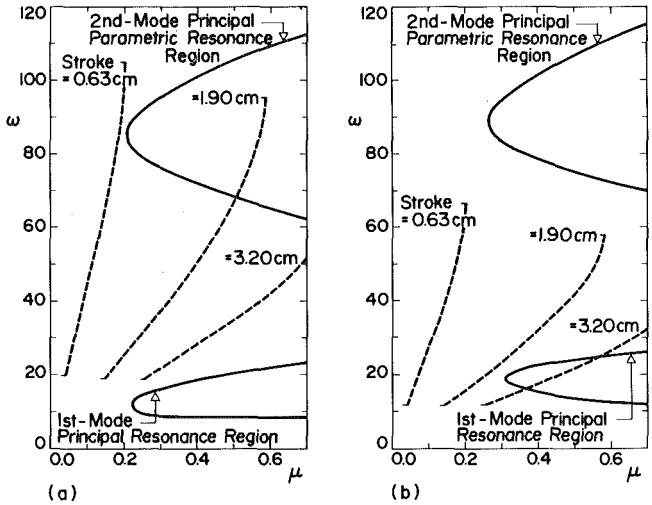


Fig. 2 Experimental pulsation frequencies and amplitude loci in the (μ,ω)-plane for three strokes of the reciprocating mechanism, and the theoretical boundaries of parametric resonances for a system with $\beta = 0.466$, $\kappa = 1.90$, $\chi = 1.25$, $h = 0.5$, $\epsilon c_f = 0.50$. (*a*) for a cylinder with $\gamma = 4.18$, hysteretic damping coefficient $\nu = 0.30$ (equivalent α for first and second mode: 0.025 and 0.0036), and $u_0 = 4.50$; (*b*) for a cylinder with $\gamma = 2.10$, $\nu = 0.36$, and $u_0 = 3.71$

lations occur. Nevertheless, when such high strokes *were* used, no clearcut first-mode resonance developed on the scale of amplitude and regularity of motion that is associated with second-mode resonance. Thus, although in the neighborhood of the theoretical region, there were definite indications of greater regularity of small quasi-harmonic motion in the generally stochastic response of the cylinder, no clearcut limit-cycle motion could be identified.

The conclusion that must be reached from the experiments of Fig. 2(*b*) is that there may indeed have been a small first-mode resonance region there, but either it was impeded from developing by interference from the relatively high turbulent buffeting (see also final paragraph of this section), or that it had an extremely small limit-cycle amplitude which could not be detected visually. This matter was not pursued at great length, and experiments concentrated on second-mode parametric resonance oscillations which were relatively easy to generate with the existing apparatus.[2]

A final comment on Fig. 2 is that the second-mode regions are larger *vis-à-vis* the first-mode ones than is typical in the figures of Part 1 [1]. The reason is that the calculations here were done for a hysteretically damped system, while those of Part 1 considered a Kelvin-Voigt viscoelastic dissipative model.

Finally, it is instructive to discuss some early experiments where the piping leading to the test section was of much smaller diameter than shown in Fig. 1; in fact, it was of smaller diameter than that of the test-section itself, so that there was a diffusing section in between. Moreover, no special care had been exercised to render uniform the flow entering the test-section. As a result, the mean flow about the cylinder was very noisy (indicating a high level of turbulence) and perhaps nonuniform as well. Following the usual experimental procedure, no parametric resonance oscillations whatsoever were observed with this setup, even in regions of the (μ, ω)-plane where theory predicted their existence [1] (and where they were subsequently found to exist after the apparatus was modified, to the form previously described). Instead, there was small random vibration of the cylinder throughout these tests, where the cylinder probably responded mainly to the random pressure fluctuations of the highly turbulent near field.

[1] This is true, provided these parameters were within the ranges given in the section, "Description of Apparatus;" otherwise, the regions of resonance could move beyond the frequency and amplitude capabilities of the apparatus.

[2] It must be stressed that these observations are not claimed to be applicable to all systems. Perhaps for another system with lower flow turbulence and/or lower internal dissipation, first-mode resonance oscillations would have been more prominent.

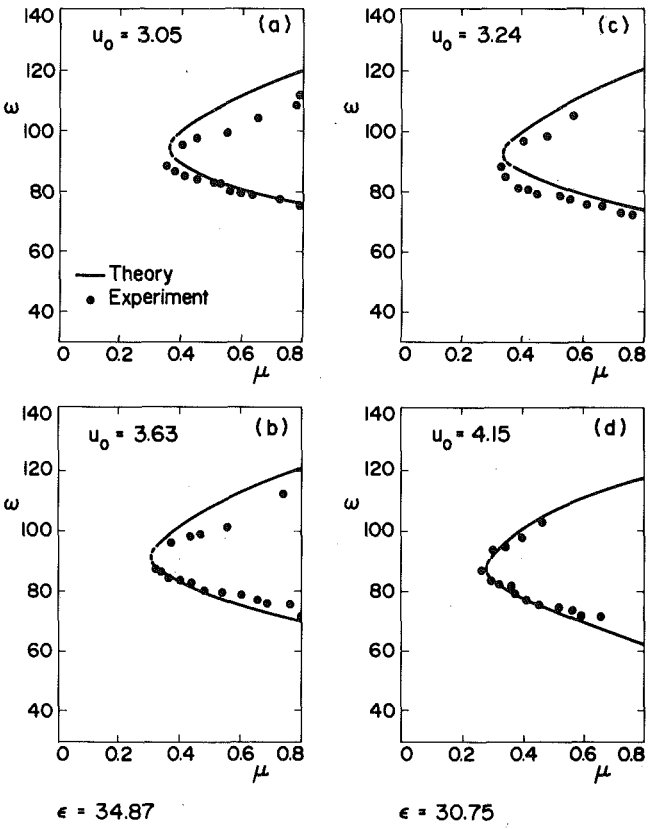


Fig. 3 Experimental boundaries of second-mode principal parametric resonance compared with theory; system parameters: $\beta = 0.465$, $\alpha = 0.0049$ ($\nu \simeq 0.49$), $\kappa = 1.9$, $\chi = 1.22$, $h = 0.462$; for (a) and (b) $\gamma = 5.45$ and for (c) and (d) $\gamma = 3.65$

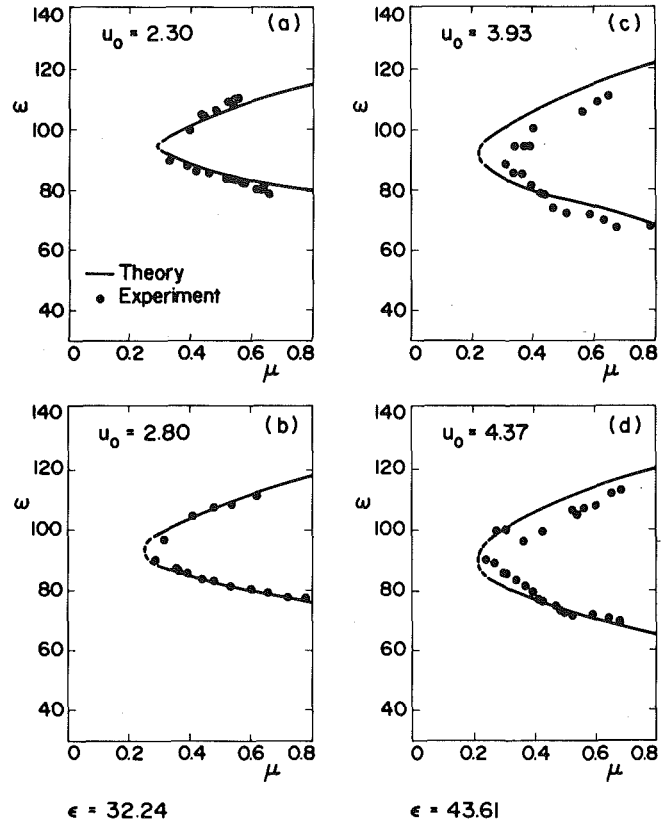


Fig. 4 Experimental boundaries of second-mode principal parametric resonance compared with theory; system parameters: for (a) and (b) $\beta = 0.465$, $\alpha = 0.0030$, $\kappa = 1.9$, $\chi = 1.265$, $h = 0.519$, $\gamma = 3.98$; for (c) and (d) $\beta = 0.446$, $\alpha = 0.0036$, $\kappa = 1.9$, $\chi = 1.236$, $h = 0.481$, $\gamma = 11.09$

Comparison Between Theory and Experiment

By comparing the band of frequencies of the observed limit-cycle resonance oscillations of the system to the theoretically predicted eigenfrequencies, it was quite clear that they were principal primary resonances associated with the second mode.

The theoretical and experimental boundaries of the resonance regions are compared quantitatively in Figs. 3 and 4. The results shown and the degree of accord between theory and experiment are typical of many more such sets of results obtained, but not shown here for the sake of brevity; a report will be published in the near future containing all such results.

The parameters used in these figures are nondimensional system parameters defined in Part 1 [1]. Thus

$\omega = [(m + \rho A)/EI]^{1/2}\Omega L^2$ = dimensionless frequency
$u_0 = (\rho A/EI)^{1/2}U_0 L$ = dimensionless mean flow velocity
$\epsilon = L/D$ = slenderness ratio

for cylinders of length L, diameter D, mass per unit length m, cross-sectional area A, and flexural rigidity EI, immersed in fluid of density ρ, flowing with mean velocity U_0. The pulsation amplitude μ is defined in equation (1), and Ω is the pulsation circular frequency. The other system parameters given in the captions are as follows:

α = dimensionless viscoelastic damping of the cylinder
κ = dimensionless spring constant of the effective rotational spring at the downstream end (cf. equation (2))
γ = dimensionless gravity parameter = $(m - \rho A)gL^3/EI$
β = mass parameter = $\rho A/(\rho A + m)$
χ = virtual (or hydrodynamic) mass coefficient, equal to 1 for unconfined flow, and >1 for confined (annular) flow
$h = D$/(hydraulic diameter of the annular flow)

The results shown in Fig. 3 were all obtained with the same cylinder (D = 12.6 mm) in the same test section (D_{ch} = 39.8 mm), but with different cylinder lengths: L = 36.1 cm in Figs. 3(c) and (d), while L = 41.3 cm in Figs. 3(a) and (b). Each case corresponds to a different flow velocity. For the shorter cylinder (figures on the right) the dimensional flow velocities are 20 to 30 percent higher—relatively larger than comparison of the dimensionless u_0 would indicate (*vide* definition of u_0 in the foregoing). Clearly, agreement between theory and experiment may be said to be reasonably good.

Similarly, Fig. 4 shows the principal second-mode instability regions for two cylinders: Figs. 4(a) and (b) are for a cylinder of diameter D = 13.6 mm and length L = 41.6 cm in the wider test section (D_{ch} = 40 mm); Figs. 4(c) and (d) are for a cylinder with D = 10.4 mm, L = 43.9 cm in the narrower test-section (D_{ch} = 32 mm). The two cases, therefore, involve different hydrodynamic mass, slenderness, and flow velocities. Once again, the degree of accord between theory and experiment is reasonably good.

In these experiments an attempt was made to vary the parameters as much as possible, in order to assess the effect of some of them on the degree of agreement between theory and experiment. From the results obtained it was concluded that there is no significant deterioration of this accord as any one of these system parameters was changed. Unfortunately, limitations of the apparatus did not permit a very wide variation of most parameters. For example, if $D_{ch} - D$ were very small, it was found that in the course of parametric resonance the cylinder would impact on the test-section, which rendered the theoretical assumptions invalid. If D_{ch} were large, on the other hand, the beneficial effect of flow convergence upstream of the test section would be eliminated, resulting in unacceptably high turbulence levels. Of course, some of these limitations of the apparatus could have been designed out of the system. However, the main purpose of these experiments was not to test extensively the effect of

parameters, but (*i*) to observe the behavior of the system, and (*ii*) for some arbitrary sets of parameters, to test the validity of the theoretical model. These aims may be considered to have been satisfied.

Conclusions

From the work presented in the two parts of this study a number of general conclusions may be drawn.

In the presence of harmonic perturbations to the mean flow velocity about a cylindrical body, parametric resonance oscillations may develop. Of these, the most important are associated with principal primary resonances. For the experiments of this study, the principal primary resonances associated with the second mode were predominant, although it cannot be concluded that this is a general result.

It was found that the theoretical model is capable of predicting the pulsation amplitudes and frequencies necessary to induce the observed parametric resonance oscillations. Accordingly, one may use this theoretical model with some confidence. It is especially important to note that the model agrees reasonably well with the experiments in relatively narrow annular passages, which together with the work of [4] indicates that the simple manner utilized by theory for accounting for the increased hydrodynamic mass is basically sound.

One important finding was that if the turubulence level in the flowing fluid is sufficiently large, parameteric resonances may be suppressed. An additional manifestation of this may be the observation that first-mode resonances did not materialize in these experiments. It may also contribute to the paucity of practical observations of occurrence of parametric resonances in the field.

Another reason for this last statement relates to the fact that one of the main diagnostic tools for attributing flow-induced vibration is the measurement of the cross-spectral-density between structural vibration and pressure in the flowing stream. As was shown here, however, the dominant parametric resonances are the principal ones, where the frequency of the pressure pulsations is twice that of vibrations. Hence, in the absence of coincidence between the two frequencies, this diagnostic technique would consistently miss parametric resonances, even if they are present. Of course, if such measurements are complemented by temporal cross-correlations between the two quantities, which is not always done, these resonances may then be detected.

Acknowledgments

The authors gratefully acknowledge the financial support given to this research by the Natural Sciences and Engineering Research Council of Canada and Le programme ≪Formation de chercheurs et action concertée≫ (D.G.E.S.-F.C.A.C.) of the ministère de l'Education du Québec, as well as by the Whiteshell Nuclear Research Establishment of Atomic Energy of Canada Ltd., at the early stages of this work; they are also indebted to Mr. Frank Shum for assistance with some of the experiments.

References

1 Paidoussis, M. P., Issid, N. T., and Tsui, M., "Parametric Resonance Oscillations of Flexible Slender Cylinders in Harmonically Perturbed Axial Flow. Part 1: Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, Dec. 1980, pp. 709–714.

2 Bishop, R. E. D., and Johnson, D. C., *The Mechanics of Vibration.* Cambridge University Press, 1960.

3 Paidoussis, M. P., and Issid, N. T., "Experiments on Parameteric Resonance of Pipes Containing Flow," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, 1976, pp. 198–202.

4 Paidoussis, M. P., and Pettigrew, M. J., "Dynamics of Flexible Cylinders in Axisymmetrically Confined Axial Flow," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 37–44.

Y. Matsuzaki
Principal Research Engineer.
Mem. ASME

T. Ueda
Research Engineer.

National Aerospace Laboratory,
Jindaiji, Chofu,
Tokyo, Japan

Reexamination of Unsteady Fluid Dynamic Forces on a Two-Dimensional Finite Plate at Small Mach Numbers

The Fourier transform theory is applied to the analytical determination of the disturbance velocity potential and pressure acting on a two-dimensional plate, in order to reexamine those of previous analyses by other investigators. Simplified expressions of the generalized forces are presented for incompressible and nearly incompressible flows. As the Mach number tends to zero, the virtual mass induced by an oscillating fluid becomes infinitely large for a natural mode symmetric with respect to the midchord point. It is recommended to take into account a symmetric mode which changes no fluid volume contained in a control surface, when a coupled flutter oscillation at very low Mach numbers is analyzed. Incorrectness in the generalized forces of the previous analyses is pointed out by comparing with the present analysis.

Introduction

Stability of a finite plate or cylinder exposed to subsonic flows has been examined by many investigators. Prerequisite to such a stability analysis is the determination of unsteady fluid pressures acting on the plate or cylinder. At an early stage of research on the plate stability the pressure approximation for an infinite wavy wall was used by Flax [1], and Dugundji, Dowell, and Perkin [2]. Making use of the source functions, Ishii [3, 4] derived expressions for the generalized forces of inviscid subsonic flow past a two-dimensional plate, which is connected to semi-infinite rigid walls at the leading and trailing edges. The pressure for a nearly incompressible flow past the same plate as Ishii treated was examined by Weaver and Unny [5] with the aid of the Fourier transform, and the generalized forces were numerically calculated in an approximate manner. Applying a Laplace transform with respect to time and a double Fourier transform to space variables, Dowell [6] evaluated the generalized forces for a rectangular plate undergoing a transient motion. Simple expressions of the two and three-dimensional generalized pressures in low frequencies are presented by Ellen [7], using the asymptotic expansions. Ellen claims that Weaver and Unny's approximation [5] overlooks an important contribution to the generalized forces, and that Ishii [3] has errors in his forces for the flow at finite Mach numbers.

As for an incompressible flow in a finite cylinder which is connected to semi-infinite rigid pipes, Weaver and Unny [8] numerically calculated the generalized forces by applying the Fourier transform theory like in their previous paper concerning the flat plate. Analytical evaluation of the generalized forces for the incompressible flow in the cylinder was presented by Matsuzaki and Fung [9] with the aid of the Fourier transform theory. Extending their approach, Matsuzaki and Fung [10, 11] evaluated analytically the generalized forces for incompressible and small Mach number flows in a two-dimensional, flexible channel which is connected to semi-infinite rigid channels.

When plates or cylinders are of finite length, difficulties are encountered in determining the pressure acting on them:

1 In an incompressible flow problem, the integral for pressure may become boundless and no pressure expression can be obtained. This happens, for instance, when the plate oscillates with an eigenmode symmetric with respect to the midchord point [5]. It is also true when the upper and lower panels of the two-dimensional channel oscillate symmetrically with respect to the center line of the channel as well as the midchord point [10]. We may find the same situation in a tube analysis [9] when the tube is assumed to oscillate in an axisymmetric mode, that is, the circumferential wave number $n = 0$. The analytic difficulty is considered to occur for a mode of oscillation which does not preserve constant fluid volume. In order to avoid the difficulty for such a mode, the effect of compressibility is taken into account in reference [11].

2 Tedious numerical calculations are required for evaluating the

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, September, 1979; final revision, April, 1980.

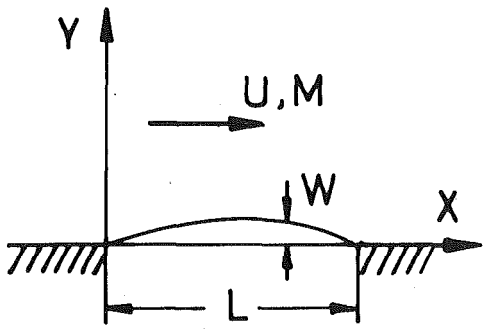


Fig. 1 The geometry of the panel and the coordinate system

generalized force at finite Mach numbers which is expressed in an integral form, for example, equation (8*b*) of reference [3].

In spite of the analytical problem mentioned in item 1, Ishii [3, 4] derived the generalized forces for the incompressible flow with an oscillatory mode symmetric with respect to the midchord point. If his analysis is carefully examined, then we see that the perturbation velocity potential, the pressure and the generalized forces, given by equations (2), (5*a*), and (8*a*) of reference [3], respectively, may diverge as x tends to infinity, although the counterparts for the compressible flow certainly attenuate. In other words, the radiation condition is not necessarily satisfied for the incompressible flow.

The objectives of the present paper are to reexamine the fluid pressures for subsonic flow past the same two-dimensional plate as used in references [3–5, 7], and, further, to derive simple expressions for the generalized forces at small Mach numbers by applying an asymptotic expansion. Details of derivation of the disturbance velocity potential and disturbance pressure shall be given in order to compare with the Ishii's and Weaver and Unny's analyses which are criticized by Ellen. No strong limitation on the reduced frequency is imposed unlike in reference [7], so that the generalized pressures proposed are applicable to analysis of dynamic instability.

Mathematical Formulation

An inviscid fluid with the free-stream velocity U flows past a two-dimensional panel in the X-direction as shown in Fig. 1. The panel is connected to semi-infinitely long, flat rigid walls at the leading and trailing edges, i.e., $X = 0$ and L. Let us assume that the space under the elastic panel and rigid walls ($Y < 0$) is empty.

The linearized velocity potential $\phi(X, Y, t)$ is described by

$$\partial^2\phi/\partial X^2 + \partial^2\phi/\partial Y^2 - (1/a_0^2)(\partial/\partial t + U\partial/\partial X)^2\phi = 0 \quad \text{for } Y \geq 0 \tag{1}$$

where a_0 is the velocity of sound in the fluid. The velocity potential is related to the surface deflection $w(X, t)$ by

$$(\partial\phi/\partial Y)_{Y=0} = \partial w/\partial t + U\partial w/\partial X \tag{2}$$

The potential must satisfy a radiation condition

$$\lim \phi(X, Y, t) = 0 \quad \text{as} \quad \sqrt{X^2 + Y^2} \to \infty \tag{3}$$

The perturbation pressure, $p(X, t)$, on the panel is given by

$$p = -\rho(\partial\phi/\partial t + U\partial\phi/\partial X)_{Y=0} \tag{4}$$

where ρ is the density of the free stream. We assume the plate deflection, the potential and the pressure to be in the forms of

$$w(X, t) = W(X)e^{i\omega t}, \quad \phi(X, Y, t) = \Phi(X, Y)e^{i\omega t}, \quad p(X, t) = P(X)e^{i\omega t} \tag{5}$$

Then, equations (1), (2), and (4) can be rewritten as

$$\partial^2\Phi/\partial x^2 + \partial^2\Phi/\partial y^2 - M^2(\partial/\partial x + ik)^2\Phi = 0 \tag{6}$$

$$(\partial\Phi/\partial y)_{y=0} = U(d/dx + ik)W \tag{7}$$

$$P(x) = -\{(\rho U/L)(\partial/\partial x + ik)\Phi\}_{y=0} \tag{8}$$

where

$$x = X/L, \quad y = Y/L, \quad k = \omega L/U, \quad M = U/a_0 \tag{9}$$

Applying the Fourier transformation, with respect to x, to Φ and W, i.e.,

$$\Phi^*(\xi, y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\Phi(x, y)e^{i\xi x}dx, \quad W^*(\xi) = \frac{1}{\sqrt{2\pi}}\int_0^1 W(x)e^{i\xi x}dx \tag{10}$$

we obtain from equations (6) and (7)

$$\partial^2\Phi^*/\partial y^2 - \zeta^2\Phi^* = 0 \tag{11}$$

$$(\partial\Phi^*/\partial y)_{y=0} = iU(-\xi + k)W^* \tag{12}$$

where

$$\zeta = \beta[(\xi - \xi_+)(\xi - \xi_-)]^{1/2}, \quad \xi_+ = Mk/(1 + M), \quad \xi_- = -Mk/(1 - M) \tag{13}$$

$$\beta = (1 - M^2)^{1/2} \tag{14}$$

The solution to equation (11) is given by

$$\Phi^*(\xi, y) = Ae^{\zeta y} + Be^{-\zeta y} \quad \text{for } y \geq 0 \tag{15}$$

When $\xi_- < \xi < \xi_+$, ζ is pure imaginary so that we can put $\zeta = i\bar{\zeta}$, where $\bar{\zeta}$ is real and positive. Since $\exp\{i(\omega t - \bar{\zeta}y)\}$ and $\exp\{i(\omega t + \bar{\zeta}y)\}$ represent disturbances moving in the positive and negative y-directions, respectively, the first term of equation (15) must be discarded. When $\xi < \xi_-$ or $\xi > \xi_+$, the first term diverges as y tends to infinity. Therefore, we must put

$$A = 0 \tag{16}$$

irrespective of the value of ξ. Equation (15), satisfying equation (12), is

$$\Phi^*(\xi, y) = \{iU(\xi - k)/\zeta\}W^* \exp(-\zeta y) \tag{17}$$

The inverse transform of equation (17) is written as

$$\Phi(x, y) = -(U/\sqrt{2\pi})\int_{-\infty}^{\infty}\{i(-\xi + k)W^*(\xi)/\zeta\}e^{-\zeta y}e^{-i\xi x}d\xi \tag{18}$$

Let us assume $Mk \neq 0$ and introduce a variable λ, defined by

$$\lambda = (\xi - \delta)/\gamma \tag{19}$$

where

$$\gamma = Mk/(1 - M^2) \quad \text{and} \quad \delta = -M^2k/(1 - M^2) \tag{20}$$

By virtue of the convolution theorem, equation (18) becomes

$$\Phi(x, y) = -(U/\sqrt{2\pi})\int_0^1 I(x - \eta, y)(d/d\eta + ik)W(\eta)d\eta \tag{21}$$

where

$$I(v, y) = (e^{-i\delta v}/\sqrt{2\pi}\beta)\int_{-\infty}^{\infty}\exp\{-(\gamma\beta\sqrt{\lambda^2 - 1}\,y + iv\gamma\lambda)\}/\sqrt{\lambda^2 - 1}\,d\lambda \tag{22}$$

Some manipulation on equation (22) yields

$$I(v, y) = -i\sqrt{\pi/2}\,e^{-i\delta v}H_0^{(2)}[\gamma\sqrt{\lambda^2 + (\beta y)^2}] \tag{23}$$

In deriving equation (23), we have used [12, 13]

$$\int_{-a}^{a}\cos(b\sqrt{a^2 - z^2})e^{-i\xi z}/\sqrt{a^2 - z^2}dz = \pi J_0(a\sqrt{\xi^2 + b^2}),$$

$$\int_{-a}^{a} \frac{\sin(b\sqrt{a^2-z^2})}{\sqrt{a^2-z^2}} e^{-i\xi z}dz + \left\{\int_{-\infty}^{-a} + \int_{a}^{\infty}\right\} \frac{\exp(-b\sqrt{z^2-a^2})}{\sqrt{z^2-a^2}} e^{-i\xi z}dz = \pi Y_0(a\sqrt{\xi^2+b^2})$$

$$H_0^{(2)}(z) = J_0(z) - iY_0(z)$$

J_0, Y_0, and $H_0^{(2)}$ are, respectively, the Bessel functions of the first, second, and third kinds. Substituting equation (23) into equation (21) gives

$$\Phi(x, y) = (iU/2\beta) \int_0^1 e^{-i\delta(x-\eta)} \times H_0^{(2)}[\gamma\sqrt{(x-\eta)^2+(y\beta)^2}] \left(\frac{d}{d\eta} + ik\right) W(\eta)d\eta \quad (24a)$$

$$= (iU/2\beta) \int_0^1 W(\eta)e^{-i\delta(x-\eta)}[i(k-\delta)H_0^{(2)}[\gamma\sqrt{(x-\eta)^2+(y\beta)^2}] - \gamma \operatorname{sign}(x-\eta)H_1^{(2)}[\gamma\sqrt{(x-\eta)^2+(y\beta)^2}]]d\eta \quad (24b)$$

where $W(0) = W(1) = 0$ have been used.

The disturbance velocity potential, equation (24a), is the same as equation (3)[1] of reference [3], where Ishii derived it by using the source function. The disturbance potential at any point (x, y) can be determined by using equations (24).[2] It is evident from equation (24b) that $\Phi(x, y)$ vanishes as $\sqrt{x^2+y^2}$ tends to infinity as long as the fluid is compressible, i.e., $M \neq 0$.

Low Mach Number Approximation. Now, let us assume that

$$M \ll 1 \quad (25)$$

and that the order of k is of unity at most, i.e.,

$$0(k) \lesssim 1 \quad (26)$$

If we confine ourselves to evaluating a distribution of the velocity potential on the plate $(0 < x < 1)$, then it follows that

$$\gamma|x-\eta| \ll 1 \quad \text{and} \quad \delta(x-\eta) \ll 1 \quad (27)$$

for $0 < \eta < 1$. From the definitions of the Bessel functions [13], we obtain, for small z,

$$J_0(z) = 1 - (z/2)^2 + 0(z^4)$$

$$Y_0(z) = (2/\pi)[J_0(z)\{\gamma^* + \ln(z/2)\} + (z/2)^2 + 0(z^4)] \quad (28)$$

where γ^* is Euler's constant. Upon taking into account equations (25)–(28), equations (24) are reduced, for $0 < x < 1$, to

$$\Phi(x, 0) = (U/\pi) \int_0^1 (\ln|x-\eta| + c)(d/d\eta + ik)W(\eta)d\eta \quad (29)$$

where

$$c = \ln(Mk/2) + \gamma^* + i\pi/2 \quad (30)$$

c may have a large negative value in its real part since $M \ll 1$ and $0(k) \lesssim 1$. Integrating by part equation (29) yields

$$\Phi(x, 0) = \frac{U}{\pi}\left[\int_0^1 \left\{ik \ln|x-\eta| + \frac{1}{x-\eta}\right\} \times W(\eta)d\eta + ikc \int_0^1 W(\eta)d\eta\right] \quad (31)$$

It is noted that the last integral, which is multiplied by c, is independent of x.

Substituting equation (31) into equation (8), we have

$$P(x) = \frac{\rho U^2}{\pi L}\left[\int_0^1 \left\{\frac{1}{(x-\eta)^2} - \frac{2ik}{x-\eta} - (ik)^2 \ln|x-\eta|\right\} W(\eta)d\eta - c(ik)^2 \int_0^1 W(\eta)d\eta\right] \quad (32)$$

Now, letting

$$W(x) = \sum_{q=1}^{N} a_q W_q(x) \quad (33)$$

we will evaluate the generalized fluid forces defined by

$$Q_{mn} = L \int_0^1 P_m(x)W_n(x)dx/(\rho U^2/2) \quad (34)$$

where $P_m(x)$ is the pressure due to the deflection $W_m(x)$. When the panel is simply supported at both ends, we put

$$W_m(x) = \sin m\pi x \quad (35)$$

Substituting equations (32) and (35) into equation (34) gives

$$Q_{mn} = -Q_{mn}^{(0)} + (ik)^2 Q_{mn}^{(2)}, \quad \text{for} \quad m+n = \text{even} \quad (36a)$$

$$Q_{mn} = ikQ_{mn}^{(1)}, \quad \text{for} \quad m+n = \text{odd} \quad (36b)$$

where

$$Q_{nn}^{(0)} = 2n[\operatorname{Si}(n\pi) - \{1-(-1)^n\}/n\pi] \quad (37a)$$

$$Q_{mn}^{(0)} = \frac{4mn}{\pi(m^2-n^2)}\{\operatorname{Ci}(m\pi) - \operatorname{Ci}(n\pi) + \ln(n/m)\}, \quad \text{for } m \neq n \quad (37b)$$

$$Q_{mn}^{(1)} = (8/\pi^2)\{n \operatorname{Si}(m\pi) + m \operatorname{Si}(n\pi)\}/(n^2-m^2) \quad (37c)$$

$$Q_{mn}^{(2)} = 4[n^2\{\operatorname{Ci}(m\pi) - \ln m\pi - \gamma^*\} - m^2\{\operatorname{Ci}(n\pi) - \ln n\pi - \gamma^*\}]/\{mn(m^2-n^2)\pi^3\} - (c/mn\pi^2)\{1-(-1)^m\}\{1-(-1)^n\} \quad (37d)$$

Si and Ci are the sine and cosine integral functions, respectively. The Mach number disappears in the coefficients of the generalized forces, except for $Q_{mn}^{(2)}$ where m and n are odd. Because of c in the last term of equation (37d), the real part of $Q_{mn}^{(2)}$ is large for a nearly incompressible flow when both m and n are odd numbers. In other words, when the plate is oscillating with one of the assumed modes symmetric with respect to $x = 0.5$, a large virtual mass is induced. Consequently, no rapid oscillation in such a mode is expected. If M tends to zero, then we may predict from equations (37d) and (30) that an infinite amount of virtual mass induced will prevent the plate from oscillating in the symmetric mode.

However, it is clear from equations (31) and (32) that a rapid oscillation in symmetric modes may occur, provided that

$$\int_0^1 W(x)dx = 0, \quad (38)$$

since the last term in equation (37d) disappears. Let us consider an arbitrary control surface enclosing the elastic panel. Then, we see that equation (38) represents no change in fluid volume contained in the control volume. The simplest symmetric mode which satisfies equation (38) is given by

$$W(x) = (1/\sqrt{10})(\sin \pi x - 3 \sin 3\pi x) \quad (39)$$

Next, we shall analytically evaluate the pressure for the incompressible flow in order to compare with the analyses by Ishii, and Weaver, and Unny.

Incompressible Flow Approximation ($M = 0$). When $M = 0$, we must return to equation (18). Because of $\xi_+ = \xi_- = 0$ and $\zeta = |\xi|$, $\Phi(x, 0)$ is written as

$$\Phi(x, 0) = -(U/\sqrt{2\pi}) \int_{-\infty}^{\infty} J(x-\eta)(d/d\eta + ik)W(\eta)d\eta \quad (40)$$

where

$$J(v) = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} e^{-iv\xi}/|\xi|d\xi \quad (41)$$

[1] i is dropped out from the right-hand side of equation (3). There are also a few errors in equation (5b) of reference [3].

[2] Whenever improper integrals appear, their principal values should be taken.

Table 1 Generalized forces [a] Q_{mn} for $M = 0$, associated with natural modes

n		m = 1	m = 2	m = 3
1	b	$-2.4304 + \infty\ (ik)^2$	$-1.3838\ ik$	$0.55461 + \infty\ (ik)^2$
	c	$-2.4247 + 0.45296\ (ik)^2$	$-1.3824\ ik$	$0.57312 + 0.06265\ (ik)^2$
2	b	d	$-5.6720 + 0.22230\ (ik)^2$	$-1.2327\ ik$
	c		$-5.6753 + 0.24173\ (ik)^2$	$-1.2320\ ik$
3	b			$-8.7759 + \infty\ (ik)^2$
	c			$-8.7755 + 0.13904\ (ik)^2$

a) $Q_{mn} = -Q_{mn}^{(0)} + ik\,Q_{mn}^{(1)} + (ik)^2\,Q_{mn}^{(2)}$

b) present analysis

c) from reference 4

d) $Q_{nm} = (-1)^{m+n} Q_{mn}$

The inverse Fourier integral, equation (41), cannot be evaluated in an ordinary manner [14]. However, if we resort to the theory of generalized functions [15], then equation (41) becomes

$$J(v) = (-2/\sqrt{2\pi})(\ln|v| + d) \tag{42}$$

where d is an arbitrary constant. Substitution of equation (42) into equation (40) yields

$$\Phi(x, 0) = (U/\pi)\int_0^1 (\ln|x - \eta| + d)(d/d\eta + ik)W(\eta)d\eta \tag{43}$$

Comparison of equation (43) with equation (29) shows that the disturbance velocity potential for $M = 0$ has the same form as that for $M \ll 1$. It should be noted, however, that equation (29) is applicable only to the small range of x about the origin, whereas equation (43) is valid for any value of x. We will make use of equations (31)–(39), instead of deriving the corresponding equations, by assuming that c is an arbitrary constant in place of d and eliminating the restriction on x.

When $|x| \gg 1$, equation (31) is reduced to

$$\Phi(x, 0) = (ikU/\pi)(\ln|x| + c)\int_0^1 W(\eta)d\eta \tag{44}$$

Let us assume $k \neq 0$. As long as equation (38) holds, the radiation condition is satisfied. If this is not the case, that is, if the fluid volume in the control surface is not constant, then the radiation condition is violated. In other words, if the deflection of the panel does not satisfy equation (38), then no oscillation in such a mode can occur because of an infinitely large virtual mass induced. This agrees with the remark concerning $M \to 0$ given in the preceding section.

As a matter of course, it follows that the generalized forces for the incompressible flow are presented by equations (36) and (37) with $M = 0$.

Numerical Results and Discussions

Let us begin with comparing Weaver and Unny's analysis [5] with ours. There are two points to be discussed. First, they introduced an assumption of a nearly incompressible flow to avoid the boundlessness of the pressure integral for $M = 0$. As is seen in equation (41), we have encountered the same problem. However, this has been solved by using the theory of generalized functions. Second, by arguing that the velocity potential does not attenuate along the z-direction for $0 < u < \omega/a$, Weaver and Unny changed the lower limit of equation (18) of reference [5] from 0 to ω/a. Let us return to equation (22) of the present paper. One may find the same situation there. If one follows Weaver and Unny, then the integration over $|\lambda| < 1$ in equation (22) had to be excluded. However, it is obvious from equations (23) and (21) that $I(v, y)$ and, therefore, $\Phi(x, y)$ are diminishing as y tends to infinity although the integration is included. As pointed out by Ellen [7], Weaver and Unny's velocity potential has a zero-order error.

In reference [7] Ellen asserts that Ishii has errors in the generalized forces for flow at finite Mach numbers, being suspicious of the calculations to be based on the result for $M = 0$. However, we can consider that Ishii's numerical calculations for finite Mach numbers are independent of the result for $M = 0$, since the generalized forces were calculated by approximating the Hankel functions which appear in equation (8*b*) of reference [3] or equation (14) of reference [4].

As for Ishii's analysis for the incompressible flow, it should be noted that his velocity potential, i.e., equation (2) of reference [3] with $z = 0$ differs from equation (43) of the present paper by the constant d. As examined in the previous section, if equation (38) does not hold, then d must be set to a negatively infinite number in equation (43). Otherwise, the radiation condition is dissatisfied. Hence, the equation (2) can not satisfy the condition by any means.

The generalized forces defined by equations (36) are numerically evaluated for the first to third eigenmodes as well as the symmetric mode described by equation (39). In Table 1 the results for the incompressible flow are presented in comparison with those of reference [4]. When m and n are odd numbers, that is, equation (38) does not hold, $Q_{mn}{}^{(2)}$ calculated by the present analysis are infinite, whereas those of reference [4] are finite. As previously mentioned, the radiation condition is violated in the latter. It is seen that the generalized force coefficients of the present analysis agree with those of Ishii, except for $Q_{mn}{}^{(2)}$ (m, n: odd). Table 2 shows $Q_{mn}{}^{(2)}$ (m, n: odd) at several small Mach numbers with $k = 1.0$. They are complex numbers. Their real parts grow slowly as M decreases, and they become infinite at $M = 0$.

The generalized forces associated with the symmetric mode described by equation (39) are given in Table 3. It is noted that $Q_{ss}{}^{(2)}$ is considerably small, compared with the real part of $Q_{11}{}^{(2)}$ or $Q_{33}{}^{(2)}$ in Table 2. We may expect that the plate will oscillate more rapidly in the mode defined by equation (39) than either in the first or third eigenmode, since the virtual mass is smaller in the former than in the latter. Therefore, it is plausible that, instead of the first or third mode, such a symmetric mode producing a small virtual mass is taken into

Table 2 $Q_{mn}^{(2)}$ **at small Mach numbers with** $k = 0$ **for** m **and** n **being odd numbers**

	Real Part				Imag. Part [a]
M	0.01	0.001	0.00001	0	
$Q_{11}^{(2)}$	1.6770	2.2711	3.4593	∞	-0.40529
$Q_{31}^{(2)}$ [b]	0.47067	0.66870	1.0648	∞	-0.13510
$Q_{33}^{(2)}$	0.27442	0.34043	0.47245	∞	-0.04503

a) Imaginary parts are independent of M.

b) $Q_{13}^{(2)} = Q_{31}^{(2)}$

Table 3 **Generalized forces for the symmetric mode defined by equation (39)**

Q_{s1} [a]	$-1.2947 + 0.083795\,(ik)^2$
Q_{s2}	$1.6070\,ik$
Q_{ss}	$-8.4741 + 0.13228\,(ik)^2$

a) $Q_{ms} = (-1)^{m+1} Q_{sm}$

account together with the second eigenmode in analyzing coupled flutter oscillations. Stability analysis of a flat or buckled plate will be presented elsewhere by using the generalized forces obtained here.

Concluding Remarks

We have examined previous analyses concerning a subsonic flow past a simply supported two-dimensional plate. The Fourier transform theory is applied in obtaining analytically the disturbance velocity potential, pressure, and generalized forces to compare with those derived by Ishii who used the source functions. The generalized forces for flow at small Mach numbers are presented in simplified forms by using the asymptotic expansion. Comparison with Ishii's analysis shows that he has errors in the case of the incompressible flow, since the radiation condition is violated when the fluid volume in the control surface changes. Consequently, we may conclude that it is necessary to reexamine the stability analysis of simply supported plates exposed to an incompressible flow. We confirm Ellen's claim that Weaver and Unny have zero-order errors in their generalized forces of a nearly incompressible flow.

References

1 Flax, A. H., *Aero and Hydroelasticity, Structural Mechanics, Proceedings of the First Symposium on Naval Structural Mechanics,* Pergamon Press, New York, 1960.

2 Dugundji, J., Dowell, E., and Perkin, B., "Subsonic Flutter of Panels on Continuous Elastic Foundations," *AIAA Journal,* Vol. 1, No. 5, 1963, pp. 1146–1154.

3 Ishii, T., "Aeroelastic Instabilities of Simply Supported Panels in Subsonic Flow," AIAA Paper No. 65-772, 1965.

4 Ishii, T., "Theoretical Study of Two-Dimensional Panel Flutter and Panel Divergence in Subsonic Flow, (I) Incompressible Case and (II) Compressible Flow Case," Technical Reports TR-83 and TR-87, National Aerospace Laboratory, Tokyo, 1965.

5 Weaver, D. S., and Unny, T. E., "The Hydroelastic Stability of a Flat Plate," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, 1970, pp. 823–827.

6 Dowell, E. H., "Generalized Aerodynamic Forces on a Flexible Plate Undergoing Transient Motion," *Quarterly of Applied Mathematics,* Vol. 24, No. 4, 1967, pp. 331–338.

7 Ellen, C. H., "The Stability of Simply Supported Rectangular Surfaces in Uniform Subsonic Flow," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, 1973, pp. 68–72.

8 Weaver, D. S., and Unny, T. E., "On the Dynamic Stability of Fluid-Conveying Pipes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, 1973, pp. 48–52.

9 Matsuzaki, Y., and Fung, Y. C., "Unsteady Fluid Dynamic Forces on a Simply Supported Circular Cylinder of Finite Length Conveying a Flow, With Applications to Stability Analysis," *Journal of Sound and Vibration,* Vol. 54, No. 3, 1977, pp. 317–330.

10 Matsuzaki, Y., and Fung, Y. C., "Stability Analysis of Straight and Buckled Two-Dimensional Channels Conveying an Incompressible Flow," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 548–552.

11 Matsuzaki, Y., and Fung, Y. C., "Nonlinear Stability Analysis of a Two-Dimensional Model of an Elastic Tube Conveying a Compressible Flow," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 31–36.

12 Moriguchi, S., Udagawa, K., and Hitotsumatsu, S., *Mathematical Formula II,* Iwanami Shoten, Tokyo, 1956.

13 Abramowitz, M., and Stegun, I. A., ed., *Handbook of Mathematical Functions,* Dover, New York, 1965.

14 Sneddon, I. N., *Fourier Transforms,* McGraw-Hill, New York, 1951.

15 Lighthill, M. J., *Fourier Analysis and Generalized Functions,* Cambridge University Press, London, 1958.

A. A. Sonin

Professor,
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
Cambridge, Mass. 02139
Mem. ASME

Rationale for a Linear Perturbation Method for the Flow Field Induced by Fluid-Structure Interactions[1]

A formal justification is developed for a method in which hydrodynamic data for a transient in a rigid-wall system (derived, for example, from a small-scale experimental simulation) is used as input in a linear computation for the perturbation flow field due to actual wall flexibility. The method is useful in problems where the basic flow transient is so complex that it can be quantified only empirically, and where the fluid-structure interaction is too complex for the fluid side to be represented by a priori defined equivalent mass.

Introduction

The analysis of loads resulting from complex flow transients in vessels is often further complicated by the effects of elastic boundaries. Numerical methods are almost invariably required, and even so, only relatively simple problems can be solved practicably [1]. The purpose of this paper is to identify a class of such problems where it is useful to separate the flow field into a component which would result if the walls were perfectly rigid, and a perturbation which arises because of wall flexibility. We will show rigorously that the effects of the wall flexibility can be derived separately by means of a perturbation analysis which in most cases is considerably simpler than the general problem. The pressure of the rigid-wall flow field appears as a forcing function at the boundary of the perturbation flow field.

This result is useful in two ways. First, it simplifies analysis. The calculation for the flow transient with assumed rigid boundaries can be done first and the additional effects of wall flexure can be derived by a separate perturbation calculation in which the fluid behaves linearly.

The second utility of our result arises in cases where the flow transient is so complex that a computation for it, or for its rigid-wall component, is difficult or impossible. In such cases, the first calculation can be replaced by experimental data from a small-scale simulation using rigid walls. Our analysis provides a formal justification for a method of using such data as input in a relatively simple calculation for the perturbation caused by wall flexibility. Such a combined empirical/analytical approach is often more practicable than a complete empirical simulation which includes both the effects of wall flexure and the fluid dynamic transient proper.

The technique of applying the experimental rigid-system load as a forcing function to compute structural oscillations has been used widely to solve problems involving flow-induced vibrations of cylinders and similar structures [2]. In these applications, the question of how one deals with the inertially induced pressure field in the fluid is resolved simply by introducing an equivalent mass, one which can be determined semiempirically for a given body geometry. The method we propose here is useful in more complex problems where an equivalent water mass cannot be specified *a priori*, and where a solution must be derived for the flow field perturbation which results from wall flexibility.

The method suggested here is not novel. Bedrosian [3], for example, has applied essentially this method to compute the fluid-structure interaction effects in pressure-suppression containment vessels of boiling-water reactors. The purpose of the present paper is to give the method a formal basis, and to specify the conditions which must be satisfied if it is to be valid.

Analysis

We consider a class of problems where an essentially inviscid motion is induced in a liquid by the transient application of pressure at one or several of the places where the liquid is bounded by gas. Elsewhere, the liquid is bounded by solid, but flexible, walls. The general case is best illustrated with an example (Fig. 1).

A vertical pipe is partially submerged in a liquid pool which is initially at rest, and bounded above by a region of gas. An event is triggered by a sudden discharge of gas or vapor into the pipe from above, causing the clearing of the liquid from the pipe, the formation of a gas bubble at the pipe end, and the rise or oscillation of the liquid in the pool. If the pool boundaries are rigid, the resulting pressure history at some point on the pool floor, for example, might be the one sketched in Fig. 2. If the boundaries are elastic, they, and together with

[1] Work supported by the U. S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Division of Reactor Safety Research, under Contract No. NRC-07-77-011.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, August, 1979; final revision, May, 1980.

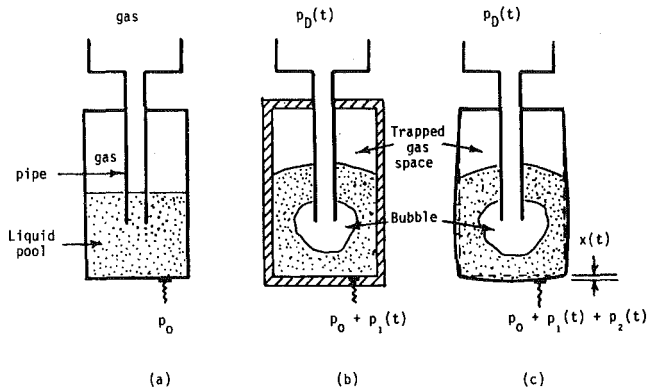


Fig. 1 Example of flow transient. (*a*) Initial condition, (*b*) time *t* in system with rigid walls; (*c*) time *t* in system with flexible walls

them the pool, would be set into oscillation, and the resulting pool acceleration and deceleration would give rise to an additional oscillatory component of pressure, as indicated on the figure. We aim to show how these two contributions to the pressure can be separated.

The liquid dynamics is governed by the equation of motion,

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p - \rho g \nabla z, \tag{1}$$

the equation of mass conservation,

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0 \tag{2}$$

and the isentropic pressure-density relation, $dp/d\rho = c^2$. We shall assume that the latter applies in the linear approximation,

$$p - p_0 = c_0^2(\rho - \rho_0) \tag{3}$$

where p_0, c_0, and ρ_0 are the pressure, speed of sound, and density, respectively, in the undisturbed fluid.

The boundary conditions must be specified for equations (1)–(3) at the free surfaces and at the solid walls. We assume that the pressure at any free surface is uniform and, for the purposes of the analysis of the liquid, given. In Fig. 1, the trapped gaseous space above the liquid is one free surface, and the bubble emerging from the pipe is another. The free-surface boundary condition can thus be express as

$$p = p_i(t) \qquad \text{at the } i\text{th free surface.} \tag{4}$$

At the solid walls one must apply a boundary condition like

$$v_\perp = \frac{dx}{dt} \tag{5}$$

where $v_\perp$ is the fluid velocity component directed perpendicularly into the wall, and x is the displacement of the wall (away from the fluid) from its initial, equilibrium position under hydrostatic conditions. The wall displacement x is governed by a structural equation of motion which can be expressed symbolically as

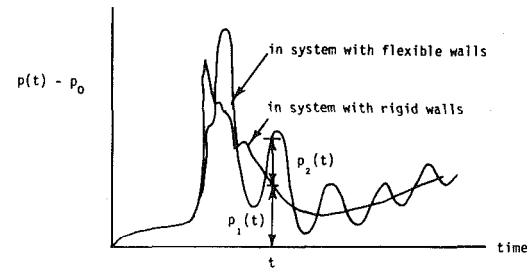


Fig. 2 Sketch of pressure histories in rigid-wall and in flexible-wall systems

$$m\frac{d^2x}{dt^2} = (p - p_0) - \sigma\left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, t\right) \tag{6}$$

where m is the local mass of the wall per unit area, p is the local instantaneous pressure in the fluid, p_0 is the initial hydrostatic pressure, and σ is a local structural restraining force per unit area, whose magnitude depends on the displacement x of the wall from its initial equilibrium position, on the time derivatives of x, and possibly also on the time t itself. The form of σ is governed by structural considerations. Note that when the displacements x are small, the boundary condition embodied in equations (5) and (6) can to a good approximation be applied at the equilibrium, or undisturbed, location of the wall rather than at the actual, instantaneous deflected position.

We separate the variables into three components by writing

$$\mathbf{v} = 0 + \mathbf{v}_1(\mathbf{r}, t) + \mathbf{v}_2(\mathbf{r}, t) \tag{7}$$

$$p = p_0(\mathbf{r}) + p_1(\mathbf{r}, t) + p_2(\mathbf{r}, t) \tag{8}$$

$$\rho = \rho_0 + \rho_1(\mathbf{r}, t) + \rho_2(\mathbf{r}, t) \tag{9}$$

where the subscript 0 refers to the values corresponding to the initial static conditions in the fluid, the subscript 1 refers to the hypothetical perturbation which *would* be caused if the imposed blowdown occurred in the system with *rigid walls*, and the subscript 2 refers to the remainder of the quantity, and represents the perturbation which can be attributed to the flexibility of the walls. The initial pressure distribution p_0 is assumed to hydrostatic,

$$p_0 = \text{constant} - \rho_0 g z. \tag{10}$$

By definition, the rigid-wall flow is the solution of equations (1)–(3) with $\mathbf{v}_2$, p_2, and ρ_2 equal to zero. Thus

$$(\rho_0 + \rho_1)\left(\frac{\partial \mathbf{v}_1}{\partial t} + \mathbf{v}_1\cdot\nabla\mathbf{v}_1\right) = -\nabla p_1 - \rho_1 g \nabla z \tag{11}$$

$$\frac{\partial \rho_1}{\partial t} + \mathbf{v}_1\cdot\nabla\rho_1 + (\rho_0 + \rho_1)\nabla\cdot\mathbf{v}_1 = 0 \tag{12}$$

$$p_1 = \rho_1 c_0^2 \tag{13}$$

where we used equation (10) to eliminate p_0, and assumed ρ_0 to be constant. The boundary conditions for the rigid-wall solution are that

$$p_0 + p_1 = p_i(t) \quad \text{at the } i\text{th free surface} \tag{14}$$

Nomenclature

c = speed of sound in liquid
g = gravitational acceleration
L = characteristic length associated with gradients in velocity, density, and pressure
m = wall mass per unit area
p = pressure
t = time
$\mathbf{v}$ = velocity
$v_\perp$ = component of velocity directed perpendicularly into wall
x = displacement of wall from equilibrium position, in perpendicular direction away from fluid
z = direction measured vertically upward, against gravity
ρ = liquid density
τ_2 = characteristic time associated with the wall flexure

Subscripts

0 = value corresponding to the initial conditions in the static fluid

1 = perturbation which would be caused if the event occurred in a rigid-wall system

2 = remainder of the quantity, i.e., additional perturbation caused by wall flexibility

and that

$$(v_\perp) \equiv 0 \quad \text{at solid walls.} \tag{15}$$

The equations for the perturbation (2) due to wall flexure is obtained by substituting equations (7)–(10) into equations (1) and (2), and subtracting equations (11) and (12), respectively. One obtains the equation of motion

$$(\rho_0 + \rho_1 + \rho_2)\left(\frac{\partial \mathbf{v}_2}{\partial t} + \mathbf{v}_2 \cdot \nabla \mathbf{v}_1 + \mathbf{v}_1 \cdot \nabla \mathbf{v}_2 + \mathbf{v}_2 \cdot \nabla \mathbf{v}_2\right) = -\nabla p_2 - \rho_2 g \nabla z - \rho_2\left(\frac{\partial \mathbf{v}_1}{\partial t} + \mathbf{v}_1 \cdot \nabla \mathbf{v}_1\right) \tag{16}$$

and the mass conservation equation

$$\frac{\partial \rho_2}{\partial t} + \mathbf{v}_1 \cdot \nabla \rho_2 + \mathbf{v}_2 \cdot \nabla \rho_1 + \mathbf{v}_2 \cdot \nabla \rho_2 + \rho_2 \nabla \cdot \mathbf{v}_1 + (\rho_0 + \rho_1 + \rho_2)\nabla \cdot \mathbf{v}_2 = 0, \tag{17}$$

and the pressure-density relation,

$$p_2 = \rho_2 c_0^2. \tag{18}$$

Assuming that the wall flexure does not actually affect the gas pressure at the free surfaces, we can write the free-surface boundary condition as

$$p_2 = 0 \quad \text{at the } i\text{th free surface.} \tag{19}$$

The boundary condition at the solid walls is

$$(v_\perp)_2 = \frac{dx}{dt} \tag{20}$$

where $x(t)$ is given by equation six.

Equations (16) and (17) can be simplified considerably under conditions which are often not very restrictive in practice. Let

p_1 = typical amplitude of p_1
p_2 = typical amplitude of p_2
L = characteristic length over which gradients in velocity and pressure occur during the transient
τ_2 = characteristic time of the oscillation caused by wall flexure
x = typical wall displacement during wall flexure

We assume that

$$\frac{p_1}{\rho_0 c_0^2} \ll 1 \tag{21}$$

$$\frac{p_2}{\rho_0 c_0^2} \ll 1 \tag{22}$$

$$\frac{gL}{c_0^2} \ll 1 \tag{23}$$

$$\frac{\tau_2}{L}\sqrt{\frac{p_1}{\rho_0}} \ll 1 \tag{24}$$

$$\frac{x}{L} \ll 1. \tag{25}$$

The implication of these assumptions becomes apparent when we analyze the orders of magnitude of the various terms in equations (11) and (12) and equations (16) and (17).

We assume that the velocity v_1 arises from the acceleration of the liquid by a pressure difference p_1 acting over a distance L. The time of the acceleration is of order L/v_1. The order of magnitude of v_1 can then be estimated from the equation of motion as

$$v_1 \sim \sqrt{\frac{p_1}{\rho_0}} \tag{26}$$

The remaining order-of-magnitude estimates are self-evident:

$$\nabla \sim \frac{1}{L} \tag{27}$$

$$v_2 \sim \frac{x}{\tau_2} \tag{28}$$

$$\frac{\partial v_2}{\partial t} \sim \frac{x}{\tau_2^2} \tag{29}$$

$$\frac{\partial \rho_2}{\partial t} \sim \frac{\rho_2}{\tau_2} \tag{30}$$

In addition, equations (13) and (18) give ρ_1 and ρ_2 in terms of p_1 and p_2, respectively.

Using these estimates, we are in a position to estimate the relative orders of magnitude of the various terms in the governing equation. We find that equations (21) and (22) imply, first of all, that ρ_1 and ρ_2 are small compared with ρ_0. Equation (21) also implies that the second term in equation (12) is small compared with the third. Equation (23) is equivalent to the assumption that the gravitational term in equation (11) is negligible compared with the pressure gradient term. Hence, the equations for the rigid-wall flow reduce in good approximation to

$$\rho_0\left(\frac{\partial \mathbf{v}_1}{\partial t} + \mathbf{v}_1 \cdot \nabla \mathbf{v}_1\right) = -\nabla p_1 \tag{31}$$

$$\frac{\partial \rho_1}{\partial t} + \rho_0 \nabla \cdot \mathbf{v}_1 = 0 \tag{32}$$

$$p_1 = \rho_1 c_0^2 \tag{33}$$

The boundary conditions for the rigid-wall flow are given by equations (14) and (15).

On the left-hand side of equation (16), the second and third velocity terms are negligible compared with the time-derivative term when equation (24) applies, and the fourth term is negligible when equation (25) applies. On the right-hand side of the same equation, the pressure-gradient term is large compared with the terms involving $\mathbf{v}_1$ when equation (21) applies, and also large compared with the gravitational term if equation (23) applies. In equation (17), the second and fifth terms on the left are small compared with the first when equation (24) applies. The third is small compared with the last one (the sixth) if equation (21) applies, and the fourth is small compared with the last one if equation (22) applies. Thus the equations for the perturbation field caused by wall flexure reduce to

$$\rho_0 \frac{\partial \mathbf{v}_2}{\partial t} = -\nabla p_2 \tag{34}$$

$$\frac{\partial \rho_2}{\partial t} + \rho_0 \nabla \cdot \mathbf{v}_2 = 0 \tag{35}$$

$$p_2 = c_0^2 \rho_2 \tag{36}$$

The boundary conditions for this perturbation field are equations (19) and (20). Consistent with the small-perturbation assumption, the free-surface boundary condition, equation (19), is to be applied at the free-surface location obtained from the rigid-wall solution (or experiment), and the wall boundary condition, equation (20), is to be applied at the initial, undisturbed wall location (which is the wall location in the rigid-wall solution). The displacement $x(t)$ which appears in the boundary condition at the wall (equation (20)) couples the perturbation field to the structural behavior of the wall. In the symbolic representation of equation (6), $x(t)$ is governed by the equation

$$m\frac{d^2x}{dt} = p_1 + p_2 - \sigma\left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, t\right) \tag{37}$$

Note that the equations (34)–(36) for the perturbation field do not themselves explicitly involve the rigid-wall solution. The perturbation field due to wall flexure is coupled to the rigid-wall flow field only through the pressure p_1 which appears in the structural dynamic equation of the wall (e.g., equation (37)) and through the instantaneous locations and shapes of the free surfaces (governed by the rigid-wall solution), where the boundary condition equation (19) must be applied. It is as if the rigid-wall pressure $p_1(t)$ appears as an externally applied transient pressure on the wall, and drives the wall

(and fluid) oscillation calculated by the perturbation (2). Thus, if one has obtained, analytically or experimentally, the rigid-wall pressure distribution history p_1 at the walls and the time dependent shapes of the free surfaces, one can apply the boundary conditions on the solution (2) and calculate the perturbations in velocity and pressure, throughout the fluid, caused by wall flexibility.

Equations (34)–(36) are the linearized acoustic equations for the liquid, and can be solved, for example, by the usual linear method of characteristics with the sound speed c_0 taken as a constant. The liquid is, as it were, regarded somewhat as a "gel" with its free-surface boundaries prescribed as a function of time by the rigid-wall flow field. In the particular case where the period τ_2 of the wall oscillations is much longer than the acoustic transit time l/c_0 across the system, l being the characteristic length of the liquid pool,

$$\frac{l}{\tau_2 c_0} \ll 1, \tag{38}$$

equations (34)–(36) reduce to the simple, linearized incompressible flow form

$$\rho_0 \frac{\partial \mathbf{v}_2}{\partial t} = -\nabla p_2 \tag{39}$$

$$\nabla \cdot \mathbf{v}_2 = 0 \tag{40}$$

This constitutes a particularly simple case, since the pressure p_2 now satisfies Laplace's equation.

Discussion of Applications

The analysis outlined here has an important application in the design of large vessels, such as the pressure-suppression containment systems of boiling-water reactors, where transient liquid flows are induced and wall flexibility effects must be accounted for. In many such cases the basic transient flow phenomena are very complex and can be quantified only by means of small-scale experimental simulation [4, 5]. Often it is impracticable to model both the effects of wall flexibility and the hydrodynamics on a small scale. The question then arises whether one can use the hydrodynamic data derived from small-scale tests with rigid walls, and derive from that the strains in the walls of a real, full-scale system where wall flexure may occur.

The present analysis gives a rigorous basis for such a procedure. Consider the situation in Fig. 1 as an example. Let us say we have available, from an experimental simulation with *rigid* walls, the pressure history $p_1(t)$ at every point on the walls and the locations of the free surfaces as functions of time. One can then obtain a numerical solution of the relatively simple, linear fluid equations, equations (34)–(36) (or equations (39)–(40), if equation (38) applies) for the perturbations caused by wall flexibility. The rigid-wall pressure $p_1(t)$ appears as a driving force in the wall boundary condition for the perturbation field (equations (20) and (37)), applied at the initial undisturbed wall location, and the specified free-surface locations define where the free-surface boundary condition, equation (19), is to be applied.

The method is valid as long as the inequalities expressed in equations (21)–(25) apply. The requirements expressed by equations (21)–(23) are satisfied in most practical cases. In water, for example, it suffices that p_1 and p_2 be small compared with 2×10^4 bar, and that the characteristic length L of the flow transient be small compared with 2×10^2 km. Equation (25) is also satisfied in many practical cases: it merely requires that the wall deflections x be small in amplitude compared with L.

The key requirement is equation (24). The characteristic length L is the length over which gradients in velocity and pressure occur in the fluid during the transient. Usually, this can be taken as the system size. The depth of the pressure suppression pool in a boiling-water reactor, for example, is about 5 m. With $L = 5$ m and $p_1 \sim 1$ bar, equation (24) would be satisfied if the wall oscillation frequency $(2\pi\tau_2)^{-1}$ is large compared with 0.3 Hz. A more conservative interpretation of the requirement would set L equal to the smallest characteristic flow length in the system. This would be of the order of 1 m in the example, and thus the conservative requirement in the example would be that the wall oscillation frequency should be large compared with 1.6 Hz. One notes that linearization schemes often work well to relatively large amplitudes, and hence the wall oscillation frequencies may not have to be very much larger than the limiting values computed from equation (24) for the method to be a good approximate.

References

1 Belytschko, T., *Nuclear Engineering and Design,* Vol. 42, 1977, pp. 41–52.

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

3 Bedrosian, B., private communication, 1978.

4 Sonin, A. A., and Huber P. W., ASME *Journal of Heat Transfer,* Vol. 100, 1978, pp. 601–604.

5 Anderson, W. G., Huber, P. W., Sonin, A. A., ASME *Journal of Heat Transfer,* Vol. 100, 1978, pp. 605–612.

T. V. Nguyen
Student Mem. ASME.

C. E. Brennen
Mem. ASME.

R. H. Sabersky
Mem. ASME

Division of Engineering
and Applied Science 104-44,
California Institute of Technology,
Pasadena, Calif. 91125

Funnel Flow in Hoppers

Detailed observations of funnel flows of dry granular materials in wedge-shaped hoppers of different geometries are presented. The variations of the flow regime with changes in the height of material in the hopper/vertical bin configuration, the width of the vertical bin, the hopper angle and the hopper opening width were investigated and a number of specific flow regimes identified (mass flow and several forms of funnel flow). In the first part of the paper particular attention is paid to the conditions for transition from one flow regime to another; in particular it is shown that the existence of a funnel depends not only on the hopper angle but is also strongly dependent on the geometry of the hopper/bin system. In the second part of the paper the variations in the shape of the funnel near the exit opening are explored in detail.

1 Introduction

This paper is concerned with the flow of dry, noncohesive, granular materials in hoppers. It is well known that for plane (wedge-shaped) or conical hoppers of fairly small included angle all of the granular material flows in a fairly uniform and regular way. Such devices have been referred to as "mass flow hoppers" and have been the subject of considerable study and analyses (for example, references [1–6]). The convergence of experimental observations and theoretical predictions suggests that there exists some understanding of the mechanics of granular media flow in these circumstances. However, as the included angle is increased and a vertical bin is added to the top of the inclined sides of the hopper changes occur in the flow pattern which are much less well understood. Most of the motion occurs in a central core, funnel or "rat-hole" and stagnant regions of material tend to occur near the walls of the bin or hopper. This paper presents experimental observations of funnel flows in plane hoppers with vertical bins since it is not only of fundamental interest but is also important to the hopper designer. In this regard comparison is made with some of the existing design criteria such as that proposed by Jenike [1].

The various types of flow pattern which were observed in the present experiments are indicated in Fig. 1 (also shown are the definitions of θ_w, D, W, and H, the hopper angle and opening width, the bin width, and the total height to the upper free surface, respectively). Fig. 2(a) is an example of type A or mass flow. The flows with stagnant regions are subdivided into two basic types, B and C. Type B has

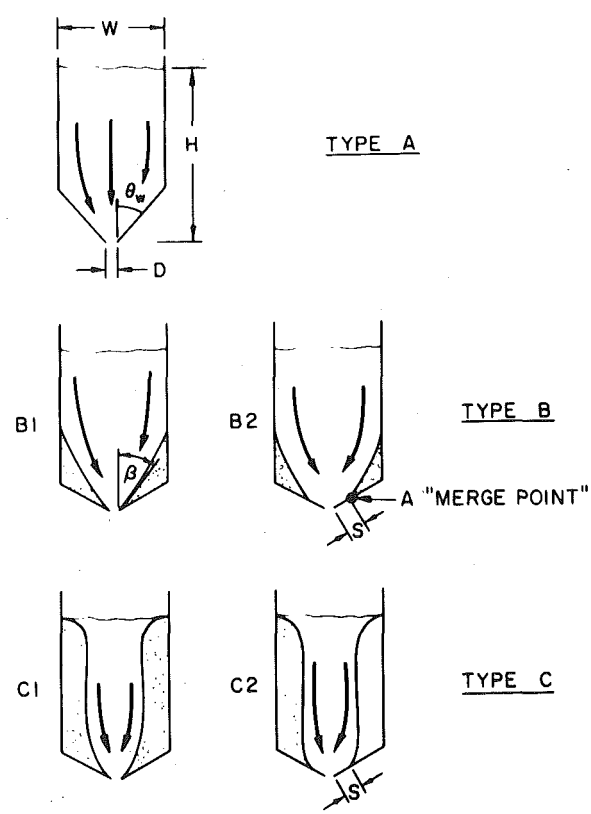


Fig. 1 Schematic indicating the different flow regimes observed. Type *A* is the mass flow regime. Types *B* and *C* are funnel flows, *B* having stagnant material in the corner and *C* having stagnant material on the sides of the bin. The geometric notation is also shown; the dimension, *b*, is the breadth of the flow or distance between the front and back vertical walls.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, November, 1979; final revision, April, 1980.

Table 1 Material properties (angles in degrees)

Material	Bulk Specific Gravity	Mean Diameter mm	Internal Friction Angle	Wall Friction Angle Lucite Wall	Wall Friction Angle Aluminum Wall	θ_{w_1} (*)	θ_{w_2} (*)
Sand	1.5	0.5 -1	31	15	18	40	70
Polystyrene	0.56	0.25-0.39	39	12	17	35	60
Glass Beads	1.46	0.325	25	15.3	17.7	30	50
Rice	0.8	-	30	-	-	55	70

(*) Values of θ_{w_1} and θ_{w_2} are for a smooth walled hopper with a thickness equal to 15.2 cm.

sliding along the upper part of the bin wall and stagnant regions near the bin/hopper wall corner (for example Fig. 2(*b*)). Type *C* has no sliding along the bin walls and larger stagnant regions on either side of the funnel (for example Fig. 2(*c*) and (*d*)). Two subtypes were also noted. In some instances slip occurred along the walls of the hopper (types *B*2 and *C*2) whereas in other cases the stagnant regions extended to the opening (types *B*1 and *C*1). Figs. 2(*b*) and 2(*d*) are representative of types *B*2 and *C*2 whereas Fig. 2(*c*) is of type *C*1.

Though there have been many observations of funnel flow patterns in both plane hoppers [1, 7–9, 11, 12] and conical hoppers [16–19] the variations with hopper geometry have not been exhaustively explored. The main objective of the present investigation was to study the flow fields occurring in plane hoppers over a wide range of geometries of a hopper/vertical bin system and for a wide range of cohesionless granular materials. Some effects of rough and smooth walls are also studied.

2 Experimental Apparatus

A plane hopper surmounted by a vertical bin was made of lucite. The dimensions H, W, D, θ_w (see Fig. 1) and the breadth or separation of the front and back vertical walls (denoted by b) were all adjustable. The hopper angle, θ_w, could be varied continuously; exit openings, D, ranged from 0.5 cm to 3.8 cm and bin widths, W, from 17.8 cm to 33 cm. Most observations were made with a breadth, b, of 15.2 cm but tests with sand were also performed with breadths of 7.6 cm and 22.9 cm. In some tests the height, H, was maintained constant during flow by means of a second supply bin; in other tests H decreased naturally with discharge. Finally, a series of experiments were conducted in hoppers with rough inclined walls and smooth vertical walls in order to observe the effects of inclined wall roughness.

The granular materials used were sand, glass beads, polystyrene pellets and rice; the grain shapes range from spherical (glass beads) to elongated (rice). The grain size distributions were fairly uniform and all the materials are practically cohesionless. Their physical properties, measured according to the procedures described by Pearce [20], are shown in Table 1.

3 Observation of the Transition Between Flow Patterns

The transition criteria for the flow of sand in a hopper of breadth, b = 15.2 cm will be discussed first. When the hopper angle, θ_w, was less than about 60° mass flow (type *A*) was observed to occur until the free surface reaced a critical height, H. Below this value of H, stagnant side flow of type *C* occurred. In Fig. 3, the critical ratio H/W is plotted

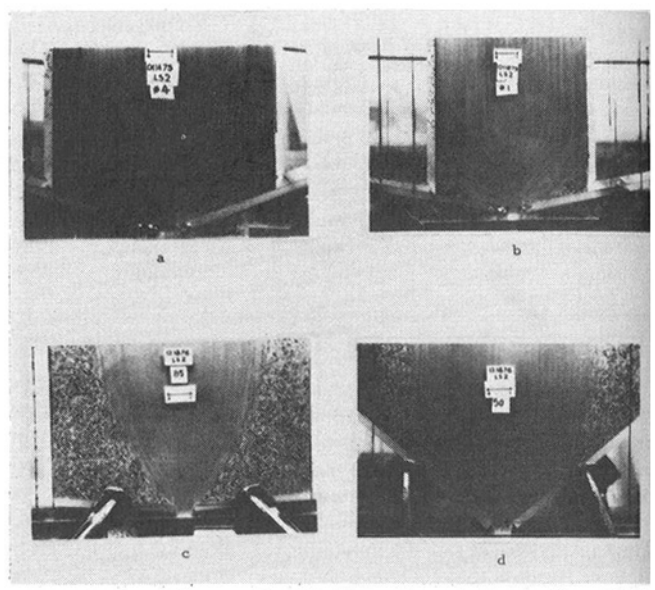


Fig. 2 Photographic examples of the flow patterns for the sand (φ = 31°). The breadth, b, is 15.2 cm in all cases. The following are the values of θ_w, H, W, D (in cm): (*a*) 70°, 36.8, 22.9, 1.91 (*b*) 80°, 58.4, 17.8, 1.37 (*c*) 70°, 35.6, 30.5, 2.54 (*d*) 50°, 35.6, 30.5, 2.54.

versus the ratio W/D for various hopper angles. For θ_w less than about 60°, the critical value of H/W is more or less constant for all angles θ_w. This implies that the transition from mass flow to type *C* is caused by the presence of the vertical bin on top of the hopper rather than by the inclination of the hopper walls.

For hopper angles greater than about 60°, flow of type *B* occurred for large H. However, the flow underwent a transition of type *C* when H/W reached the critical values plotted in Fig. 3. Again, the critical value is more or less independent of the hopper angle. (Note that even

Nomenclature

b = hopper breadth
D = width of exit opening
H = height of material above the exit opening
S = distance along side wall from edge of exit opening to merge point
W = hopper width
X = dimensionless horizontal position of funnel boundary
Y = dimensionless vertical position of funnel boundary
β = inclination of funnel to vertical at the discharge
θ_w = hopper angle

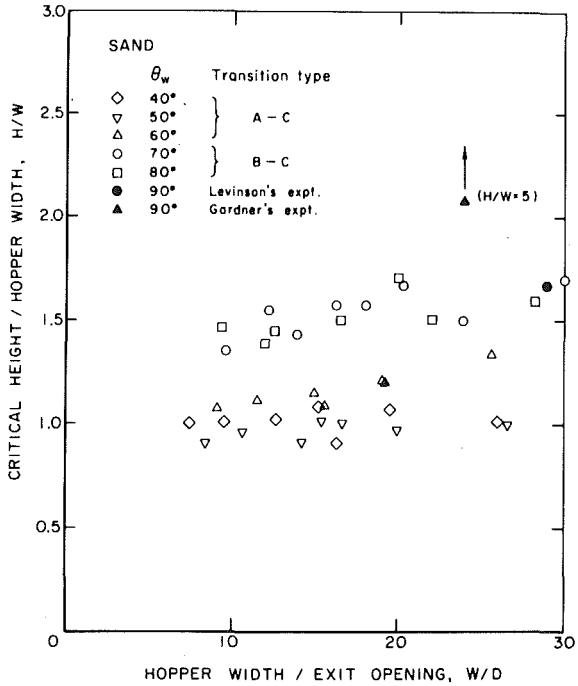


Fig. 3 Critical values of *H/W* plotted against *W/D* for various hopper angles, θ_w; the material is sand ($\varphi = 31°$, $d = 0.5 \rightarrow 1$ mm)

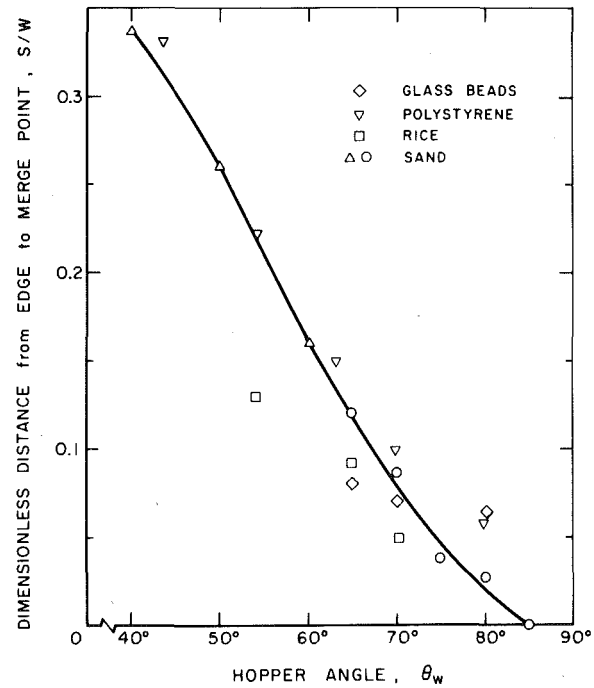


Fig. 4 Dimensionless hopper wall slip length, *S/W* plotted against the hopper angle, θ_w, for type *C* flow of the four materials used in the experiments

for hopper angles as large as 70° mass flow can occasionally occur at large H/W and small W/D).

It is important to emphasize that type C flow can be clearly distinguished from type B. The former does *not* occur simply as a result of the upper free surface engulfing the boundary between stagnant and flowing material in type B. The extent of the stagnant material distinctly increases in a B to C transition as can be seen by comparing the boundary geometries in Figs. 8 and 9.

In type B or C flow the stagnant material may either terminate at the discharge opening or it may end at a "merge point," S, which is some distance from the edge of the exit opening. In the latter case the material slides along the wall below the merge point. Consequently, mass flow is present in a localized region near the exit of the hopper. For convenience the subtypes $B2$ and $C2$ are defined as having a merge point on the hopper wall while in the subtypes $B1$ and $C1$ the merge point coincides with the edge of the discharge opening (see Fig. 1). The length of hopper wall over which slip occurs was measured for a variety of flows and will be denoted by S. Flows of type $B2$ (or $C2$) were observed to occur when the hopper angle was less than about 85°. As the angle is decreased below this the distance S increases monotonically as the flow begins a transition to mass flow. The magnitude of S also depended upon the width W. However the values of S/W for type C flows were primarily functions of θ_w as indicated in Fig. 4; the data for type B flows were limited and more scattered.

Fig. 5 is an alternative presentation of the information in Fig. 3. Here the minor variations with W/D are not explicitly shown. Rather the flow regimes are shown in a parametric map of H/W and θ_w. Fig. 5 is for sand and smooth walls; similar flow regime maps for glass beads and for sand with rough walls are presented in Figs. 6 and 7. Flow maps for other granular materials such as rice and polystyrene are included in reference [24]; they are qualitatively similar to Figs. 5 and 6.

These flow maps suggest that the phenomena are best described by defining three different ranges of hopper angle. When θ_w is between 0° and θ_{w1} mass flow (type A) occurs in the flow field irrespective of the ratio H/W. For values of θ_w between θ_{w1} and θ_{w2}, a dual transition behavior is observed. There is first a transition from type A flow into type C flow. This is followed by a second transition from type C flow into type A flow at a lower critical value of H/W. Finally,

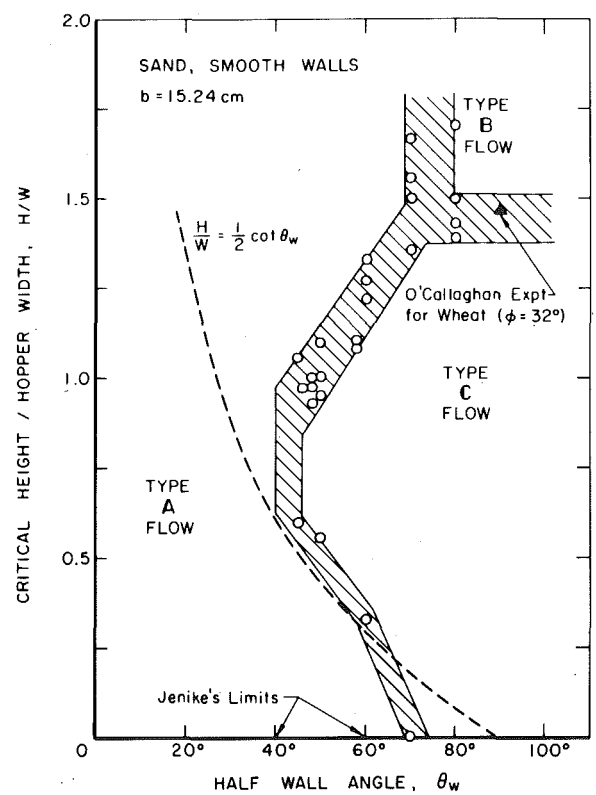


Fig. 5 Map of the flow regimes as a function of hopper geometry for the sand flowing in a smooth walled hopper. The solid lines are the approximate positions of the regime boundaries with actual transition points indicating the uncertainty in these boundaries. The dashed line locates conditions in a hopper without a bin for $W/D \gg 1$.

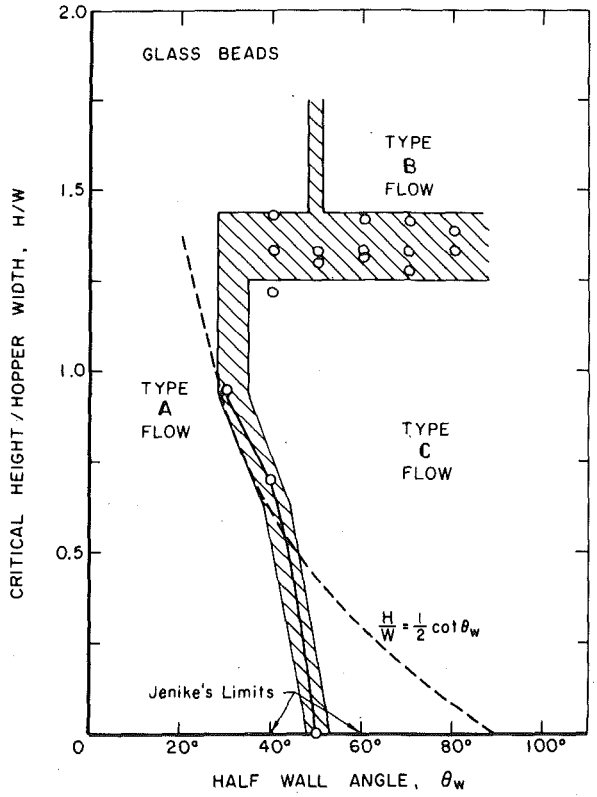


Fig. 6 Flow map as Fig. 5 but for glass beads

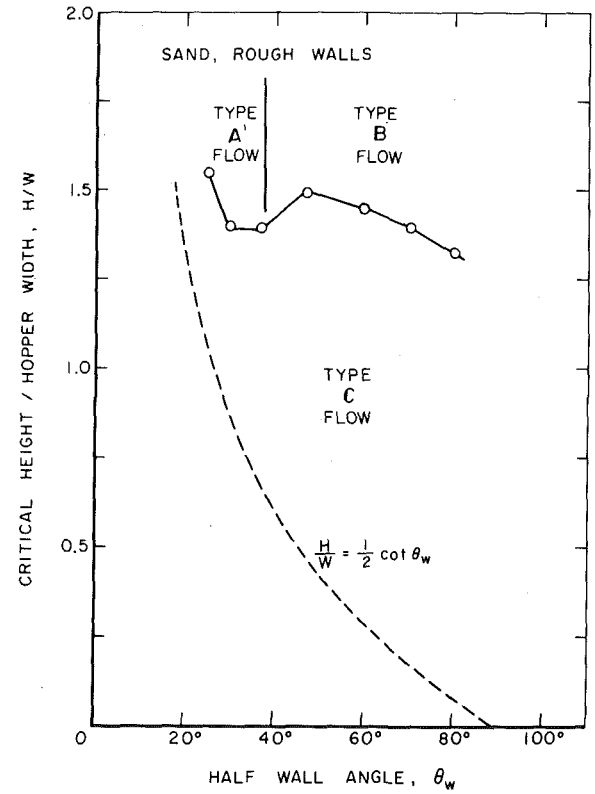


Fig. 7 Flow map as Fig. 5 but for sand in hoppers with rough inclined walls (the vertical bin walls are smooth). Type *A'* is similar to type *A* but has a thin layer of stagnant material next to the rough walls.

for values of the hopper angles θ_w greater than θ_{w2}, the transition is from type B flow into type C flow with stagnant material always present in the flow field. It should however be mentioned that some nonuniqueness can occur when the wall angles are close to the angles θ_{w1} and θ_{w2}; different flow regime histories can result for nominally identical tests. The values of θ_{w1} and θ_{w2} which appear to be functions of the frictional properties of the granular materials (and the walls) are presented in Table 1.

It follows from the foregoing that the common practice of using only values of θ_w and D as the relevant dimensions (e.g., [1]) would give a very incomplete picture of the flow field. One example is the effect of the vertical bin on the flow which can be demonstrated using the dashed lines in Figs. 5–7. These dashed lines represent the ratio H/W in a hopper without a vertical bin (for large values of W/D). Consequently such hoppers would exhibit mass flow up to the point at which the dashed line intersects that regime boundary. Note that in some cases in which a bin-less hopper exhibited mass flow (type A) the addition of a bin would change the flow regime to a funnel flow type.

As expected the material properties also play a role in determining the flow pattern. With the exception of glass beads, the higher the internal friction angle (see Table 1) of the material the lower the values of θ_{w1} and θ_{w2}. The apparent inconsistency represented by glass beads may be due to their small size (300 μm). Crewdson, Ormond, and Nedderman [21] have shown that the effects of the interstitial air can influence the discharge rate (and presumably other flow properties) when the particle size is less than about 500 μm. Other particle unique features which were noted included the tendency for the elongated rice grains to align themselves with the flow.

As previously stated the aforementioned observations were made with a breadth, b, equal to 15.2 cm. Some limited observations with breadths of 7.6 and 22.8 cm indicated the same qualitative transitional phenomena and only minor quantitative differences. Data on the variation of funnel shape with b/W are presented in Section 5 and suggests that the results have asymptoted to those of pure flow for values of b equal to 15.2 cm or greater.

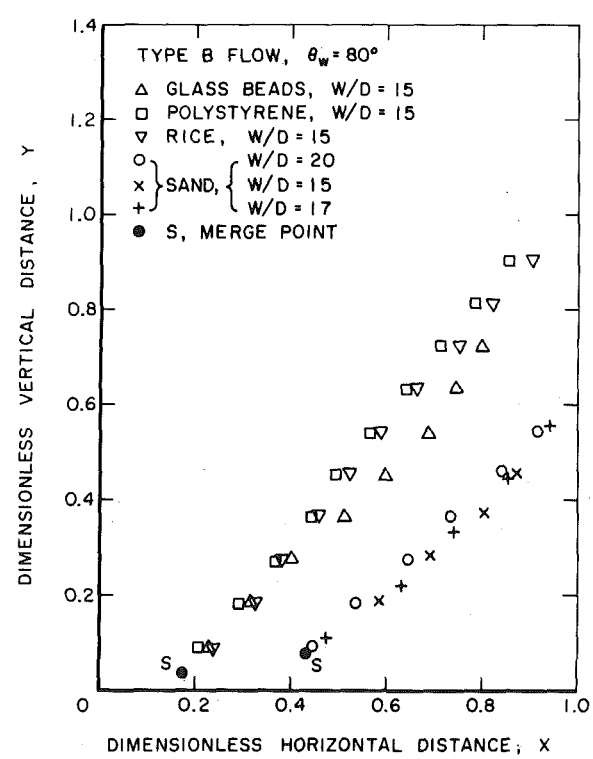


Fig. 8 Examples of funnel shapes for type *B* flows in smooth-walled hoppers (b = 15.2 cm). Shapes for given material are almost independent of θ_w, *H/W* or *W/D*.

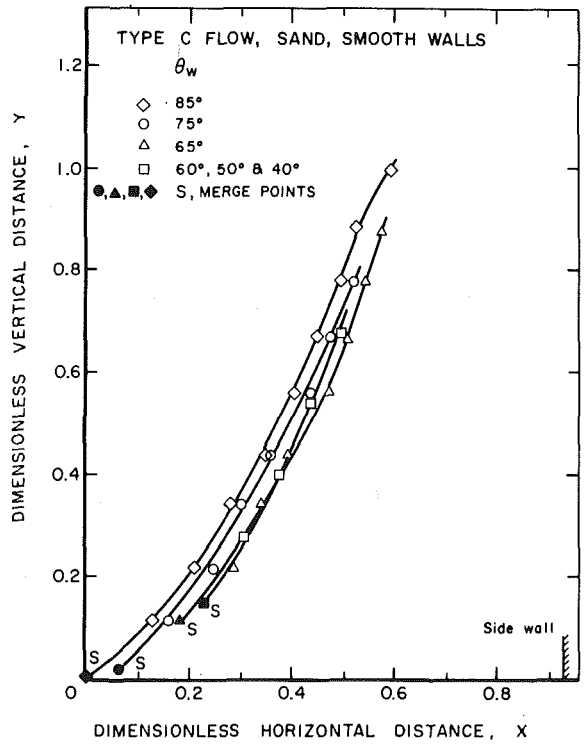


Fig. 9 Funnel shapes for type *C* flows of sand in smooth-walled hoppers of various θ_w (b = 15.2 cm)

4 Comparison With Previous Studies

Though no other complete parametric study has appeared in the literature it is valuable to compare previously observed flow patterns with those expected on the basis of the present study. O'Callaghan [7] examined the flow regimes in a flat-bottomed bin and noted the transition from type B to type C flow (which he termed "deep bin flow" and "shallow bin flow," respectively). He measured critical values for H/W of 1.47, 1.49, and 1.77 for wheat ($\varphi = 32°$), barley ($\varphi = 38°$), and fertilizer ($\varphi = 42°$), respectively. As indicated in Fig. 5 these are consistent with the present experiments. The points of operation (*not* the critical values) of the hoppers in Gardner's [8] experiments and those of Levinson, et al. [12], are shown in Fig. 3. Gardner's photographs clearly show that his flows were indeed of the type B; the values of H/W used by Levinson, et al., were marginal and this is reflected in the fact that the flows with clover seeds tended toward type B whereas, the flows with sand (which has a larger internal friction angle) were closer to type C. Though Brown and Richards [9] did not give the dimensions of their apparatus, their value of H/W appears to be about 1.5 and flows of type C were observed for a larger number of granular materials. In Toyama's [11] experiments flows of type B were encountered with large values of H/W of 10 and 6.5.

It is worth mentioning that trends similar to those previously reported for plane hoppers also appear to occur in conical hoppers. Van Zanten, et al. [16], and Giunta [18] observed flows of type C in conical systems with values of H/W of 2.57 and 1.33, respectively. On the other hand, Novosad and Surapati [19] obtained flows of type B for H/W ranging from 4 to 8. McCabe [17] observed a change in the flow field for values of H/W of about 2. Thus it would appear that conical hoppers exhibit results qualitatively similar to those reported here for wedge-shaped hoppers and that the critical value of H/W for conical hoppers is between 2 and 3 depending on the properties of the material.

Jenike [1] studied the conditions on the hopper geometry under which mass flow (type A) would occur. By balancing the stress at the exit against the strength of the material, he concluded that the *upper* limit of θ_w for which mass flow would occur was ($90° - \delta$) or 60°, whichever is the smallest (δ is the wall friction angle). He also proposed a *lower* limit based on observations for flows in hoppers with vertical bins (and a minimum exit opening D required to avoid arching). The two limiting values of θ_w are indicated in Figs. 5 and 6 and appear to correspond roughly with the angles θ_{w1} and θ_{w2} of the present study. Despite a number of studies [13–15] which have questioned the validity of Jenike's critera they have been extensively used during the past 20 years for the design of bins and hoppers. We would however suggest attention also be paid to the ratio H/W since according to the present studies this appears to be a crucial parameter; we note that Johanson and Colijin [22] have introduced the concept of a minimum height, H, required to insure mass flow.

5 Variations in the Shapes of the Funnels

Having established the conditions for the $A \rightarrow C$ and $B \rightarrow C$ transitions we now proceed to examine the changes in the shape of the funnel with geometry for both type B and type C flows. In the figures which follow these shapes are plotted nondimensionally by dividing all lengths by $W/2$; the origin of the resulting X, Y coordinates in which Y is vertical is taken at the end of the hopper wall at the discharge opening. Hence type $B1$ or $C1$ profiles end at the origin. In type $B2$ or $C2$ profiles the merge points are identified by the letter S. The objective of this section will be to identify the variations in funnel shape with the parameters θ_w, H/W, W/D, b/W, and the material properties.

The dimensionless funnel shapes for type B flows were found to be almost completely independent of θ_w, H/W or H/D for a given granular material (Gardner [8] noted the lack of dependence on θ_w in his experiments). One example of this is shown in Fig. 8 where several profiles at different W and D are included for sand: other data appears in reference [24]. The profiles for different granular materials are shown in Fig. 8. The profiles for glass beads, polystyrene pellets, and rice are quite similar. Only sand appears substantially different. However this is primarily caused by a difference in the location of the merge point; relative to that point the profiles are quite similar. It is not clear why the merge point for sand should be so different from that for the other materials in type B flows; it was not the case in type C flows.

The funnels occurring in type C flow are more variable and will first be described for sand with the understanding that the results for the other materials are similar. Fig. 9 displays the type C funnel shapes in sand for various hopper angles, θ_w, at fixed values of H, W, D, and b as indicated. As one proceeds to smaller angles the type $C2$ flow slides over longer lengths of the hopper wall; this feature was previously described in Fig. 4. The funnel becomes wider but the shape of the funnel remains much the same; indeed the profiles simply appear to have been shifted outward in the X-direction. The inclination of the funnel to the vertical at the merge point, S, appears to decrease somewhat. These trends halt at $\theta_w = 60°$ and the profiles for 60°, 50°, and 40° all correspond. Hence in the type A–C transition the resulting funnel seems to be independent of the angle whereas in the type B–C transition ($\theta_w > 60°$) the funnel shape depends on θ_w or more specifically on the position of the merge point as given by S/W.

Further type C funnel shape variations in sand are included in references [24, 25]. Briefly little or no variation in the funnel shape occurred with variations in either H/W or W/D or with changes in overall size. It was however observed that when H/W was somewhat greater than unity, the funnels were slightly larger than those for $H/W < 1$. This effect seems to reflect the tail end of the transition from type B flows to the type C flows.

In Fig. 10 comparison is made between the funnels for three different breadths, b, of hopper (b/W = 0.25, 0.5, and 0.75). The funnels for the two larger breadths are quite close. However, the funnel for the smallest breadth (b/W = 0.25) seems significantly narrower throughout its length. As mentioned previously, we have tentatively concluded from this that friction on the vertical front and back faces begins to alter the flow regime and funnel shape when b/W is less than some value between 0.25 and 0.4.

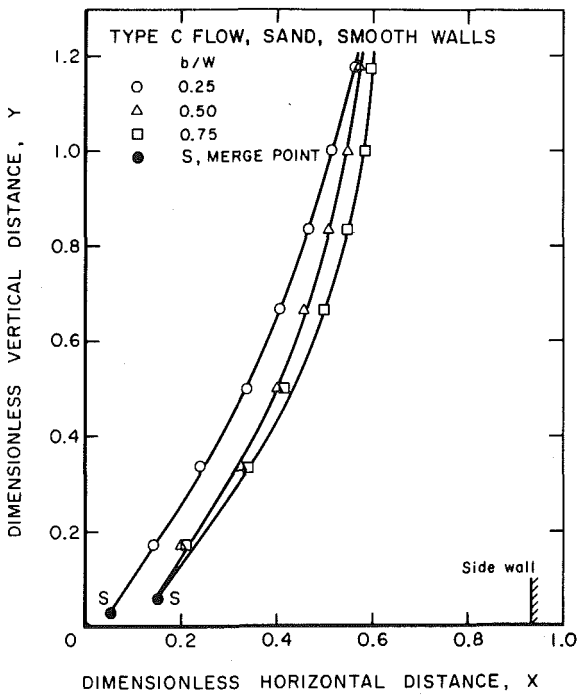


Fig. 10 Type *C* funnel shape variations with hopper breadth, *b*; for sand in smooth-walled hoppers of $\theta_w = 70°$

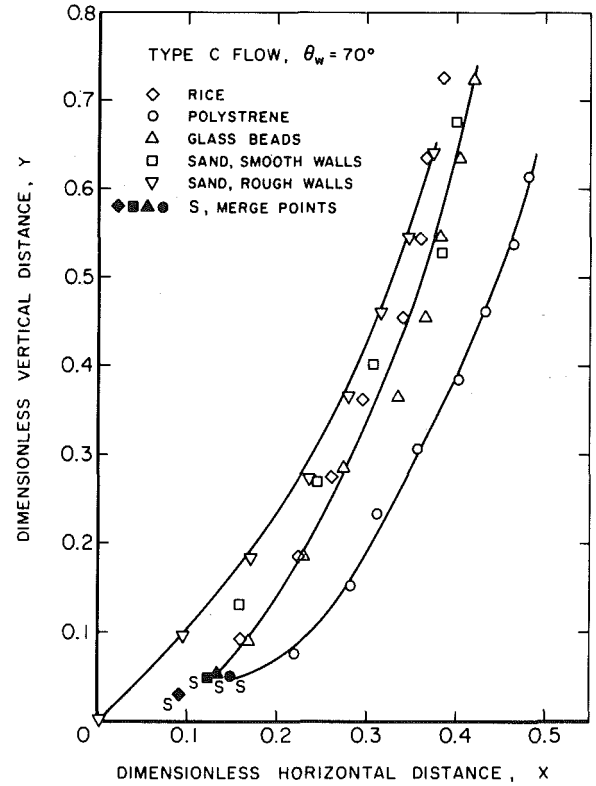


Fig. 11 Type *C* funnel shapes for different materials hopper of $\theta_w = 70°$ and $b = 15.2$ cm; funnels in sand with both smooth and rough inclined hopper walls are shown

In summary, these studies suggest that the type C funnel shapes are primarily dimensioned by the width, W, of the vertical bin. The hopper angle, θ_w, has a fairly simple effect shifting the profile outward as θ_w is decreased. Furthermore, provided the conditions are not close to the critical value of H/W for transition to another regime and provided b/W is 0.5 or larger, the funnel shape appears to be relatively independent of H/W, W/D, and b/W.

Finally a typical comparison between the type C funnel shapes in the four different granular materials is included in Fig. 10. In this respect polystyrene appears to be rather different from the other three materials.

6 The Effects of Rough Hopper Walls

It was envisaged that hopper wall roughness would affect the value of the wall friction angle and therefore the flow in the hopper, (Gardner [8], Bosley, et al. [23], and Savage and Sayed [26]). In the present study the experiments with sand were repeated in a hopper whose inclined walls were roughened by depositing sand on double-stick tape (Savage [26]). Results with vertical bin walls which were left smooth will be described here; some experiments were performed with these similarly roughened but the results dependent on the initial head of material loaded into the bin.

With rough inclined walls, a thin stagnant layer of material next to these walls was always present. Thus, at small angles θ_w, the mass flow regime will be described as type A' to denote this minor difference. Fig. 7 represents the flow map for sand in a hopper with rough inclined walls and is clearly different from that for smooth walls (Fig. 5). Even at small θ_w there appears to be a transition to type C at some critical H/W. It appears that the lower transition boundary in Fig. 7 has been rotated clockwise through about 90°. At hopper angles greater than about 40° the type B to type C transition occurs at values of H/W similar to the B–C transition for smooth walls.

The funnel boundaries in the rough-walled experiments were also insensitive to hopper geometry. It can be seen from the one example included in Fig. 11 that the primary difference between the rough and smooth wall funnels is in the location of the merge point.

It has been reported by a number of investigators (Bosley, et al. [23], and Savage and Sayed [26]), that the mass flow rate out of a hopper with rough walls can actually be *greater* than that from a similar smooth-walled hopper. Such a comparison was made in the present study, the results being presented in nondimensional form in Fig. 12 where U is the mean velocity corresponding to the discharge flow rate. It can be observed that though the smooth-walled hopper exhibits a higher flow rate at small hopper angles less than about 35°, the rough-walled hopper has the greater flow rate for $\theta_w > 35°$. Furthermore about $\theta_w = 90°$ the two converge to the same value again.

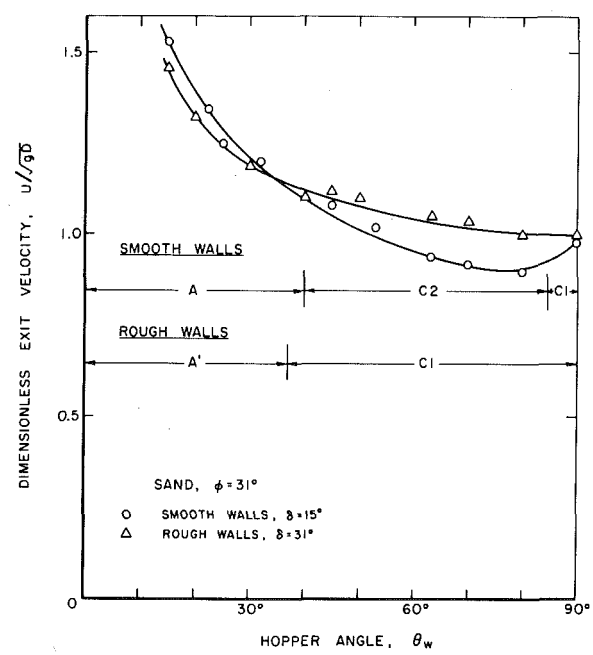


Fig. 12 The dimensionless exit velocity or discharge flow rate versus hopper angle, θ_w, for sand in hoppers with smooth and rough inclined walls

The present study revealed that such unexpected effects of wall roughness can be partially understood by reference to the different flow regimes which the two kinds of hopper exhibit over the range of hopper angles. For θ_w less than about 35°, the comparison is between types A and A'; then, as expected, the larger wall friction of the rough walls decreases the flow rate. As θ_w is increased above about 35° the rough-walled hopper produces a flow of type $C1$ in which the merge points are located at the discharge opening. On the other hand the smooth-walled hopper produces a flow of type $C2$ with merge points some considerable distance up the hopper walls. Consequently in the immediate neighborhood of discharge the flow in the rough-walled hopper experiences an "effective" hopper angle which is less than the actual hopper angle encountered in the smooth-walled hopper. Since the flow rate is primarily determined by the conditions in the flow in the immediate vicinity of discharge and since the flow rate tends to increase as the hopper angle decreases it is then possible to understand why the flow rate is greater for the rough-walled hopper.

As the hopper angle is increased further the flow rate from the rough-walled hopper changes little since its "effective" hopper angle is changing only slightly. On the other hand the actual hopper angle which is pertinent in the smooth-walled case continues to increase causing further substantial decrease in the flow rate. This continues until $\theta_w \approx 85°$ when the merge points in the smooth-walled case reach the discharge opening (see Fig. 4). At this point the flows in both types of hopper become of the type $C1$ and the flow rates converge to similar values.

Thus it can be seen that the effect of wall roughness on the flow rate is related to its effect upon the flow regime. It can also be concluded that the flow rate depends mainly upon the conditions near the exit and on the location of the merge points.

7 Concluding Remarks

In the present study, the various types of flow which exist in a hopper with a vertical bin have been identified and classified. The experimental observations show that the presence of the vertical bin will cause funnel flow to occur at lower values of the hopper wall angle θ_w. The ratio of the height of the material in the bin to its width (H/W) is important in determining the type of flow which is present and the transition from one type of flow into another.

The nondimensionalized funnel boundary is found to be independent of the hopper angle θ_w, the width of the exit opening D, and the width of the vertical bin W. It is mainly a function of the material properties. Some changes in the flow field due to the proximity of the front and back walls are observed when the hopper thickness falls below a certain limit. Finally, the presence of the wall roughness affects the flow field in the hopper by causing stagnant material to appear at lower values of θ_w. This change in the flow field is responsible for the fact that the rate of discharge from a hopper with rough walls is actually slightly *higher* than that from a hopper with smooth walls when the hopper wall angle θ_w is greater than about 35°.

Acknowledgments

The authors wish to acknowledge the contribution by Paul Mason who enthusiastically and effectively assisted in the early laboratory work. He is being sincerely missed by all of us. The authors would like to express their gratitude to the National Science Foundation for support under Grant No. ENG-7615043 and to Union Carbide Corp. for their continuing interest and financial support.

References

1 Jenike, A. W., "Storage and Flow of Solids," Bulletin No. 123, 1964, Utah Engineering Experiment Station, University of Utah, Salt Lake City, Utah.

2 Savage S. B., "Gravity Flow of Cohensionless Bulk Solids in a Converging Conical Channel," *International Journal of the Mechanical Science,* Vol. 9, 1963, p. 651.

3 Brennen, C., and Pearce, J. C., "Granular Material Flow in Two-Dimensional Hoppers," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, p. 43.

4 Nguyen, T. V., Brennen, C., and Sabersky, R. H., "Gravity Flow of Granular Materials in Conical Hoppers," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, Sept. 1979, pp. 529–535.

5 Lee, J., Cowin, S. C. and Templeton III, J. S., "An Experimental Study of the Kinematics of Flow Through Hoppers," *Transactions of the Society of Rheology,* Vol. 18, No. 2, 1974, p. 247.

6 Drescher, A., Cousens, T. W., and Bransby, P. L., "Kinematics of the Mass Flow of Granular Materials Through a Plane Hopper," *Geotchnique,* Vol. 28, No. 1, 1978, p. 27.

7 O'Callaghan, J. R., "Internal Flow in Moving Beds of Granular Materials," *Journal Agr. Engng. Res.,* Vol. 5, 1960, p. 200.

8 Gardner, G. C., "The Region of Flow When Discharging Granular Materials From Bins," *Chemical Engineering Science,* Vol. 21, 1966, p. 261.

9 Brown, R. L., and Richards, J. C., *Principles of Powder Mechanics,* Pergamon Press, 1970.

10 Wieghardt, K., "Experiments in Granular Flow," *Annual Review of Fluid Mechanics,* Vol. 7, 1975, p. 89.

11 Toyama, S., "The Flow of Granular Materials in Moving Beds," *Powder Technology,* Vol. 4, 1970–1971, p. 214.

12 Levinson, M., Shmutter, B., and Resnick, W., "Displacement and Velocity Field in Hoppers," *Powder Technology,* Vol. 16, 1977, p. 29.

13 Walker, D. M., "A Basis for Bunker Design," *Powder Technology,* Vol. 1, 1967, p. 228.

14 Eckhoff, R. K., and Leversen, P. G., "A Further Contribution to the Evaluation of the Jenike Method for Design of Mass Flow Hoppers," *Powder Technology,* Vol. 10, 1974, p. 51.

15 Wright, H., "An Evaluation of the Jenike Bunker Design Method," ASME *Journal of Engineering for Industry,* Vol. 95, 1973, p. 48.

16 Van Zanten, D. C., Richards, P. C., and Mooiji, A., "Bunker Design—Part 3: Wall Pressures and Flow Patterns in Funnel Flow," ASME *Journal of Engineering for Industry,* Vol. 99, 1977, p. 819.

17 McCabe, R. P., "Flow Pattern in Granular Material in Circular Silos," *Geotechnique,* Vol. 24, 1974, p. 45.

18 Giunta, J. S., "The Study of Flow Patterns of Granular Solids in Flat-Bottom Bins With Circular Openings," MS Thesis, University of Pittsburgh, 1966.

19 Novosad, J., and Surapati, K., "Flow of Granular Materials: Determination and Interpretation of Flow Patterns," *Powder Technology,* Vol. 2, 1968–1969, p. 82.

20 Pearce, J. C., "Mechanics of Flowing Granular Media," PhD Thesis, California Institute of Technology, 1976.

21 Crewdson, B. J., Ormond, A. L., and Nedderman, R. O., "Air Impeded Discharge of Fine Particles From a Hopper," *Powder Technology,* Vol. 16, 1977, p. 197.

22 Johanson, J. R., and Colijin, H., "New Design Criteria for Hoppers and Bins," *Iron and Steel Engineer,* Oct. 1964, p. 85.

23 Bosley, J., Schofield, C. and Shook, C. A., "An Experimental Study of Granular Discharge From Model Hoppers," *Transactions of the Institution of Chemical Engineers,* Vol. 47, 1969, p. T147.

24 Nguyen, T. V., "Studies in the Flow of Granular Materials," PhD Thesis, California Institute of Technology, Pasadena, Calif., 1979.

25 Nguyen, T. V., Brennen, C., and Sabersky, R. H. "Funnel Flow in Hoppers," *Mechanics Applied for the Transport of Bulk Materials,* ASME, AMD, Vol. 31, 1979, p. 25.

26 Savage, S. B., and Sayed, M., "Gravity Flow of Cohensionless Granular Materials in Wedge-Shaped Hoppers," *Mechanics Applied to the Transport of Bulk Materials,* ASME, AMD, Vol. 31, 1979, p. 1.

D. Durban

Department of Aeronautical Engineering,
Technion—Israel Institute of Technology,
Haifa, Israel

Drawing of Tubes

The process of the tube drawing between two rough conical walls is analyzed within the framework of continuum plasticity. Material behavior is modeled as rigid/linear-hardening along with the von-Mises flow rule. Assuming a radial flow pattern and steady state flow conditions it becomes possible to obtain an exact solution for the stresses and velocity. Useful relations are derived for practical cases where the nonuniformity induced by wall friction is small. A few restrictions on the validity of the results are discussed.

Model Presentation

We consider an axially symmetric forming process where the dimensions of a circular cylindrical tube are reduced by drawing it through rigid conical dies (Fig. 1). Of primary interest here is the working zone where severe plastic deformation is imposed upon the material. One would like in particular to find the dependence of the drawing stress t on the dies geometry, walls friction, material properties, and amount of back-pull t_b.

The present study of that drawing process employs a model which is based on three underlying assumptions

First, we assume that the material is rigid/linear-hardening according to the von-Mises flow rule

$$\frac{2}{3}E_T\mathbf{D} = \frac{\dot{\sigma}_e}{\sigma_e}\mathbf{S} \quad \sigma_e \geq Y \tag{1}$$

where $\mathbf{D}$ is the Eulerian strain rate, $\mathbf{S}$ the stress deviator, σ_e the effective stress, E_T the constant tangent modulus, and Y the yield stress. The superposed dot denotes differentiation with respect to time.

The second assumption is that the drawing process takes place in steady-state conditions so that the Eulerian form of (1) is

$$\frac{2}{3}E_T\mathbf{D} = \frac{\mathbf{V}\cdot\nabla\sigma_e}{\sigma_e}\mathbf{S} \quad \sigma_e \geq Y \tag{2}$$

where $\mathbf{V}$ is the velocity vector and ∇ the left gradient operator.

Finally we assume that the flow field within the walls is radially directed toward the origin O. Thus, introducing a spherical-polar system of coordinates (r, θ, ϕ) along with the unit triad $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi)$, we can write the velocity vector as

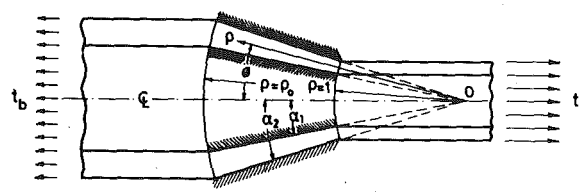


Fig. 1 Notation for tube drawing

$$\mathbf{V} = -V_r\mathbf{e}_r \tag{3}$$

where V_r is independent of ϕ because of axial symmetry.

The main advantage offered by that model is the possibility of arriving at an exact closed-form solution of the governing field equations. This has already been observed in [1] where a similar model was used to simulate drawing or extrusion of wires. The equivalent two-dimensional model was applied in [2] to sheet drawing or extrusion. The theoretical predictions of [1, 2] show good agreement with experimental results for the corresponding forming processes. We can therefore expect the present analysis for tube drawing to be of practical value as well.

Observing that the material defined by (1) is incompressible, we find from (3), via the incompressibility condition $tr\mathbf{D} = 0$, that the radial velocity has to be of the form

$$V_r = \frac{f(\theta)}{r^2} \tag{4}$$

where $f(\theta)$ is an unknown function of θ. The Eulerian strain rate is then given by

$$\mathbf{D} = \frac{1}{r^3}\left[f(\theta)(3\mathbf{e}_r\mathbf{e}_r - \mathbf{I}) - \frac{1}{2}f'(\theta)(\mathbf{e}_r\mathbf{e}_\theta + \mathbf{e}_\theta\mathbf{e}_r)\right] \tag{5}$$

where $\mathbf{I}$ is the second-order unit tensor and the prime denotes differentiation with respect to θ.

Turning now to the stress field we have that $\tau_{r\phi} = \tau_{\theta\phi} = 0$ as a result

Contributed by the applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, December 1979; final revision, April 1980. Paper No. 80-WA/APM-17.

of axial symmetry and that $\sigma_\theta = \sigma_\phi$ because of the coaxiality of tensors **D** and **S**. The stress deviator is therefore reduced to the form

$$\mathbf{S} = \frac{1}{3}(\sigma_r - \sigma_\theta)(3\mathbf{e}_r\mathbf{e}_r - \mathbf{I}) + \tau_{r\theta}(\mathbf{e}_r\mathbf{e}_\theta + \mathbf{e}_\theta\mathbf{e}_r) \tag{6}$$

so that

$$\sigma_e{}^2 = \frac{3}{2}\mathbf{S}\cdot\cdot\mathbf{S} = (\sigma_r - \sigma_\theta)^2 + 3\tau_{r\theta}{}^2 \tag{7}$$

Substitution of (3)–(6) into the constitutive equation (2) followed by some simple manipulations, described in [1], gives the three basic relations

$$\tilde{\sigma}_r - \tilde{\sigma}_\theta = -\xi \ln \rho + \Phi \tag{8}$$

$$\tilde{\tau}_{r\theta} = \frac{\sqrt{3}}{3}\beta(\xi \ln \rho - \Phi) \tag{9}$$

$$\tilde{\sigma}_e = (1 + \beta^2)^{1/2}(-\xi \ln \rho + \Phi) \tag{10}$$

where the superposed tilde indicates nondimensionalization with respect to the yield stress Y, $\xi = 2E_T/Y$ is the nondimensional hardening parameter, ρ is a suitably nondimensionalized radial coordinate, Φ is an unknown function of θ and

$$\beta = \frac{\sqrt{3}}{6}\frac{f'(\theta)}{f(\theta)} \tag{11}$$

Thus the basic relations, (8)–(10), obtained from the constitutive equation and kinematics, contain two unknown functions of θ. These functions will be determined in the next section where the equilibrium requirements are discussed.

Equilibrium Equations and Solution

Neglecting inertial effects and noting the axial symmetry of the stress field, we are left with the two quasi-static equilibrium equations

$$\frac{\partial \tilde{\sigma}_r}{\partial \rho} + \frac{1}{\rho}\frac{\partial \tilde{\tau}_{r\theta}}{\partial \theta} + \frac{2(\tilde{\sigma}_r - \tilde{\sigma}_\theta)}{\rho} + \frac{1}{\rho}\tilde{\tau}_{r\theta}\cot\theta = 0 \tag{12}$$

$$\frac{\partial \tilde{\tau}_{r\theta}}{\partial \rho} + \frac{1}{\rho}\frac{\partial \tilde{\sigma}_\theta}{\partial \theta} + \frac{3\tilde{\tau}_{r\theta}}{\rho} = 0 \tag{13}$$

Inserting relations (8)–(9) into (12) and integrating over ρ gives, for the radial stress component,

$$\tilde{\sigma}_r = -\xi\left[\frac{\sqrt{3}}{6}(\beta' + \beta\cot\theta) - 1\right]\ln^2\rho + \left\{\frac{\sqrt{3}}{3}[(\beta\Phi)' + \beta\Phi\cot\theta] - 2\Phi\right\}\ln\rho + F(\theta) \tag{14}$$

where F is a new unknown function of θ. By now we have three unknown functions of θ: β, Φ, and F. These functions are determined by three-differential equations that result from the second equilibrium equation (13). It is easily verified, [1], upon substituting $\tilde{\tau}_{r\theta}$ from (9) and $\tilde{\sigma}_\theta$ from (8) and (14), into (13), that equilibrium in the θ-direction is maintained if

$$(\beta' + \beta\cot\theta)' = 0 \tag{15}$$

$$\frac{\sqrt{3}}{3}[(\beta\Phi)' + \beta\Phi\cot\theta]' - 2\Phi' + \sqrt{3}\xi\beta = 0 \tag{16}$$

$$F'' - \Phi' - \sqrt{3}\beta\Phi + \frac{\sqrt{3}}{3}\xi\beta = 0 \tag{17}$$

The solution of equations (15)–(17) provides explicit expressions for the stresses and the velocity profile. The method of solution is essentially as in [1] only that here we present the *complete solution* of the system (15)–(17), namely,

$$\beta = -\frac{2\sqrt{3}k}{\sin\theta}\left(\sin^2\frac{\theta}{2} - \frac{k_1}{2k}\right) \tag{18}$$

$$\Phi = \frac{\xi}{k+2}\left[h(\theta) - \sqrt{3}\,I\left(\theta, k, \frac{k_1}{2k}\right)\right] + A_1 g(\theta) + A \tag{19}$$

$$F = \Phi + \frac{\xi}{k+2}\left[\frac{1}{2}h^2(\theta) - 3J\left(\theta, k, \frac{k_1}{2k}\right)\right] + \left(A - \frac{1}{3}\xi\right)h(\theta) + \sqrt{3}\,A_1 K\left(\theta, k, \frac{k_1}{2k}\right) + B \tag{20}$$

where k, k_1, A, A_1, B are integration constants,

$$g(\theta) = \left(\sin^2\frac{\theta}{2} - \frac{k_1}{2k}\right)^{-(k+2)/k} \tag{21}$$

$$h(\theta) = 3\ln\left(\cos\frac{\theta}{2}\right)^{2k}\left(\tan\frac{\theta}{2}\right)^{k_1} \tag{22}$$

and

$$I\left(\theta, k, \frac{k_1}{2k}\right) = g(\theta)\int\frac{\beta(\theta)}{g(\theta)}d\theta \tag{23}$$

$$J\left(\theta, k, \frac{k_1}{2k}\right) = \int I\left(\theta, k, \frac{k_1}{2k}\right)\beta(\theta)d\theta \tag{24}$$

$$K\left(\theta, k, \frac{k_1}{2k}\right) = \int g(\theta)\beta(\theta)d\theta \tag{25}$$

The restricted version given in [1] is obtained from this general solution with $k_1 = 0$. In that case it is more convenient to express integrals (23)–(25) by means of the incomplete Beta function.

Combining (18) with (11) and solving for $f(\theta)$ gives the velocity profile

$$f(\theta) = Ue^{2h(\theta)} = U\left(\cos\frac{\theta}{2}\right)^{12k}\left(\tan\frac{\theta}{2}\right)^{6k_1} \tag{26}$$

where the constant U stands for some reference velocity.

Explicit expressions for the stresses are now readily obtained. Inserting relations (18)–(19) into (14) yields

$$\tilde{\sigma}_r = \frac{1}{2}\xi(k+2)\ln^2\rho - [\xi h(\theta) + (k+2)A]\ln\rho + F \tag{27}$$

Similarly, from (8) and (27),

$$\tilde{\sigma}_\theta = \frac{1}{2}\xi(k+2)\ln^2\rho - [\xi h(\theta) + (k+2)A - \xi]\ln\rho + F - \Phi \tag{28}$$

The corresponding expressions for $\tilde{\tau}_{r\theta}$ and $\tilde{\sigma}_e$ are given by (9) and (10).

The solution of the governing field equations is now complete. It remains, though, to determine the five integration constants k, k_1, A, A_1, B through the introduction of proper stress boundary data.

It should be mentioned that the general solution presented here is, in general, not correct when $\xi = 0$, i.e., for rigid/perfectly plastic materials. This curious situation is revealed at once by equation (10). With $\xi = 0$ we do not obtain the basic constitutive relation, of a rigid/perfectly plastic material, $\tilde{\sigma}_e \equiv 1$ but rather a violation of that relation. The correct solution for axially symmetric radial flow of rigid/perfectly plastic materials has been given by Shield in [3].

It can be shown however, that in the absence of friction our general solution does agree, at the limit of zero hardening, with Shield's solution. Taking $k = k_1 = A_1 = 0$ we have the nonvanishing stress components

$$\tilde{\sigma}_r = \xi\ln^2\rho - 2A\ln\rho + A + B \tag{29}$$

$$\tilde{\sigma}_\theta = \xi\ln^2\rho - (2A - \xi)\ln\rho + B \tag{30}$$

along with the uniform velocity profile, (26),

$$f(\theta) \equiv U \tag{31}$$

The corresponding effective stress is then simply

$$\tilde{\sigma}_e = -\xi\ln\rho + A \tag{32}$$

Equations (29)–(32) agree, at the limit of zero hardening, with Shield's solution for uniform flow provided that $A \to 1$ as $\xi \to 0$. More details on that point can be found in the Appendix of [1].

Further Investigation of the Solution

It is instructive, at this stage of the inquiry, to suggest a simple interpretation of the quantity $k_1/2k$. Observing from (9)–(10) that the ratio between the shear stress and the effective stress is constant along a streamline, we may introduce the shear factors m_1, m_2 at the walls (Fig. 1) so that

$$\sqrt{3}\,\tilde{\tau}_{r\theta} = -m_1\tilde{\sigma}_e \quad \text{at} \quad \theta = \alpha_1 \tag{33a}$$

$$\sqrt{3}\,\tilde{\tau}_{r\theta} = m_2\tilde{\sigma}_e \quad \text{at} \quad \theta = \alpha_2 \tag{33b}$$

Substituting $\tilde{\tau}_{r\theta}$ and $\tilde{\sigma}_e$ from (9)–(10) into (33), with β given by (18), and solving for k, k_1 gives

$$k = \frac{\sqrt{3}}{6}\,\frac{\overline{m}_2 \sin\alpha_2 + \overline{m}_1 \sin\alpha_1}{\sin^2\frac{\alpha_2}{2} - \sin^2\frac{\alpha_1}{2}} \tag{34}$$

$$k_1 = \frac{\sqrt{3}}{3}\,\frac{\overline{m}_2 \sin\alpha_2 \sin^2\frac{\alpha_1}{2} + \overline{m}_1 \sin\alpha_1 \sin^2\frac{\alpha_2}{2}}{\sin^2\frac{\alpha_2}{2} - \sin^2\frac{\alpha_1}{2}} \tag{35}$$

where $\overline{m}$ is the modified shear factor defined by

$$\overline{m} = m/\sqrt{1-m^2} \tag{36}$$

Recalling that in the tube drawing process both shear factors, at the walls, are positive we deduce from (34)–(35) that

$$k,\, k_1 > 0 \tag{37}$$

Furthermore, the relation

$$\frac{k_1}{2k} = \frac{\overline{m}_2 \sin\alpha_2 \sin^2\frac{\alpha_1}{2} + \overline{m}_1 \sin\alpha_1 \sin^2\frac{\alpha_2}{2}}{\overline{m}_2 \sin\alpha_2 + \overline{m}_1 \sin\alpha_1} \tag{38}$$

implies that

$$\sin^2\frac{\alpha_1}{2} < \frac{k_1}{2k} < \sin^2\frac{\alpha_2}{2} \tag{39}$$

Thus the angle α_m defined by

$$\sin^2\frac{\alpha_m}{2} = \frac{k_1}{2k} \tag{40}$$

is located within the walls of the dies

$$\alpha_1 < \alpha_m < \alpha_2 \tag{41}$$

Just to get a feeling about the location of α_m, consider the case where α_1, α_2 are small and $\overline{m}_1 = \overline{m}_2 = \overline{m}$. From (38) and (40) we then obtain the elegant relation

$$\alpha_m = \sqrt{\alpha_1\alpha_2}$$

An immediate consequence of (41) is that the function $g(\theta)$, defined by (21), becomes singular at $\theta = \alpha_m$. That singularity is carried further, via (19)–(20), into the stress field and in order to avoid it we shall take the value of

$$A_1 = 0 \tag{42}$$

The angle $\theta = \alpha_m$ presents itself as a natural lower limit for the integrals (23)–(25). The third one, (25), is now of course irrelevant because of (42). As for the other two integrals, it is worth mentioning that both admit a regular expansion near $\theta = \alpha_m$. Indeed, with

$$u = \sin^2\frac{\theta}{2} - \frac{k_1}{2k} = \sin^2\frac{\theta}{2} - \sin^2\frac{\alpha_m}{2}$$

we find from (23)–(24) that near $u = 0$

$$I(\theta, k, \alpha_m) = g(\theta)\int_{\alpha_m}^{\theta}\frac{\beta(\theta)}{g(\theta)}\,d\theta \simeq -\frac{4\sqrt{3}\,k^2}{2+3k}\left(\frac{u}{\sin\alpha_m}\right)^2 \tag{43a}$$

$$J(\theta, k, \alpha_m) = \int_{\alpha_m}^{\theta} I(\theta, k, \alpha_m)\beta(\theta)\,d\theta \simeq \frac{12k^3}{2+3k}\left(\frac{u}{\sin\alpha_m}\right)^4 \tag{43b}$$

It is now seen that the shear stress (9) vanishes along the streamline that passes through $\theta = \alpha_m$. Also, as observed from (26) and (11), the velocity profile has its highest value at that particular streamline. We may conclude therefore that the nonuniformity of the flow field, induced by the friction along the walls, decreases as $\theta \to \alpha_m$ and attains its smallest value at $\theta = \alpha_m$.

A possible measure of the local nonuniformity is simply the local shear factor $m(\theta)$ defined as

$$\sqrt{3}\,|\tilde{\tau}_{r\theta}| = m(\theta)\tilde{\sigma}_e$$

with the obvious values

$$m(\alpha_1) = m_1, \quad m(\alpha_m) = 0, \quad m(\alpha_2) = m_2$$

The measure $m(\theta)$ can be interpreted in a different way. Recalling Truesdell's measure of vorticity, [4],

$$M = \left(\frac{|\mathbf{\Omega}\cdot\cdot\mathbf{\Omega}|}{\mathbf{D}\cdot\cdot\mathbf{D}}\right)^{1/2}$$

where $\mathbf{\Omega} = \frac{1}{2}(\nabla\mathbf{V} - \mathbf{V}\nabla)$ is the spin tensor, we find that for the radial flow field

$$\mathbf{\Omega} = -\frac{1}{2r^3}f'(\theta)(\mathbf{e}_r\mathbf{e}_\theta - \mathbf{e}_\theta\mathbf{e}_r)$$

Inserting the last relation along with (5) into the formula for M, and observing (9)–(11), gives the remarkable identity

$$M \equiv m$$

The local shear factor is precisely Truesdell's measure of vorticity.

A different and perhaps more practical measure of the local nonuniformity is the local modified shear factor $\overline{m}(\theta)$ defined, in the spirit of (36), as

$$\overline{m}(\theta) = m(\theta)/[1 - m^2(\theta)]^{1/2} = \frac{\sqrt{3}|\tilde{\tau}_{r\theta}|}{\sigma_e}\Big/\left(1 - \frac{3\tilde{\tau}_{r\theta}^2}{\sigma_e^2}\right)^{1/2}$$

Or, in view of (9)–(10),

$$\overline{m}(\theta) \equiv |\beta(\theta)| \tag{44}$$

Thus the absolute value of the function $\beta(\theta)$ is a measure of local nonuniformity. That measure has the advantage of appearing directly in the expressions for the stresses.

The difference between the two measures decreases of course as the nonuniformity becomes smaller, and for very small deviations from uniformity they are practically the same

$$\overline{m}(\theta) \simeq m(\theta) \ll 1$$

Tube Drawing With Small Nonuniformity

In practice, the process of tube drawing is performed with very well-lubricated walls. We can therefore take advantage of this fact and proceed with an approximated version, of the general solution, suitable for the case of small nonuniformity

$$|\beta(\theta)| \ll 1 \tag{45}$$

In addition, we assume that the half-die angles α_1, α_2 are sufficiently small to admit approximations of the type $\sin\theta \simeq \theta$. This assumption is in accordance with practice where α_2 is seldom higher than 15°.

Observing now, from (18) and (40), that for small angles

$$\beta(\theta) \simeq -\frac{\sqrt{3}}{2}k\,\frac{\theta^2 - \alpha_m^2}{\theta} \tag{46}$$

we find that (45) implies the restriction

$$k|\theta^2 - \alpha_m^2| \ll \theta \tag{47}$$

With these simplifications in mind, we may replace the exact expressions for the stresses by simple yet consistent and fairly close approximations.

Consider first integrals (43a–b). Noting that both $|\beta(\theta)|$ and $|g(\theta)|^{-1}$ increase monotonously with $|\theta - \alpha_m|$ we have the obvious bound

$$\left|\int_{\alpha_m}^{\theta} \frac{\beta(\theta)}{g(\theta)} d\theta\right| < \left|\frac{\beta(\theta)}{g(\theta)}\right| |\theta - \alpha_m|$$

Hence

$$|I(\theta, k, \alpha_m)| < |\theta - \alpha_m| \; |\beta(\theta)| \tag{48a}$$

Likewise, combining (48a) with (43b) gives

$$|J(\theta, k, \alpha_m)| < (\theta - \alpha_m)^2 \beta^2(\theta) \tag{48b}$$

Next, consider the function $h(\theta)$ defined by (22). Here it is worth noting, via (11) and (26), the differential relation $h'(\theta) = \sqrt{3}\,\beta(\theta)$ or

$$h(\theta) = h(\alpha_m) + \sqrt{3} \int_{\alpha_m}^{\theta} \beta(\theta) d\theta$$

where the integral is again bounded by

$$\left|\int_{\alpha_m}^{\theta} \beta(\theta) d\theta\right| < |\theta - \alpha_m| \; |\beta(\theta)| \tag{49}$$

Turning now to functions Φ and F, given by (19)–(20), and introducing the two new constants

$$\overline{A} = \frac{\xi}{k+2} h(\alpha_m) + A \tag{50a}$$

$$\overline{B} = \frac{\xi}{2(k+2)} h^2(\alpha_m) + \left(A - \frac{1}{3}\xi\right) h(\alpha_m) + B \tag{50b}$$

we find that for small nonuniformity the consistent first-order approximations are

$$\Phi \simeq \overline{A}, \quad F \simeq \overline{A} + \overline{B} \tag{51}$$

and the stresses (9)–(10), (27)–(28) follow immediately as

$$\tilde{\tau}_{r\theta} = \frac{k}{2} \frac{\theta^2 - \alpha_m{}^2}{\theta} (\overline{A} - \xi \ln \rho) \tag{52}$$

$$\tilde{\sigma}_r = \tfrac{1}{2}\xi(k+2) \ln^2 \rho - (k+2)\overline{A} \ln \rho + \overline{A} + \overline{B} \tag{53}$$

$$\tilde{\sigma}_\theta = \tfrac{1}{2}\xi(k+2) \ln^2 \rho - [(k+2)\overline{A} - \xi] \ln \rho + \overline{B} \tag{54}$$

$$\tilde{\sigma}_e = \overline{A} - \xi \ln \rho \tag{55}$$

The four constants k, α_m, $\overline{A}$, $\overline{B}$ are determined from four stress boundary conditions. These conditions, motivated by the same arguments as in [2, 1], are the following:

1 The arc $\rho = \rho_0$ at the entry is the rigid/plastic interface where complete yielding of the material takes place

$$\tilde{\sigma}_e = 1 \quad \text{at} \quad \rho = \rho_0 \tag{56}$$

2 The loading parameter η defined as

$$\eta = \frac{t + t_b}{t - t_b} \tag{57}$$

is specified. Here t and t_b are the drawing tension and the amount of back-pull, respectively, given by

$$t = \tilde{\sigma}_r \quad \text{at} \quad \rho = 1 \tag{58}$$

$$t_b = \tilde{\sigma}_r \quad \text{at} \quad \rho = \rho_0 \tag{59}$$

The loading condition (57) is more conveniently handled in the form

$$(1 + \eta)\tilde{\sigma}_r(\rho = \rho_0) + (1 - \eta)\tilde{\sigma}_r(\rho = 1) = 0 \tag{60}$$

Note that pure drawing is obtained with $\eta = 1$.

3 The remaining two conditions incorporate the effective Coulomb friction coefficients, along the walls, into the solution. It is possible of course to use equations (33) as alternative boundary conditions that bring the friction effects into the solution. However, the Coulomb friction coefficient is a more sound physical quantity than the friction factor and the conditions used here are

$$\int_1^{\rho_0} \tilde{\tau}_{r\theta} \rho^d \rho = \mu_1 \int_1^{\rho_0} \tilde{\sigma}_\theta \rho^d \rho \quad \text{at} \quad \theta = \alpha_1 \tag{61a}$$

$$\int_1^{\rho_0} \tilde{\tau}_{r\theta} \rho d\rho = -\mu_2 \int_1^{\rho_0} \tilde{\sigma}_\theta \rho d\rho \quad \text{at} \quad \theta = \alpha_2 \tag{61b}$$

where μ_1, μ_2 are the respective average Coulomb friction coefficients.

Conditions (56), (60)–(61) provide the four relations necessary for the determination of k, α_m, $\overline{A}$, $\overline{B}$. Inserting the stresses (52)–(55) into these conditions gives, by a straightforward solution

$$\overline{A} = 1 + \xi \ln \rho_0 \tag{62a}$$

$$\overline{B} = (1 + \tfrac{1}{2}k)(1 + \eta)(1 + \tfrac{1}{2}\xi \ln \rho_0) \ln \rho_0 - (1 + \xi \ln \rho_0) \tag{62b}$$

$$\alpha_m{}^2 = \alpha_1\alpha_2 \left(\frac{\mu_1 + \mu_2}{\mu_{av}} - 1\right) \tag{63c}$$

$$k = \frac{2\mu_{av}}{(\alpha_2 - \alpha_1 + \mu_{av})Q - \mu_{av}} \tag{63d}$$

where

$$Q = \frac{\tfrac{1}{2}(\rho_0{}^2 - 1)(1 + \tfrac{1}{2}\xi) - \tfrac{1}{2}\xi \ln \rho_0}{[1 + \tfrac{1}{2}(1 - \eta)(\rho_0{}^2 - 1)](1 + \tfrac{1}{2}\xi \ln \rho_0) \ln \rho_0} \tag{64}$$

and μ_{av} is an averaged friction coefficient defined as

$$\mu_{av} = \frac{\mu_1\alpha_1 + \mu_2\alpha_2}{\alpha_1 + \alpha_2} \tag{65}$$

with $\mu_1 = \mu_2 = \mu$ we have from (65) that $\mu_{av} = \mu$ and

$$\alpha_m = \sqrt{\alpha_1\alpha_2} \tag{66}$$

The stresses (52)–(55) are now completely determined. The drawing tension, (58), is given by

$$t = (1 + \tfrac{1}{2}k)(1 + \eta)(1 + \tfrac{1}{2}\xi \ln \rho_0) \ln \rho_0 \tag{67}$$

or in a more compact form

$$t = (1 + \tfrac{1}{2}k)t_u \tag{68}$$

where

$$t_u = (1 + \eta)(1 + \tfrac{1}{2}\xi \ln \rho_0) \ln \rho_0 \tag{69}$$

is the drawing tension in a uniform (frictionless) flow field. Relation (69) is easily derived from (29) and (32) in conjunction with boundary conditions (56) and (60).

Discussion

The structure of the basic result of this study—the formula for the drawing tension (67)—is the same as the equivalent expression obtained in [1] for the process of wire drawing. This similarity is quite expected as the analysis in [1] is based on a restricted version of the general solution employed here. The main difference between the two results is in the expression for k. For wire drawing we have (equation (57) in [1])

$$k = \frac{2\mu}{(\alpha + \mu)\,Q - \mu} \tag{70}$$

instead of (63d). Here μ is the friction coefficient and α the semiangle of the die. The charts presented in [1] for pure drawing of wires ($\eta = 1$) are therefore applicable to pure drawing of tubes via the transformation

$$\alpha \to \alpha_2 - \alpha_1, \quad \mu \to \mu_{av} \tag{71}$$

Similar charts that include the effect of back-pull ($\eta > 1$) on tube drawing can be prepared without difficulty.

Turning now to the limits of validity that should be imposed on the results we have in the first place restriction (45) as well as the as-

sumption of small angles. We have seen already that these two requirements, when combined, lead to (47).

A different restriction results from the fact that our model cannot describe the actual conditions at the entry and exit of the die. The boundary conditions used here, (56) and (58)–(60), are clearly of an averaged nature. We may expect, however, that for sufficiently tapered working zones the inaccuracy caused by that approximation is small. This geometrical restriction means essentially that $(\alpha_2 - \alpha_1)$ should be much smaller thatn $(\rho_0 - 1)/\rho_0$—a situation usually met in practice.

There are two additional limits on the validity of the results. The first one is associated with the possibility of necking instability at the exit of the die. That limit can be stated as, [1],

$$t < \tfrac{1}{2}\xi \quad \text{for} \quad \xi \geq 2$$

and (72)

$$t < 1 \quad \text{for} \quad \xi \leq 2$$

No necking will occur as long as the drawing tension (67) does not violate (72).

The last limit is set in order to prevent separation between the tube and the faces of the dies. That phenomenon is due to begin at the exit from die when $\tilde{\sigma}_\theta(\rho = 1) = 0$ or, from (54) and (62*b*), when

$$1 + \frac{1}{2}k = \frac{1 + \xi \ln \rho_0}{(1 + \eta)\left(1 + \frac{1}{2}\xi \ln \rho_0\right) \ln \rho_0} \equiv Q_s \tag{73}$$

The drawing tension, (67), at separation is simply $(1 + \xi \ln \rho_0)$ so that the separation limit reads

$$t < 1 + \xi \ln \rho_0 \tag{74}$$

Put otherwise, we can combine (73) with (63*d*), eliminate k, and obtain the critical relation for separation

$$\alpha_2 - \alpha_1 = \left[\frac{Q_s}{(Q_s - 1)Q} - 1\right] \mu_{\text{av}} \tag{75}$$

The restriction imposed by (72) is stronger than the one imposed by (74) provided that $\xi \leq 2$. For higher hardening parameters, however, (72) is more restrictive than (74) only as long as

$$\rho_0 > \sqrt{e}\, e^{-1/\xi} \tag{76}$$

References

1 Durban, D., "Axially Symmetric Radial Flow of Rigid/Linear-Hardening Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, June 1979, pp. 322–328.

2 Durban, D., and Budiansky, B., "Plane-Strain Radial Flow of Plastic Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 26, 1979, pp. 303–324.

3 Shield, R. T., "Plastic Flow in a Converging Conical Channel," *Journal of the Mechanics and Physics of Solids*, Vol. 3, 1955, pp. 246–258.

4 Truesdell, C., "Two Measures of Vorticity," *Journal of Rational Mechanics and Analysis*, Vol. 2, 1953, pp. 173–217.

S. Takezono
Professor,
Department of Mechanical Engineering,
Kumamoto University,
Kumamoto 860, Japan
Mem. ASME

E. Nakamachi
Lecturer,
Department of Mechanical
and Electrical Engineering,
Yatsushiro National College of Technology,
Yatsushiro 866, Japan

T. Yamaguchi
Engineer,
Scientific Crime Laboratory,
Kumamoto Prefectural Police Headquarters,
Kumamoto 862, Japan

Elasto/Viscoplastic Analysis of Thin Circular Plates Under Large Strains and Large Deformations

In this paper the elasto/viscoplastic analysis of thin circular plates under large strains and large deformations is studied by the use of the finite-element method based on the membrane shell theory. As the constitutive relation for the materials Perzyna's equation which in the plastic range takes into account the viscosity of the material is employed. The criterion for yielding used in this analysis is the von Mises yield theory. The geometric nonlinearity is treated with incremental method and the solutions at any stage are obtained by summation of the incremental values. The experiments are carried out for the thin circular aluminum plates, and the variations of the deformations and the stresses with loading rate are analyzed. The elasto/viscoplastic solutions from the prediction method agree fairly well with the values experimentally determined for the circular plates. The method can be used to generate plasticity solutions in a simple manner when stationary conditions are reached.

Introduction

The analysis of mechanical behavior of machinery and structures in the postyield stress range have been almost based on the elasto-plastic theory, in which the material viscous effects are disregarded. But it is often observed in the experiments that the actual behavior of the material is evidently governed by the viscous effects even at room temperature after plastic state has been reached. Especially in the plastic forming the influence of viscous properties on deformation process may be significant.

The authors have analyzed the dynamic problem [1] and the quasi-static problems [2, 3] of axisymmetrical shells, employing the elasto/viscoplastic theory proposed by Perzyna [4], which takes viscous effect into account in a plastic range.

The present paper describes a finite-element analysis of quasi-static large strain[1] and large displacement response of thin elasto/viscoplastic circular plates subjected to hydraulic pressure on the basis of the membrane shell theory [5, 6]. This problem, for example, is practically concerned with the bulge forming of a thin plate.

Assuming the material is homogeneous, isotropic, and isotropic hardening, the authors employ the constitutive relation of Perzyna type [4], in which elastic strains are obtained from Hooke's law and the viscoplastic strain rates are related to the excess value above the static yield stress. The finite-element formulation is derived from the exact kinematic, incremental treatment, and denoted by the surface convected coordinate. The ring-plate element is used owing to the restriction of axisymmetrical deformation.

Up to date, there are few experimental investigations to observe the viscous effect on the deformation in a plastic range [7]. The authors have carried out the bulge experiment of the thin circular aluminum plates under constant increase rates of hydraulic pressure to compare with the numerical results. Displacements, strains, and stresses of the plate have been measured by Moiré-topography [8] and a grid method.

Incremental Finite-Element Formulation

Neglecting the change of thermal energy and the inertia force, the authors formulate the equation of motion for a finite element. Since they are concerned only with the thin plate, the plate may be assumed to behave approximately as a membrane.

This assumption imposes the geometrical restriction on the applicability of the theory. The authors have not examined rigorously the geometrical restriction. But it seems from the comparison between the experimental results and the calculated ones that the maximum ratio of thickness to diameter is about 1/100 [9].

A ring-plate element, which is currently in the state C displaced by $\mathbf{U}$ from the initial undeformed state C_0, is displaced into an unknown state $\overline{C}$ due to an incremental change of external condition, as shown in Fig. 1.

At the state C_0 a surface convected coordinate system θ^α ($\alpha = 1, 2$), where θ^1 is a coordinate axis of meridional direction and θ^2 of circumferential direction, is embedded in the middle surface [5, 9], and the third coordinate axis θ^3 perpendicular to the θ^1–θ^2 plane is also

[1] The maximum Green strain in the surface of the plate in numerical example mentioned later is about 50 percent.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, December, 1979; final revision, April, 1980.

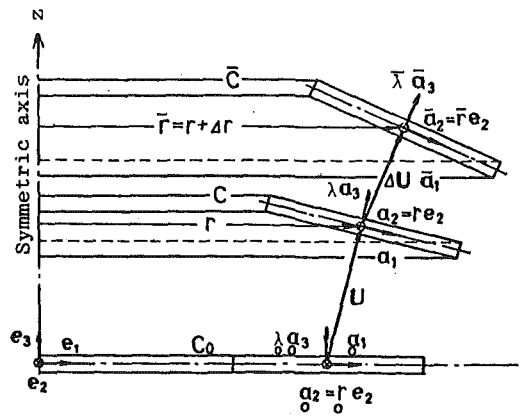


Fig. 1 Displacement of ring-plate element

fixed in the plate. The base vectors of this coordinate system at the center of gravity of the ring-plate element, denoted by ${}_0\mathbf{a}_i$, $\mathbf{a}_i$, $\bar{\mathbf{a}}_i$ (i = 1, 2, 3), corresponding to the state C_0, C, $\bar{C}$, respectively, have the following relation to the base vectors $\mathbf{e}_j$ (j = 1, 2, 3) in the fixed Cartesian coordinate system:

$$\mathbf{a}_\alpha = \Theta^i_\alpha \mathbf{e}_i, \quad \lambda\, \mathbf{a}_3 = \Theta^i_3 \mathbf{e}_i \tag{1}$$
$$\mathbf{a}^\alpha = \Omega^\alpha_j \mathbf{e}_j, \; \mathbf{a}^3/\lambda = \Omega^3_j \mathbf{e}_j$$

where $\mathbf{a}_3 = \mathbf{a}^3$, $|\mathbf{a}_3| = 1$, and $\lambda = h/h_0$, in which h_0 and h are the thickness of the plate in C_0 and C, respectively. Θ^i_j and Ω^i_j are obtained from the known nodal incremental displacements at each state. The authors follow the convention that Greek indices range over the values 1, 2 and Latin indices the values 1, 2, 3.

The natural base vectors defined by the following relations to the base vectors $\mathbf{a}_j$ are introduced:

$$\mathbf{G}_\alpha = \mathbf{a}_\alpha, \; \mathbf{G}_3 = \lambda \mathbf{a}_3 \tag{2}$$
$$\mathbf{G}^\alpha = \mathbf{a}^\alpha, \; \mathbf{G}^3 = \mathbf{a}^3/\lambda$$

In the incremental deformation from C to $\bar{C}$, a generic point P in the plate is displaced by $\Delta\mathbf{U}$, which is expressed by

$$\Delta\mathbf{U} = \Delta u^\alpha(\theta^\beta)\mathbf{a}_\alpha + \Delta u^3(\theta^\beta)\lambda\mathbf{a}_3 \tag{3}$$
$$= \Delta u_\alpha(\theta^\beta)\mathbf{a}^\alpha + \Delta u_3(\theta^\beta)\mathbf{a}^3/\lambda$$

where $\mathbf{a}^i$ is the contravariant base vector, $\mathbf{a}^\alpha \cdot \mathbf{a}_\beta = \delta^\alpha_\beta$ and $\mathbf{a}^3 = \mathbf{a}_3$.

The base vector $\bar{\mathbf{a}}_i$ in $\bar{C}$ is

$$\bar{\mathbf{a}}_\alpha = \bar{\mathbf{G}}_\alpha = (\delta^\nu_\alpha + \Delta u^\nu|_\alpha)\mathbf{a}_\nu + \Delta u^3|_\alpha \lambda \mathbf{a}_3 \tag{4}$$
$$\bar{\lambda}\bar{\mathbf{a}}_3 = \bar{\mathbf{G}}_3 = (1 + \Delta u^3|_3)\lambda \mathbf{a}_3 + \Delta u^\alpha|_3 \mathbf{a}_\alpha$$

where the symbol $|_i$ denotes the covariant differentiation with respect to θ^i in C.

The incremental displacement field in the finite element is considered to be given from the nodal values Δu^i_N by the Lagrangian interpolation function $\psi^N(\theta^\alpha)$ as follows:

$$\Delta u^i(\theta^\alpha) = \sum_N \psi^N(\theta^\alpha)\Delta u^i_N \quad (N = 1, 2) \tag{5}$$

where

$$\psi^N(\theta^\alpha) = \alpha^N + \beta^N_1\theta^1, \quad \alpha^N, \beta^N_1\text{: constants} \tag{6}$$

and the summation is taken over the nodes belonging to the element under consideration.

The Green strain increments defined at the center of gravity of the ring-plate element are expressed by

$$\Delta\gamma_{11} = \Delta u_1|_1 = \beta_{N1}\Delta u^N_1 = \beta^N_1 G_{11}\Delta u^1_N$$
$$\Delta\gamma_{22} = \Delta u_2|_2 = {}^N\textstyle\sum_{22i}\Delta u^i_N = {}_N\Pi^i_{22}\Delta u^N_i \tag{7}$$
$$\Delta\gamma_{ij} = 0 \quad (i \neq j)$$

in which ${}^N\sum_{22i} = r^N\Theta^1_i/2$, ${}_N\Pi^i_{22} = r_N\Omega^i_1/2$, i = 1, 3, where r^N and r_N are the radius of a circle through the nodal point N.

The strain increment $\Delta\gamma_{33}$ is obtained as follows:

$$\Delta\gamma_{33} = \lambda\Delta\lambda = \sum_{\alpha=1}^{2} A^{\alpha\alpha}_{33}\Delta\gamma_{\alpha\alpha} + P_{33}\Delta t \tag{8}$$

where Δt is the incremental time, and $A^{\alpha\alpha}_{33}$ and P_{33} are the coefficient tensors in the constitutive equation (26) of the materials.

The covariant differentiation of the incremental displacements is represent as

$$\Delta u_j|_\beta = {}_N\Psi^i_{j\beta}\Delta u^N_i, \; \Delta u_3|_3 = {}_N\Psi^i_{33}\Delta u^N_i + P_{33}\Delta t \tag{9}$$
$$\Delta u^j|_\beta = {}^N\Phi^i_{i\beta}\Delta u^i_N, \; \Delta u^3|_3 = {}^N\Phi^3_{i3}\Delta u^i_N + P_{33}/\lambda^2\Delta t$$

where

$$ {}_N\Psi^1_{11} = \beta_{N1}, \quad {}_N\Psi^1_{22} = {}_N\Pi^1_{22}$$
$$ {}_N\Psi^1_{33} = A^{11}_{33}\beta_{N1} + A^{22}_{33}{}_N\Pi^1_{22}, \quad {}_N\Psi^3_{31} = \beta_{N1}$$
$$ {}_N\Psi^3_{22} = {}_N\Pi^3_{22}, \quad {}_N\Psi^3_{33} = A^{22}_{33}{}_N\Pi^3_{22}$$
$$ {}^N\Phi^1_{11} = \beta^N_1, \quad {}^N\Phi^2_{12} = G^{22N}\textstyle\sum_{221}$$
$$ {}^N\Phi^3_{33} = G^{33}A^{22N}_{33}\textstyle\sum_{223}, \quad {}^N\Phi^3_{31} = \beta^N_1$$
$$ {}^N\Phi^3_{13} = G^{33}(G_{11}A^{11}_{33}\beta^N_1 + A^{22N}_{33}\textstyle\sum_{221})$$
$$ {}^N\Phi^2_{32} = G^{22N}\textstyle\sum_{223}$$

The other coefficients ${}_N\Psi^i_{jk}$, ${}^N\Phi^i_{jk}$ are vanished.

Let the stress vectors acting on the unit area in the states C and $\bar{C}$ whose unit normals are $\mathbf{n}$ and $\bar{\mathbf{n}}$, be denoted by $\mathbf{t}$ and $\bar{\mathbf{t}}$, respectively. The stress vector ${}_0\bar{\mathbf{t}}$ in $\bar{C}$ is similar to $\bar{\mathbf{t}}$, but measured per unit area in C. They are

$$\mathbf{t} = n_i\tau^{ij}\mathbf{G}_j = n_i s^{ij}\mathbf{G}_j \tag{10}$$
$$\bar{\mathbf{t}} = \bar{n}_i\bar{\tau}^{ij}\bar{\mathbf{G}}_j, \quad {}_0\bar{\mathbf{t}} = n_i s^{ij}\bar{\mathbf{G}}_j$$

where n_i and $\bar{n}_i$ are the components of the unit normal vectors in C and $\bar{C}$, respectively. τ^{ij} and s^{ij} are the true stress tensor and Kirchhoff stress tensor, respectively. $\bar{\tau}^{ij}$ and $\bar{s}^{ij}$ are expressed as follows:

$$\bar{\tau}^{ij} = \tau^{ij} + \Delta\tau^{ij}, \quad \bar{s}^{ij} = s^{ij} + \Delta s^{ij} \tag{11}$$
$$\Delta s^{ij} = \Delta\tau^{ij} + \tau^{ij}\Delta u^m|_m$$

The thin plate is assumed to be an axisymmetric membrane so that

$$\tau^{\alpha 3} = \Delta\tau^{\alpha 3} = \tau^{33} = \Delta\tau^{33} = \tau^{12} = \Delta\tau^{12} = 0 \tag{12}$$
$$s^{\alpha 3} = \Delta s^{\alpha 3} = s^{33} = \Delta s^{33} = s^{12} = \Delta s^{12} = 0$$

On the surface and the boundary, the stress vector and the surface force vector $\mathbf{T}$ are in equilibrium, namely,

$$\mathbf{t} = \mathbf{T} = T^i\mathbf{G}_i \quad (\text{in } C) \tag{13}$$
$${}_0\bar{\mathbf{t}} = {}_0\bar{T} = {}_0\bar{T}^i\,\bar{\mathbf{G}}_i, \quad {}_0\bar{T}^i = T^i + \Delta\hat{T}^i \quad (\text{in } \bar{C})$$

If the surface force is hydraulic pressure, it acts in the normal direction to the surface of the plate. In this case $\Delta\hat{T}^i$ is related to the increment of pressure Δp by

$$\Delta\hat{T}^i = \Delta p n^i + p\Delta u^m|_m n^i - p\Delta u^m|_l G_{mk}G^{li}n^k \tag{14}$$

Since the components of the unit normal vector $\mathbf{n}$ are $n^\alpha = 0$, $n^3 = 1/\lambda$, equation (14) becomes

$$\Delta\hat{T}^1 = -p/\lambda\beta^N_1\lambda^2 G^{11}\Delta u^3_N, \quad \Delta\hat{T}^2 = 0 \tag{15}$$
$$\Delta\hat{T}^3 = \Delta p/\lambda + p/\lambda\{(\beta^N_1 + G^{22N}\textstyle\sum_{221})\,\Delta u^1_N + G^{22N}\textstyle\sum_{223}\Delta u^3_N\}$$

Now the velocity vectors in C and $\bar{C}$ are denoted as

$$\mathbf{v} = v^i\,\mathbf{G}_i, \quad \bar{\mathbf{v}} = \bar{v}^i\,\bar{\mathbf{G}}_i \tag{16}$$

If the change of thermal energy, inertia forces, and body forces are neglected, the energy conservative equations in C and $\bar{C}$ are expressed as, respectively,

$$\int_V \tau^{ij}v_j|_i dV = \int_A T^i v_i dA, \tag{17}$$

$$\int_V (\bar{s}^{ij} + \tau^{im}\Delta u^j|_m)\bar{v}_j|_i dV = \int_A (T^i + \Delta\hat{T}^i)\bar{v}_i dA \quad (17)$$
(Cont.)

where V is the volume and A is the surface area.

The foregoing energy conservative equations give the equation of motion of the finite element in the following incremental form:

$$(K^{iM}_{jN} + {}^{K}_{(G)}{}^{iM}_{jN} + {}^{K}_{(\sigma)}{}^{iM}_{jN})\,\Delta u^j_M = \Delta P^i_N + {}^{R}_{(S)}{}^{iM}_{jN}\,\Delta u^j_M + (Y^i_N - {}^{Y}_{(S)}{}^{i}_{N})\,\Delta t \quad (18)$$

In equation (18) K^{iM}_{jN} is the incremental stiffness matrix, ${}^{K}_{(\sigma)}{}^{iM}_{jN}$ is the initial-stress matrix, ${}^{K}_{(G)}{}^{iM}_{jN}$ is the initial-rotation matrix, ${}^{R}_{(S)}{}^{iM}_{jN}$ is the initial-load matrix [10], and ΔP^i_N is the incremental generalized nodal forces and $(Y^i_N - {}^{Y}_{(S)}{}^{i}_{N})\Delta t$ is the incremental generalized apparent forces due to viscosity. They are calculated by the following equations:

$$K^{iM}_{jN} = \sum_{\alpha=1}^{2}\sum_{\beta=1}^{2} E^{\alpha\alpha\beta\beta} G_{\beta\beta}{}^{M}\Phi^{\beta}_{j\beta N}\Psi^{i}_{\alpha\alpha}V_e$$

$${}^{K}_{(G)}{}^{iM}_{jN} = \sum_{\alpha=1}^{2} \tau^{\alpha\alpha M}\Phi^{m}_{jmN}\Psi^{i}_{\alpha\alpha}V_e$$

$${}^{K}_{(\sigma)}{}^{iM}_{jN} = \sum_{\alpha=1}^{2} \tau^{\alpha\alpha M}\Phi^{m}_{j\alpha N}\Psi^{i}_{m\alpha}V_e$$

$$\Delta P^i_N = \Delta p n_m G^{mi} A_N$$

$${}^{R}_{(S)}{}^{iM}_{jN} = p({}^{M}\Phi^{\delta}_{j\delta}n^i - {}^{M}\Phi^{3}_{j\delta}\lambda^2 G^{\delta i}n^3)A_N$$

$$(Y^i_N - {}^{Y}_{(S)}{}^{i}_{N})\Delta t = \left(\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2} P^{\alpha\alpha}_{\beta\beta}\tau^{\beta\beta}{}_N\Psi^{i}_{\alpha\alpha}V_e - \sum_{\alpha=1}^{2}\tau^{\alpha\alpha}/\lambda^2 P_{33N}\Psi^{i}_{\alpha\alpha}V_e\right)\Delta t$$

$$V_e = \sqrt{G}\int_{V_0} dV, \quad G = \det(G_{ij}), \quad A_N = \int_{A_e}\psi_N dA$$

where $i = 1, 3$, V_0 is the volume of the finite element in C_0, V_e and A_e are volume and area of the element in C, respectively, and $E^{\alpha\alpha\beta\beta}$ represents the material constants as shown in equation (25). The term ${}^{R}_{(S)}{}^{iM}_{jN}$ introduces nonsymmetry in the coefficient matrix associated with the incremental loading(time) approach. To solve the nonsymmetric set of equations the Cholesky method using full-band matrix is employed.

Applying equation (18) to all the elements and assembling them according to the compatibility and the balance conditions at each nodal point, the simultaneous equations of the nodal incremental displacements can be obtained, which represent approximately the mechanical behavior of the circular plate.

Constitutive Equation

The constitutive equation of the homogeneous and isotropic material in the axisymmetrical plane stress condition are discussed.

When the elastic part of deformation is not so large, the increment of the total strain $\Delta\boldsymbol{\gamma}$ can be considered consistently as sum of the increments of elastic strain $\Delta\boldsymbol{\gamma}^{(e)}$ and viscoplastic strain $\Delta\boldsymbol{\gamma}^{(vp)}$,

$$\Delta\boldsymbol{\gamma} = \Delta\boldsymbol{\gamma}^{(e)} + \Delta\boldsymbol{\gamma}^{(vp)} \quad (19)$$

The elastic strain increment is related to the stress increment by Hooke's law,

$${}_J\Delta\boldsymbol{\tau} = \mathbf{B}\Delta\boldsymbol{\gamma}^{(e)} \quad (20)$$

where

$${}_J\Delta\boldsymbol{\tau} = \begin{Bmatrix} {}_J\Delta\tau^{11} \\ {}_J\Delta\tau^{22} \end{Bmatrix}, \quad \Delta\boldsymbol{\gamma} = \begin{Bmatrix} \Delta\gamma_{11} \\ \Delta\gamma_{22} \end{Bmatrix}$$

and ${}_J\Delta\boldsymbol{\tau}$ represents the Jaumann stress increment [6]. The elastic constant matrix $\mathbf{B}$ is given by

$$\mathbf{B} = \frac{2\mu}{(1-\nu)}\begin{bmatrix} G^{11}G^{11} & \nu G^{22}G^{11} \\ \nu G^{22}G^{11} & G^{22}G^{22} \end{bmatrix}$$

where μ is the shear modulus and ν is Poisson's ratio.

Combining Von Mises yield condition and the isotropic hardening hypothesis, the viscoplastic strain increment during a time interval Δt is given by Perzyna as follows [4]:

$$\Delta\boldsymbol{\gamma}^{(vp)} = \gamma\langle\Phi(F/F_0)\rangle\frac{\partial F}{\partial\boldsymbol{\tau}}\Delta t \quad (21)$$

where $\Phi(\ )$ is a positive monotonically increasing function and F is a viscoplastic potential, for which $F = \sqrt{3J_2} - F_0$, $2J_2 = \tau'^{ij}\tau'_{ij}$ are chosen, and $F_0 = \sigma_s = f(\epsilon)$ is the elastoplastic stress-strain relation in the uniaxial tension test. $1/\gamma$ means the coefficient of the viscosity and notation $\langle\rangle$ means

$$\begin{aligned}\langle\Phi(F/F_0)\rangle &= 0 \qquad F \leqq 0 \\ &= F/F_0 \quad F > 0\end{aligned} \quad (22)$$

The complete equation giving the strain increment can be written as

$$\Delta\boldsymbol{\gamma} = \mathbf{B}^{-1}{}_J\Delta\boldsymbol{\tau} + \mathbf{V}\cdot\boldsymbol{\tau}\Delta t \quad (23)$$

where

$$\mathbf{V} = \gamma\langle\Phi(F/F_0)\rangle\frac{\sqrt{3}}{2\sqrt{J_2}}\begin{bmatrix} 2/3G_{11}G_{11} & -1/3G_{11}G_{22} \\ -1/3G_{11}G_{22} & 2/3G_{22}G_{22} \end{bmatrix}$$

The Oldroyd stress increment $\Delta\tau^{\alpha\beta}$ is related to ${}_J\Delta\tau^{\alpha\beta}$ by

$$\Delta\boldsymbol{\tau} = {}_J\Delta\boldsymbol{\tau} - \mathbf{L}\,\Delta\boldsymbol{\gamma} \quad (24)$$

where

$$\Delta\boldsymbol{\tau} = \begin{Bmatrix} \Delta\tau^{11} \\ \Delta\tau^{22} \end{Bmatrix}, \quad \mathbf{L} = \begin{bmatrix} 2\tau^{11}G^{11} & 0 \\ 0 & 2\tau^{22}G^{22} \end{bmatrix}$$

Now in order to substitute the constitutive equation into the energy conservative equations, it is convenient to rewrite equation (19) by the use of equations (23) and (24) as follows:

$$\begin{aligned}\Delta\boldsymbol{\tau} &= (\mathbf{B} - \mathbf{L})\,\Delta\boldsymbol{\gamma} - \mathbf{B}\cdot\mathbf{V}\cdot\boldsymbol{\tau}\,\Delta t \\ &= \mathbf{E}\,\Delta\boldsymbol{\gamma} - \mathbf{P}\cdot\boldsymbol{\tau}\,\Delta t\end{aligned} \quad (25)$$

where

$$\mathbf{B} - \mathbf{L} = \mathbf{E} = \begin{bmatrix} E^{1111} & E^{1122} \\ E^{2211} & E^{2222} \end{bmatrix}$$

$$\mathbf{B}\cdot\mathbf{V} = \mathbf{P} = \frac{2\mu}{1-\nu}\gamma\,\langle\Phi(F/F_0)\rangle\,\frac{\sqrt{3}}{2\sqrt{J_2}} \times \begin{bmatrix} 1/3(2-\nu) & 1/3(2\nu-1)G^{11}G_{22} \\ 1/3(2\nu-1)G_{11}G^{22} & 1/3(2-\nu) \end{bmatrix}$$

The strain component $\Delta\gamma_{33}$ is calculated by

$$\Delta\gamma_{33} = \sum_{\alpha=1}^{2} A^{\alpha\alpha}_{33}\Delta\gamma_{\alpha\alpha} + P_{33}\,\Delta t \quad (26)$$

where

$$A^{\alpha\alpha}_{33} = \sum_{\beta=1}^{2} -\nu/E\lambda^2 G_{\beta\beta}B^{\beta\beta\alpha\alpha} \quad (\alpha\text{: not summed})$$

$$P_{33} = \sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}\{\nu/E\lambda^2 G_{\beta\beta}P^{\beta\beta}_{\alpha\alpha}\tau^{\alpha\alpha} - \gamma\langle\Phi(F/F_0)\rangle\frac{\lambda^2}{2\sqrt{3J_2}}\cdot G_{\alpha\alpha}\tau^{\alpha\alpha}\}$$

Comparison Between Numerical Results and Experimental Ones

The aforementioned finite-element method is applied to the quasi-static bulge deformation of thin circular plate made of aluminum (Al 1100P-0), which shows strain rate sensitivity in a plastic range. The numerical results are compared with the experimental results.

Experimental Method. The authors have performed uniaxial tension test for the aluminum specimens by Instron-type testing machine to obtain the material constants needed for the calculations. The true stress-strain relations of the aluminum under constant strain rate are illustrated in Fig. 2, where σ_s is the static stress and ϵ^{vp} is the

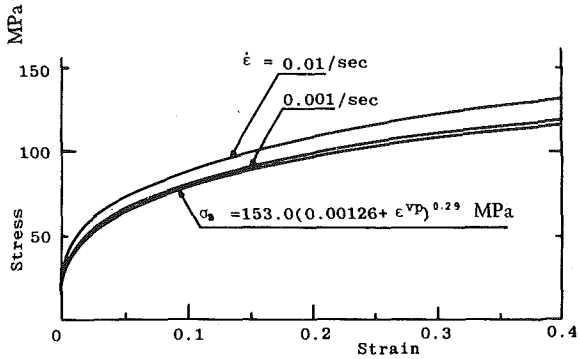


Fig. 2 Stress-strain relation

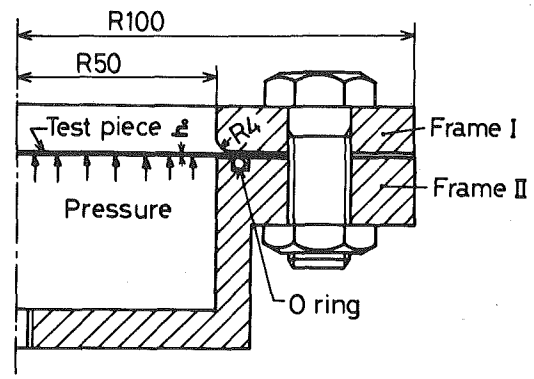


Fig. 3 Supporting frame

plastic strain. The stress-strain relation obtained under constant nominal strain rate 0.00005 (1/sec) has been regarded as the statical relation.

The dimensions of the circular plate test specimen for comparison between numerical results and experimental ones are 200 mm in diameter and $h_0 = 0.31$ mm in thickness as shown in Fig. 3.

The measurement of deformation, strain, and stress of the circular plate has been performed by a Moiré-topography [8] and a grid method without contact.

In Figs. 4 and 5, photograph of the experimental apparatus and its schematic view are shown, respectively. The Moiré-topography utilizes optical measurements to determine the deformed shape of the plate using contour lines.

As shown in Fig. 5, the parallel beams are attained by the use of a collimating lens and a field lens. If the shadow of an equispaced reference plane grating is projected onto the deformed surface of the plate and observed through the grating, contour lines showing equidifference of height Δz are obtained. Δz is

$$\Delta z = s_0 \tan \theta \tag{27}$$

where s_0 is an interval between the gratings and θ is an incident angle of the beam. In this experiment 0.5 and 1.0 mm have been used for s_0.

The comparison between the results of an almost spherical surface measured by Moiré-topography and the comparator[2] is shown in Fig. 6. At the steep slope the maximum difference is about 2 percent, but it may be concluded that the good agreement between them is obtained and the accuracy of Moiré-topography is recognized.

The relation between pressure p and deflection z_{max} at the center of the circular plate under a constant increase rate of hydraulic pressure 19.6 kPa/s (1/5 kg/(cm^2 s)) is shown in Fig. 7. The values of deflection measured by a displacement gage[3] are slightly smaller than ones by the Moiré-topography, because of the contact pressure of the displacement gage.

The variance of hydraulic pressure with time has been recorded automatically in the X-Y recorder through electric pressure transducer. Thus the constant increase rate of the pressure can be obtained, if the pressure-time lines prescribed on the X-Y recorder are followed by manual operation of a pressure pump.

The membrane displacements of the deformed plate can be measured by a reading microscope from the photographs of the grid points, which have been drawn on the surface of the undeformed plate. Thus the principal plane strain and curvature distributions can be obtained. From the condition of constant volume the thickness strain

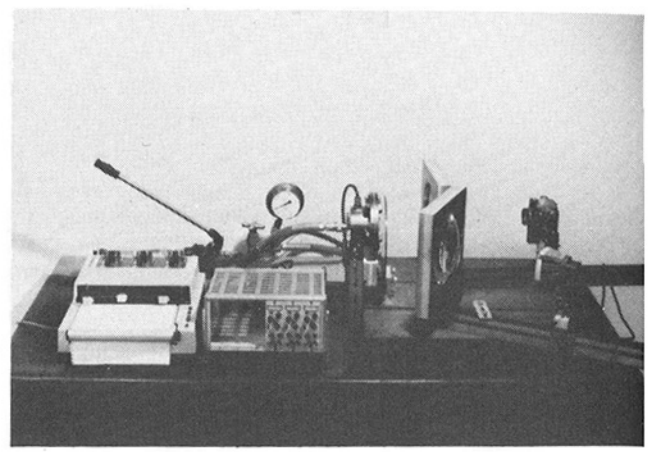


Fig. 4 Photograph of apparatus

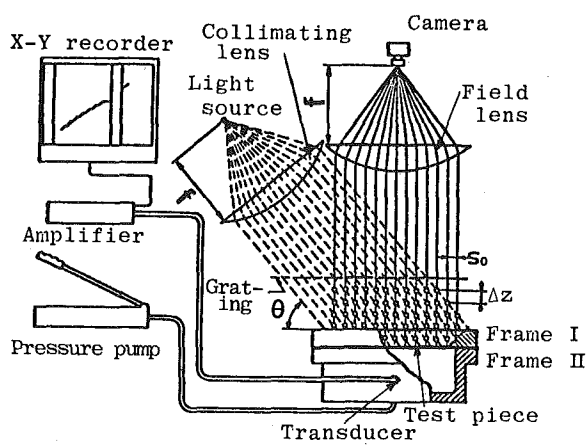


Fig. 5 Schematic view of experiment

also can be taken.

The meridional and circumferential stresses, σ_r and σ_θ, are calculated by

$$\sigma_r = pr/(2h \sin \phi) \tag{28}$$
$$\sigma_\theta = \{1 - r/(2\rho_r \sin \phi)\}\, p\rho_\theta/h$$

where $1/\rho_r$ and $1/\rho_\theta$ are the meridional and circumferential curvatures, and h, p, and ϕ are the thickness, hydraulic pressure and the slope of the tangent at the point of horizontal distance r.

[2] Microscopic type whose accuracy is 1/100 mm (Shimazu Seisakusho, Ltd.).

[3] Electric transducer type measurable up to 30 mm (Kyowa Electronic Instruments Co., Ltd.).

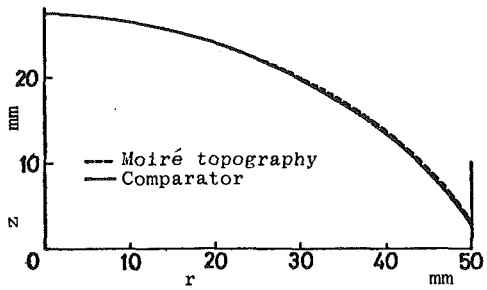


Fig. 6 Comparison between the measured values

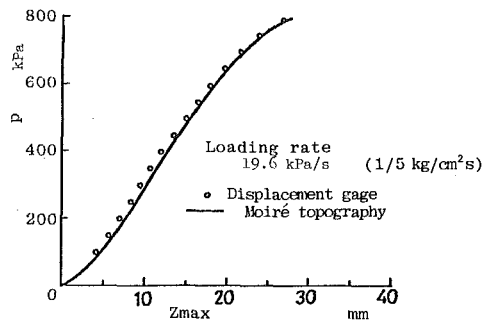


Fig. 7 Comparison between measured values of deflection at the center

Calculating Method. The circular plate has been divided into 10 ring-plate elements with equal width. This has been determined in consideration of the fact that the calculating results from 10 elements and 50 elements coincided well with each other except for the strain and the stress near the boundary in the previous elastoplastic analysis [9]. Because of the membrane approximation, the proper initial deformation at the start of calculation is needed. In this study the infinitesimal displacements $\mathbf{u}' = u'_0\mathbf{a}_1 + u'_0\mathbf{a}_3$ are assumed to be approximated as,

$$u'^1 = C_1 r(1 - r^2/R^2) \qquad (29)$$
$$u'^3 = C_2(1 - r^2/R^2)$$

where C_1 and C_2 are unknown constants, R is the radius of the circular plate, and r is the coordinate of radius.

Using the elastic stress σ'^{ij}-strain e'_{ij} relation the principle of stational energy of elastic membrane under pressure load p_0 has been expressed by Washizu [11] as follows:

$$\delta\Pi = 0, \quad \partial\Pi/\partial C_1 = 0, \quad \partial\Pi/\partial C_2 = 0 \qquad (30)$$
$$\Pi = \int_V \sigma'^{ij} e'_{ij} dV - \int_S p_0 u'^3 dS$$

Substituting (29) into (30), the values C_1 and C_2 are obtained by

$$C_1 = C_2{}^2 R^2 (3 - \nu)/4 \qquad (31)$$
$$C_2 = \{3(1-\nu)Rp_0(3-\nu)/G/h_0/2/(\nu^3 - 9\nu^2 + 11\nu + 11)\}^{1/3}/R^3$$

In this analysis the authors choose the appropriate value (about $h_0/50$) for C_2, then initial pressure load p_0 is derived from equation $(31)_2$. The initial displacement, strain and stress are also obtained.

Without introduction of this initial deflection, the stiffness in the direction $\mathbf{a}_3$ ($=\mathbf{e}_3$) becomes zero, and the deflection reaches infinity in calculation.

In the calculation the prescribed pressure-time curve is traced in a stepwise fashion, as shown in Fig. 8. That is, after an elastic analysis of the deformation to incremental pressure Δp, the response to total incremental time ΔT under the constant pressure is analyzed. The

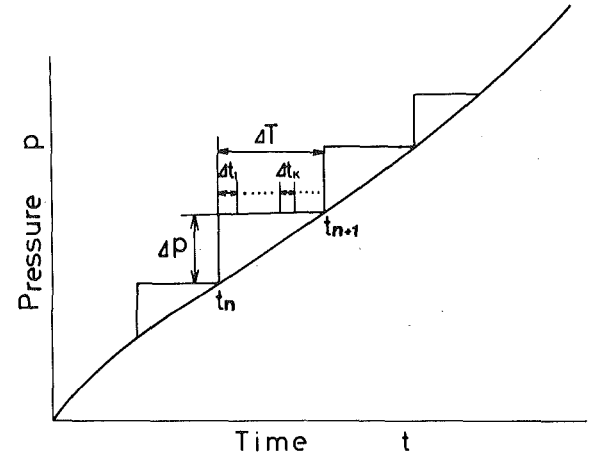


Fig. 8 Increments of pressure and time

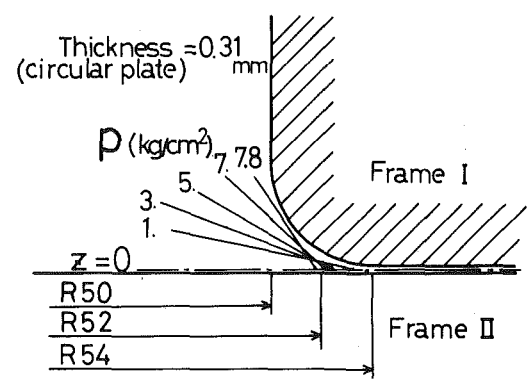


Fig. 9 Round corner of supporting frame I

time increment ΔT is obtained from summation of the subdivided time increment Δt_k,

$$\Delta T = \sum_k \Delta t_k$$

where the initial time increment Δt_1 has been determined to produce the viscoplastic strain increment of 1/120 elastic strain increment, and next time increments are increased by 50 percent, $\Delta t_k = 1.5\ \Delta t_{k-1}$, in consideration of the convergency of the solutions and computing time.

Comparison and Discussion. The calculation has been carried out for 104 mm in diameter in consideration of the round corner of the supporting frame I shown in Fig. 3, which is necessary to prevent the shear rupture at the boundary in small deformation range. It is seen from the locations of the intersections of $z = 0$ and the tangents of the deflection curves near the boundary obtained from Moiré-topography (Fig. 9) that the diameter 104 mm for calculation is reasonable through whole deformation process.

In Fig. 10, the relations between pressure and deflection at the center of the circular plate under constant increase rates of pressure, 19.6 and 1.6 kPa/s (1/5 and 1/60 kg/(cm² s)), are shown. From the comparison between the numerical results and the experimental ones, the values of the deflection in the experiment are slightly smaller than the values of the calculation. The results from the classical elastoplastic analysis [9] for the material constants in Table 1 and the numerical results under pressure rate 98.1 kPa/s (1/1 kg/(cm² s)) are also shown. Increasing of the pressure rate produces smaller deflections because of the corresponding increase in the material stiffness.

The deformed shapes measured by Moiré-topography in comparison with the calculated ones are shown in Fig. 11. The marks ○ and

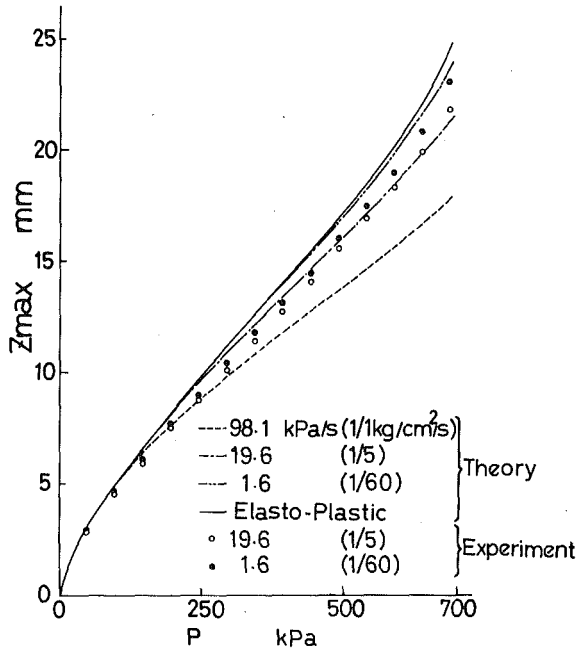


Fig. 10 Deflection at the center of the circular plate

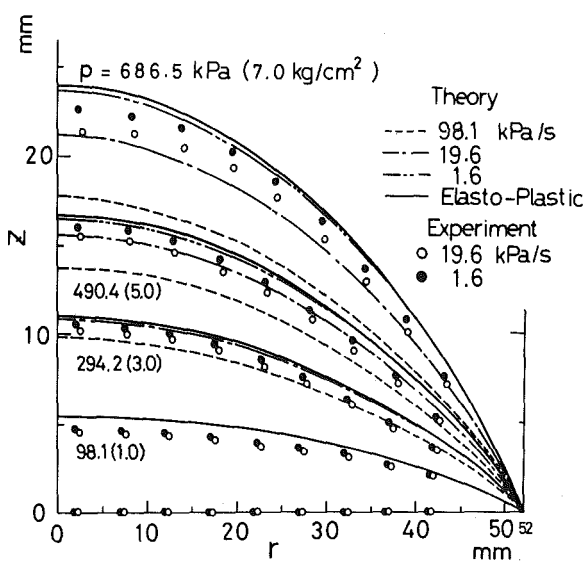


Fig. 11 Deformed shape

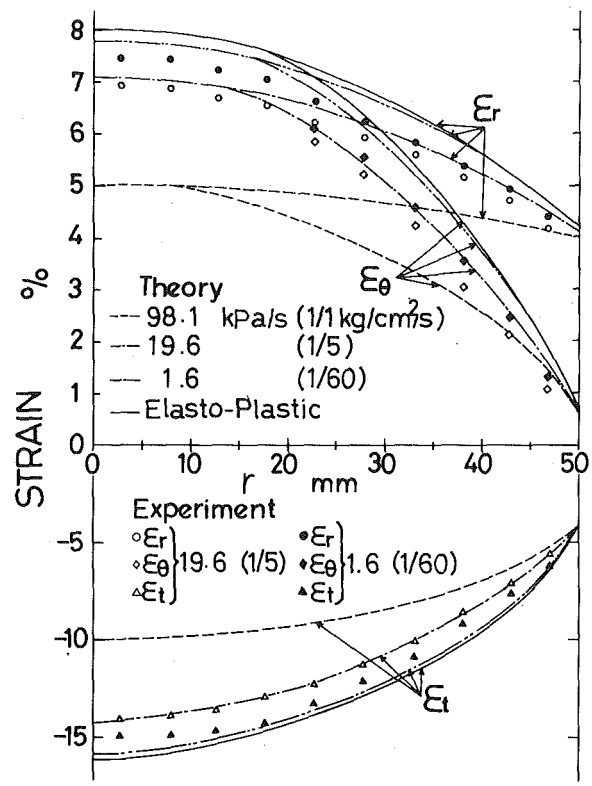


Fig. 12 Strain distribution P = 490 kPa (5 kg/cm²)

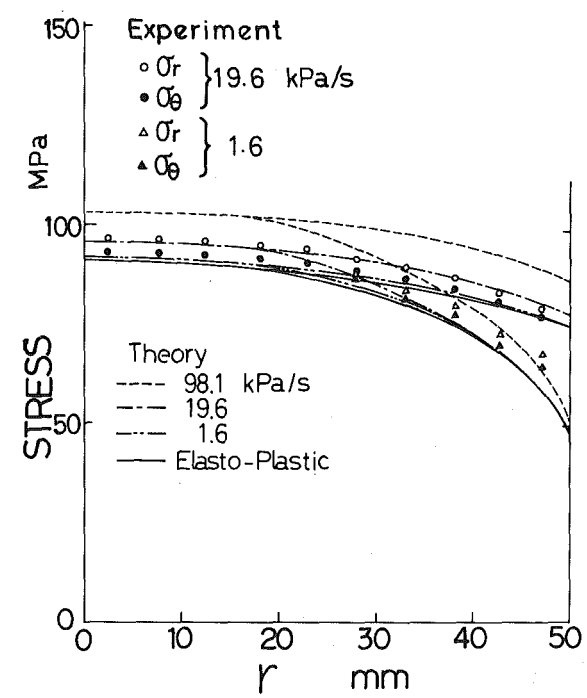


Fig. 13 Stress distribution P = 490 kPa (5 kg/cm²)

Table 1 Material properties

Young's modulus E = 68650. MPa	Yield stress σ_y = 22.073 MPa	Poisson's ratio ν = 0.314
Viscosity constant of material γ = 0.081 1/sec	Static true stress-strain relation $\sigma_s = 153.04(0.00126+\epsilon^{vp})^{0.29}$ MPa	

● indicate the material points of the deformed plate under pressure rates 19.6 and 1.6 kPa/s (1/5 and 1/60 kg/(cm² s)), respectively. From the both results of the calculation and the experiment, it is recognized that the deflection become smaller under the greater pressure rate.

The distributions of the strains and the stresses at pressure 490 kPa (5 kg/cm²) under pressure rates 98.1, 19.6, and 1.6 kPa/s (1/1, 1/5 and 1/60 kg/(cm² s)) and by elastoplastic theory are shown in Figs. 12 and 13, respectively. With increase of pressure rates the strains become small and the stresses become great.

It is seen from the calculation that as the pressure rate become lower, the results of the elasto/viscoplastic analysis become closer to ones of the elastoplastic analysis. It may be concluded from Figs. 10–13 that both results of experiment and calculation coincide fairly well in general.

Conclusion

The quasi-static large strains and large deformations of thin circular plate under hydraulic pressure have been analyzed by the incremental finite-element method, employing the membrane approximation and introducing the viscous effect of material in the plastic range.

Employing the constitutive relation of Perzyna type, the authors have formulated the incremental method with reference to the surface convected coordinate system, which is successful for such problem as including geometrical nonlinearity.

The results of calculation of the thin circular aluminum plate bulging under two constant rates of hydraulic pressure have agreed fairly well with the experimental ones in general. We have shown theoretically and experimentally the difference in deformation and stress distribution due to the difference between pressure rates.

The following matters have been recognized through the calculations and the experiments: increasing of the pressure rate produces smaller strains and deformations and larger stresses, and as the pressure rates become lower, the numerical results approach to the results from the classical elastoplastic analysis.

References

1 Takezono, S., and Tao, K., "Elasto/Viscoplastic Dynamic Response of Shells of Revolution," *Transactions of the 4th International Conference on Structural Mechanics in Reactor Technology,* Vol. M, 1977, M7/7.

2. Takezono, S., and Akashi, T., "Elasto/Viscoplastic Deformations of Thin Shells of Revolution," *Transactions of the 5th International Conference on Structural Mechanics in Reactor Technology,* Vol. M, 1979, M4/7.

3 Tao, K., and Takezono, S., "Elasto/Viscoplastic Analysis of Axisymmetrical Shells Under Asymmetrical Loading," Ibid., Vol. M, 1979, M4/8.

4 Perzyna, P., "Fundamental Problems in Viscoplasticity," *Advances in Applied Mechanics,* Academic Press, New York, Vol. 9, 1966, pp. 243–371.

5 Green, A. E., and Zerna, W., *Theoretical Elasticity,* Oxford University Press, London, 1968.

6 Kitagawa, H., and Tomita, Y., "An Incremental Finite-Element Analysis of Two-Dimensional Large Strain and Large Displacement Problems for Elastoplastic Material," *Proceedings of the 21st Japan National Congress for Applied Mechanics,* 1971, pp. 243–255.

7 Findlay, G. E., Moffat, D. G., and Stanley, P., "Torispherical Drumhead: A Limit-Pressure and Shakedown Investigation," *Journal of Strain Analysis,* Vol. 6, 1971, pp. 147–166.

8 Takasaki, H., "Moiré-Topography," *Applied Optics,* Vol. 9, 1970, pp. 1467–1472.

9 Kitagawa, H., Nakamachi, E., and Tomita, Y., "Static and Dynamic Analysis of Large Deflection of Elastic-Plastic Thin Plate," *Proceedings of the 6th North American Metalworking Conference,* 1978, pp. 236–243.

10 Oden, J. T., *Finite Elements of Nonlinear Continua,* Pergamon Press, Oxford, 1972.

11 Washizu, K., *Variational Method in Elasticity and Plasticity,* Pergamon Press, Oxford, 1968.

B. J. Hsieh

Components Technology Division,
Argonne National Laboratory,
Argonne, Ill. 60439
Mem. ASME

On the Uniqueness and Stability of Endochronic Theory

Several numerical and analytical analyses are described which evaluate the uniqueness and stability of solutions of mechanical models whose material behaviors are governed by the endochronic theory of plasticity. It has been found that the simplest form of this theory does show some "material instability" in the sense it does not satisfy Drucker's postulate when subjected to certain conditions. In other words, an endochronic material creeps under the action of applied force for dynamic problems. However, this "instability" or "lack of a hysteresis loop" can be circumvented by using more general forms of the endochronic theory when necessary. It is also shown that the endochronic solution is at least as unique as that of the elastoplastic theory. No numerical difficulties particular to this theory are observed even when the simplest form is used.

Introduction

The endochronic theory of material behaviors has been proposed by Valanis [1], and uses the principle that the history of deformation is defined in terms of a "time scale" which is not the real time, but is in itself a property of the material. No use of the classical yield surface concept is required in this theory. It is used by Valanis [2] to predict the mechanical response of aluminum and copper under conditions of complex strain histories. One constitutive equation described many phenomena, such as cross-hardening, loading and unloading loops, cyclic hardening as well as the effect of preshearing on tension behavior. Bazant and his coworkers further develop the theory to describe the liquefaction of sand [3] and the inelastic behavior and failure of concrete [4]. The use of a two-dimensional endochronic constitutive relation in dynamic transient analysis of shells is first considered by Lin [5].

Basically, the endochronic theory uses an "intrinsic time" in place of the real time in the viscoelastic constitutive equations. In other words, the time convolution integrals present in the viscoelasticity theory are replaced by the "intrinsic time" convolution integrals. Obviously, the accuracy of an endochronic model in describing the behavior of a real material is dependent on the form of the "relaxation" function, G, and on the definition of the intrinsic time. Important material phenomena may be lost when either is not defined properly. In such a case, the applicability of an endochronic model must be limited to a restricted class of problems.

Sandler [6] recently points out that the use of a simple endochronic model implies the material to be unstable and, hence, nonuniqueness of problem solutions can result. In this study, we show that the stability of a material is a matter of definition, and the stability of an endochronic model can be improved by using a more realistic "relaxation function." No nonuniqueness of solution exists. Furthermore, in a practical sense, unique and stable numerical solutions can be obtained for a structure whose material may violate Drucker's criterion of stability.

Stability of Materials

Drucker's stability criterion for materials as used in Sandler's paper is simply that no material is allowed to have negative material damping. The simple endochronic models used by both Sandler and Valanis, however, show negative damping when subjected to certain strain or stress histories. In order for a material to have positive damping, the constitutive law should be able to form a hysteresis loop upon a complete loading-unloading-reloading cycle. To show the inability of some endochronic models to form hysteresis loops, consider the simplest endochronic model used by Valanis [2] which is given by

$$E_0 d\epsilon = d\sigma + \sigma\alpha dz; \quad \alpha dz = \alpha_1\beta_1|d\epsilon|/(1+\beta\xi); \quad \beta d\xi = \beta_1|d\epsilon| \tag{1}$$

where σ and ϵ are the stress and the strain, E_0 is the initial slope of the one-dimensional stress-strain curve, α_1 and β_1 are material properties that are determined from a one-dimensional tension test, and z and ξ are measures of intrinsic time. If we define loading and unloading as when $d\epsilon > 0$ and when $d\epsilon < 0$, respectively, then we have

$$d\sigma/d\epsilon = E_0 \mp \alpha_1\beta_1\sigma/(1+\beta\xi); \quad \begin{matrix}(-)\text{ loading}\\(+)\text{ unloading}\end{matrix}. \tag{2}$$

For a hysteresis loop to be formed as shown in Fig. 1, the unloading and loading paths must intersect each other. Let the two points of

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, February, 1980; final revision, April, 1980. Paper No. 80-WA/APM-24.

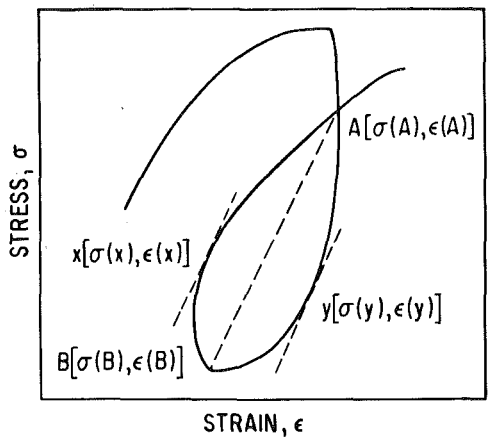


Fig. 1 Condition for forming a hysteresis loop

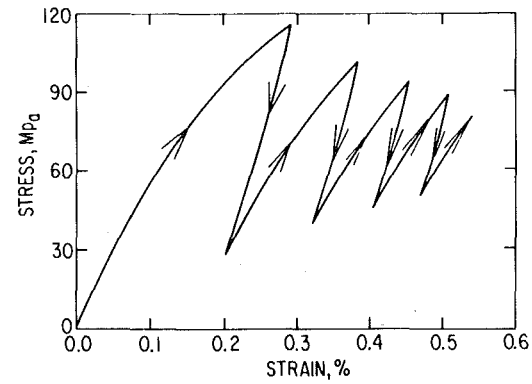


Fig. 2 "Unstable" endochronic stress-strain curve

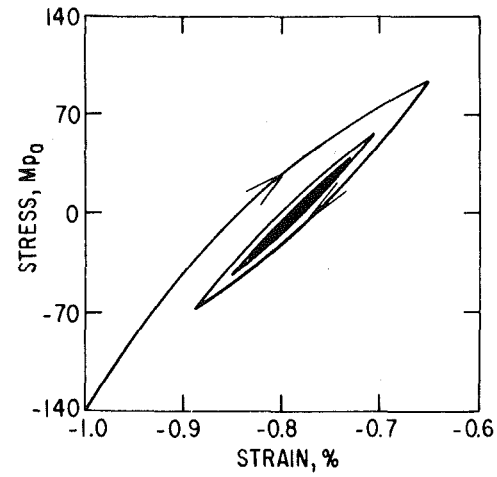


Fig. 3 Stable endochronic stress-strain curve with hysteresis loops

extreme stress value on the loop be denoted by A and B. The mean value theorem of calculus then implies that there exists points, x and y, whose coordinates are $(\sigma(x), \epsilon(x))$ and $(\sigma(y), \epsilon(y))$ on each side of the line AB, and their slopes are the same. By assuming the position for x and y and using equation (2) with appropriate signs, we find the following condition:

$$-\sigma(x)/[1 + \beta\xi(x)] = \sigma(y)/[1 + \beta\xi(y)] \tag{3}$$

for the formation of a hysteresis loop for the model depicted by equation (1). Since ξ is a monotonically increasing positive number, equation (3) indicates that $\sigma(x)$ and $\sigma(y)$ must be opposite in sign for a hysteresis loop formation. When the stress or strain history is such that a material always stays in either tension or compression, the endochronic constitutive law given by equation (1) may behave in an unstable manner in the sense described by Sandler.

Dynamic stress-strain curves for an endochronic model subjected to different strain histories are shown in Figs. 2 and 3 by using equation (1). These results are obtained by connecting a mass to a massless endochronic spring. The values for E_0, α_1, β_1 are 6.895×10^{10} Pa (1×10^7 psi) 49, and 8.75; the mass used is 4.378×10^{-2} Kg (3×10^{-3} slug). The unstable behavior shown in Fig. 2 is obtained by subjecting this spring-mass system to a suddenly applied constant force of magnitude 4.448×10^4 N (1×10^4 lb_f); while the stable hysteresis loops are obtained by subjecting the system to an initial displacement that is equivalent to 1 percent of strain in the spring. These results confirm the condition derived in equation (3) for the formation of hysteresis loops.

It should be noted that even though the "unstable" constitutive relation shown in Fig. 2 may not be a realistic representation of a real material behavior as we know it, it does not mean that such a material may never exist. However, we'll show that such unstable behavior of an endochronic model can be improved.

The endochronic model represented by equation (1) is obtained by assuming the intrinsic relaxation function G to be a single exponential function, i.e.,

$$G(z) = E_0 e^{-\alpha z}. \tag{4}$$

It is, therefore, reasonable to improve the accuracy of this constitutive model by including more terms in the relaxation function, say

$$G(z) = E_0 + E_1 e^{-\alpha z}. \tag{5}$$

The differential constitutive equation corresponding to the foregoing equation is simply

$$(E_0 + E_1)d\epsilon + E_0\epsilon(\alpha dz) = d\sigma + (\alpha dz)\sigma. \tag{6}$$

where E_0, E_1 are material parameters and αdz is defined in equation (1). The equation for the slope of this stress-strain curve is, therefore, given by

$$d\sigma/d\epsilon = (E_0 + E_1) \pm (E_0\epsilon - \sigma)\alpha_1\beta_1/(1 + \beta\xi); \quad \begin{matrix}(+)\text{ loading}\\(-)\text{ unloading}\end{matrix} \tag{7}$$

The necessary condition for a hysteresis loop can now be derived as in the previous case to be

$$[E_0\epsilon(x) - \sigma(x)]/[1 + \beta\xi(x)] = -[E_0\epsilon(y) - \sigma(y)]/[1 + \beta\xi(y)]. \tag{8}$$

From this, we see that there are possibilities for a hysteresis loop to be formed even for the cases where $\sigma(x)$ and $\sigma(y)$ have the same sign. A hysteresis loop for a copper specimen subjected to unloading and reloading in tension is excellently reproduced by using equation (6) in [2]. Fig. 4 shows the effect on the formation of hysteresis loops by the use of various ratio of E_1/E_0 for the same problem [7]. It is reasonable to assume that other types of loading-unloading-reloading may be described by using more general forms of the relaxation function.

It should be emphasized that no single constitutive law as yet can describe all the material behaviors observed in laboratories. Most constitutive laws are accurate in describing certain materials under certain conditions only. It is important to use appropriate constitutive laws for practical problems. However, the inability to represent a material in some situation by a constitutive law should not exclude its usefulness in representing the material in other situations. It is shown in [7,8] that the simple one-term relaxation function used in equation (1) can predict reasonably good results for many practical problems.

We reiterate that a constitutive law is simply a model used to approximate the behavior of a material. Certain models may represent

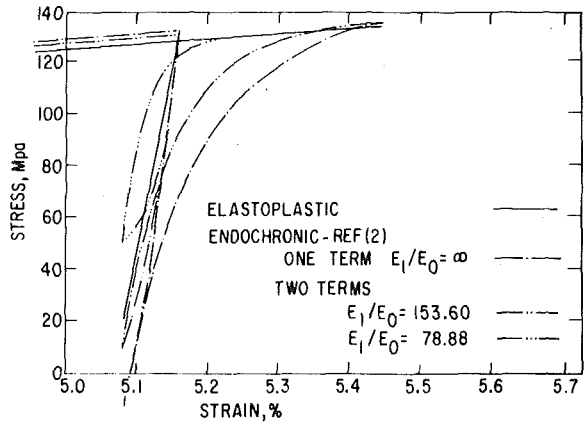


Fig. 4 Elastoplastic and endochronic uniaxial stress-strain curves for a copper specimen subjected to certain strain history. The endochronic curves are obtained by having the "relaxation function" $G(z)$ the following form: $G(z) = E_0 + E_1 e^{-\alpha z}$. The maximum and minimum values of E_1/E_0 consistent with the formulation in [2] are ∞ and 78.88 for this material, respectively.

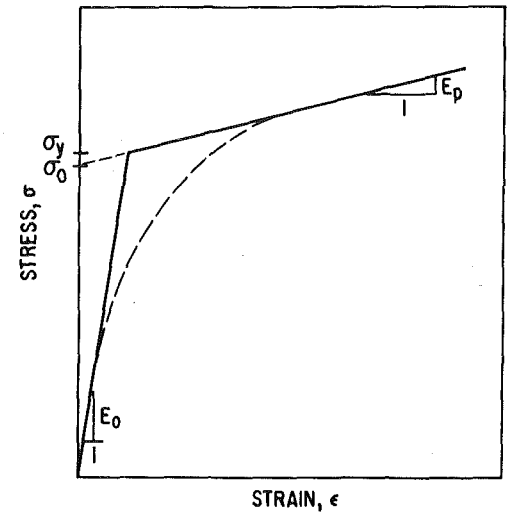


Fig. 5 One-dimensional endochronic and its associated bilinear stress-strain curve

some materials better than others in some situations and vice versa. The unstable behavior of the one-term endochronic model, hence, should be treated carefully when it occurs. The cause of this unstable behavior will be discussed later.

Uniqueness of Solution

The uniqueness of solution to a system of equations implies that one, and only one, set of results is obtained when the system is subjected to a set of initial/boundary conditions and external loadings. Nonuniqueness of solution exists if more than one set of results are possible when the system is subjected to the same situation. The question of uniqueness of an endochronic model has been discussed by Sandler [6] for both a spring-mass system and a continuum system. The spring-mass system in [6] is formed by slowly adding weights to a one-term endochronic spring and then subjecting this system to a small excitation; the continuum system is a prestressed rod initially at rest.

The spring-mass system is studied using various weights and initial disturbance, and the detailed results are available in [8]. The response of such a spring-mass system obviously depends on the magnitude of weight, direction of excitation, etc. The unstable response predicted in [6] is clearly observed in all results. In other words, the stresses oscillate about certain values while the strains drift as predicted in [6]. However, this drifting of displacements (strains) may not go unbounded within a finite time, since the response histories clearly indicate the vanishing of velocities of the mass as time increases. It is worth mentioning that the stress-strain curves for these results all have forms similar to that of Fig. 2.

These results are obtained by assuming that the system is in a gravitational field. That is the mass is obtained by dividing its weight by the gravitational constant and the mass is always subjected to a force caused by its own weight. To study the cause of the unstable behavior of the endochronic model more clearly, a constant mass is used for further investigation. Also, it is assumed that the only force available is through the external loading. This new system is subjected to various initial conditions or external loadings from its initially undisturbed state. The responses obtained are compared to those obtained by the bilinearly elastoplastic model that is related to the endochronic one, equation (1), by the following equation:

$$n = E_0/E_p; \quad \beta_1 = E_p/\sigma_0; \quad \alpha_1 = n - 1; \quad \sigma_y = \sigma_0/(1 - 1/n) \quad (9)$$

where E_0, E_p are the elastic and plastic moduli and σ_y is the initial yield stress of the bilinearly elastoplastic model. The relations between these constitutive models are shown in Fig. 5.

Practically identical results are obtained for the two models when the strains are extremely small. That is the endochronic model can indeed predict the linearly elastic response for practical purposes. Very reasonably good agreements are observed when the strains are large. The most severe discrepancies occur when the strain is intermediate particularly when it lies around the knee of the bilinearly elastoplastic stress-strain curve. The endochronic displacements are always greater than those of the elastoplastic one. The velocity of the mass has constant and decaying amplitude for the elastoplastic and endochronic one, respectively. From this, we see that the endochronic model is softer than the elastoplastic one when they are related by equation (9). This can be easily confirmed from the graphical representation of Fig. 5. It is also seen that the endochronic model has damping effect on the oscillation of the system.

The interesting phenomenon of the unstable behavior can be studied by examining the dynamic stress-strain curve of the simple spring-mass system. It is found that characteristically different behaviors are observed depending on whether the system is subjected to initial conditions or external loadings. Hysteresis loops similar to that shown in Fig. 3 are present in the stress-strain curves of those obtained by imposing an initial displacement or velocity to the spring-mass system. The orientation and aspect ratio of this loop depend on the magnitude of the strain at the center of the loop. No hysteresis loop is observed while the system is being subjected to the action of the external force. The resulting stress-strain curve shows a zigzag pattern as that shown in Fig. 2. Again the orientation and shape of this zigzag pattern depend on the magnitude of the strain. It seems reasonable to conclude that the unstable behavior is caused by the lack of hysteresis loop, and this only happens while the system is under the influence of an external force. This is very similar to the creep phenomenon of a viscoelastic material under the action of external force.

The next natural question is: Will the endochronic model restore to stable behavior when the action of external force is ceased? Figs. 6–8 show the displacement histories, the velocity histories and the dynamic stress-strain curves of the spring-mass system subjected to a rectangular pulse whose duration is six times the period of the corresponding linearly elastic system and whose magnitude is such that plastic deformation is obtained for the bilinearly elastoplastic system. The results for other pulse loadings are available in [8]. The arrows on the time axes of these figures indicate the instant the pulse ceases. These results clearly show the characteristic difference between the endochronic and elastoplastic models. They indicate that the endochronic model "creeps" under the action of external force, and it "stabilizes" immediately upon the removal of the external force. The dynamic stress-strain curve shows the zigzag pattern while the force is present and it forms hysteresis loops immediately upon the removal

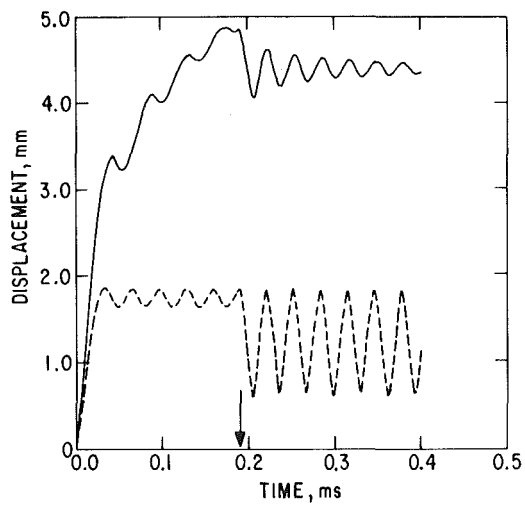


Fig. 6 Dynamic responses of a spring-mass system that is subjected to a rectangular pulse whose magnitude is 8.896 × 10⁴ N and whose duration is 6 *T*. Solid line—endochronic solution; chain-dashed line—elastoplastic solution.

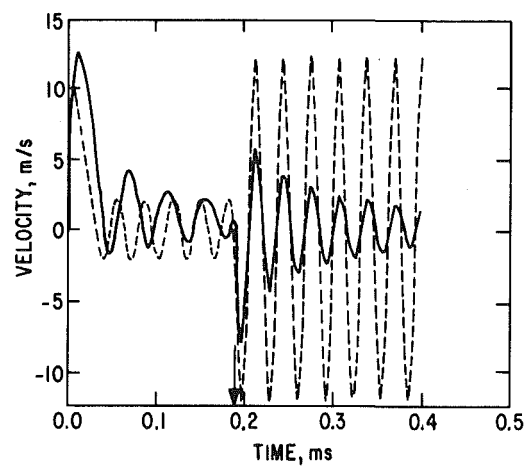


Fig. 7 Dynamic responses of a spring-mass system that is subjected to a rectangular pulse whose magnitude is 8.896 × 10⁴ N and whose duration is 6 *T*. Solid line—endochronic solution; chain-dashed line—elastoplastic solution.

of force. The damping effect on the oscillation of the endochronic model is clearly observed in the displacement and velocity histories in Figs. 6 and 7.

In all the study of the spring-mass system, unique solutions are obtained without the slightest numerical difficulty. Compared to the elastoplastic solutions, the only conclusion is that the endochronic solutions are different. This is not surprising, since the two constitutive models are different as shown in Fig. 5. It is important to note that the unstable behaviors shown in earlier examples exist only when external forces are present through either the use of a weight in a gravitational field or the application of an external load.

This unstable behavior is actually the creep phenomenon of the endochronic model with respect to its intrinsic time instead of real time, as can be seen by the similarity between endochronic and viscoelastic theories. Whether a creeping material should be termed as unstable is a question of definition; it does not present any difficulty in problem solving. It is interesting to mention that the same conclusion is observed by Bazant [9] from a different approach.

The possibility of nonuniqueness of solution to the spring-mass system proposed by Sandler [6] does not exist. He presents an interesting and useful model to uncover some undesired behavior of a

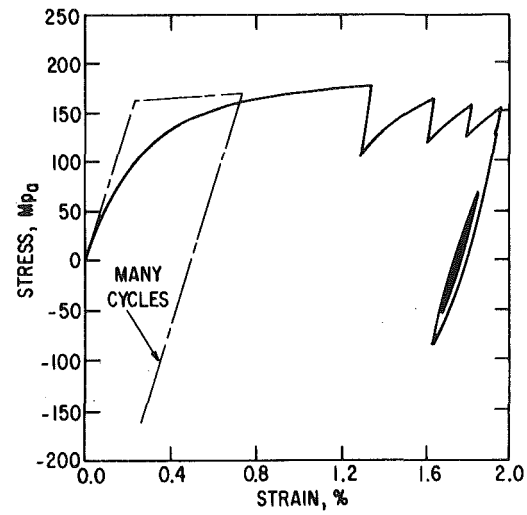


Fig. 8 Dynamic responses of a spring-mass system that is subjected to a rectangular pulse whose magnitude is 8.896 × 10⁴ N and whose duration is 6 *T*. Solid line—endochronic solution; chain-dashed line—elastoplastic solution.

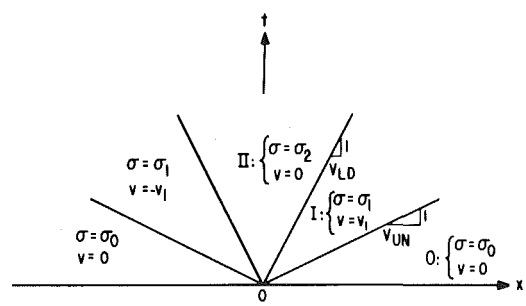


Fig. 9 Wave solution proposed in [6] for an endochronic rod

simple endochronic theory. However, the small disturbance used in his example and previous studies should not be interpreted as an error which expresses the difference between the numerical and analytical exact solutions. A computer does not know the real problem, it sees only a specific model and solves it with certain accuracy. Therefore, the exact solution is meaningless to a computer. A numerical solution is obtained in a computer by applying prescribed methods to the equilibrium equation or equation of motion. Hence, the numerical solution always satisfies the equilibrium equation exactly within the accuracy of the said computer. We reiterate that though the erroneous numerical solution may be different than the exact solution to a real problem it is the exact solution that satisfies equilibrium within the accuracy of a computer. This solution will be unique and the difference between this solution and the exact solution will not cause any unstable behavior even with the presence of external force, as confirmed by all the previous results.

As shown in Fig. 2, it is possible to have a situation where the stiffness of the unloading is greater than that of reloading for an endochronic model. In [6], Sandler considers a rod initially at rest and under a compressive stress σ_0. He proposes that alternative solutions can be constructed for the endochronic rod at least for short periods around a generic point 0. His solution is reproduced in Fig. 9. The stress and velocity in different regions are

$$\begin{aligned} &\sigma = \sigma_0; \quad v = 0 && \text{for Region 0,} \\ &\sigma = \sigma_1; \quad v = v_1 = -(\sigma_1 - \sigma_0)/(\rho V_{UN}) && \text{for Region I,} \\ &\sigma = \sigma_2; \quad v = 0 && \text{for Region II.} \end{aligned} \tag{10}$$

Here the following relation also holds:

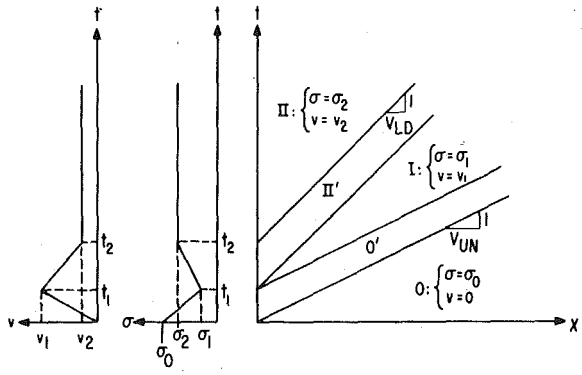


Fig. 10 Wave solution for a semi-infinitely long endochronic rod subjected to certain boundary condition

$$\sigma_2 = \sigma_0 + \rho(V_{UN} - V_{LD})v_1 \tag{11}$$

where ρ is the density of the rod, $V_{UN} = \sqrt{E_{UN}/\rho}$, $V_{LD} = \sqrt{E_{LD}/\rho}$; E_{UN} and E_{LD} are the slopes of the stress-strain curve on the unloading and reloading segments. It should be noted that E_{UN} and E_{LD} are assumed constants in deriving the aforementioned solutions. Sandler claims that equations (10) and (11), i.e., Fig. 9, is also solution of the rod subjected to the same conditions and loading as that of the trivial solution of $\sigma = \sigma_0$ and $v = 0$, and concludes that the endochronic rod may admit more than one solution.

As mentioned earlier, the nonuniqueness of solution can exist only when more than one solution may be obtained under identical initial and boundary conditions for the same governing equations. The solution presented in equations (10) and (11) is valid only before any wave arrives at the boundary ends. In other words, it is valid only for an infinitely long rod. Hence, we only have initial conditions and the governing equation to study. Equations (10) and (11) will confirm the existence of nonunique solution for the endochronic rod if they all reduce to the same initial condition and governing equation. However, we see that the initial conditions for these solutions are the same except at the point 0, from Fig. 9. Therefore, the question of uniqueness of solution can be resolved if we know whether the point 0, i.e., (0,0) on the $x - t$ plane is part of the solution, part of the initial condition or part of the external loading. In other words, this is a cause and effect problem.

To answer this question, let's first consider a semi-infinitely long endochronic rod that is initially at rest and subjected to a compressive stress σ_0. At the left boundary point 0, the stress is linearly reduced from σ_0 at time zero to σ_1 at time t_1 and then linearly increases to σ_2 at time t_2; and the stress stays at σ_2 after t_2. The boundary conditions and the solution are presented in Fig. 10. The stress and velocity in the various regions of Fig. 10 are as follows:

$$\sigma = \sigma_0;\quad v = 0 \qquad \text{for Region 0,}$$

$$\sigma = \sigma_0 + (\sigma_1 - \sigma_0)t/t_1;\quad v = -(\sigma_1 - \sigma_0)t/(\rho V_{UN} t_1) \qquad \text{for Region 0}',$$

$$\sigma = \sigma_1;\quad v = -(\sigma_1 - \sigma_0)/(\rho V_{UN}) \qquad \text{for Region I,}$$

$$\sigma = [(\sigma_2 - \sigma_1)t + (t_2\sigma_1 - t_1\sigma_2)](t_2 - t_1);$$

$$v = -\{[(\sigma_2 - \sigma_1)t + (t_2\sigma_1 - t_1\sigma_2)]/(t_2 - t_1) - \sigma_1\}/(\rho V_{LD}) - (\sigma_1 - \sigma_0)/(\rho V_{UN}) \qquad \text{for Region II}',$$

$$\sigma = \sigma_2;\quad v = -(\sigma_2 - \sigma_1)/(\rho V_{LD}) - (\sigma_1 - \sigma_0)/(\rho V_{UN}) \qquad \text{for Region II} \tag{12}$$

where the same symbols and assumption as in [6] are used. The solution for an elastoplastic rod subjected to the same boundary conditions can be obtained as a special case by letting $V_{UN} \to V_{LD}$ in equation (12).

Interesting phenomenon is observed for the endochronic one when we let the time t_1, and t_2 both go to zero. In this case, regions 0′ and II′ collapse into shock fronts, and the solution is composed of regions 0, I, and II only. The solution in these regions are still given by equation (12). Furthermore, if we assume that the stress σ_2 has the same value as σ_0, then equation (12) implies that the velocity at the boundary will have the following value:

$$v = (\sigma_1 - \sigma_0)(V_{UN} - V_{LD})/(\rho V_{UN} V_{LD}). \tag{13}$$

In other words, if the stress at the end of an endochronic rod is subjected to a disturbance which is not trivial, i.e., $\sigma_0 \to \sigma_1 \to \sigma_0$ where $\sigma_1 \neq \sigma_0$, in zero time the end will attain a velocity after the event. However, an elastoplastic one will not obtain a velocity when $V_{UN} = V_{LD}$. That is to say, the elastoplastic rod may have identical solution as the initial one if it is subjected to a process of unloading and reloading at its boundary point in zero time, while the endochronic one does not. The latter knows the disturbance, and a new solution corresponding to this disturbance will be generated. From this we conclude that when the boundary stress is a constant all the time and when the boundary stress is constant except at one instant in time are two different boundary conditions, at least for the endochronic rod. Alternatively speaking, we can say that a singular source point with a strength $\sigma_1 - \sigma_0$ exists at the point 0 at time 0.

Also, by choosing an appropriate value for σ_2, we can make the velocity in Region II vanish for an endochronic rod. This stress value can be obtained from equation (12) by substituting $v = 0$. It is given by

$$\sigma_2 = \sigma_0 + (\sigma_1 - \sigma_0)(1 - V_{LD}/V_{UN}). \tag{14}$$

That is, we can uniquely find a nontrivial stress σ_2 for each strength $\sigma_1 - \sigma_0$ of the singular source. Physically, if the end of the endochronic rod is brought to σ_1 and then to σ_2 that is given by equation (14) in zero time, this end will not have a velocity after this sequence of events. The work or energy input to the rod through the end stress variation is zero, because the end does not have a velocity. However, from equation (12), we see that Region I clearly has a velocity. The kinematic energy of this region is obtained from the strain energy of the rod. It is easy to show that the energy for the rod after this event is the same as that of the undisturbed rod. Since the end of the rod where the event occurs does not have a velocity, we can combine two identical rods where their connecting ends are subjected to the aforementioned sequence of events in zero time. A solution identical to the one proposed by Sandler in [6], i.e., Fig. 9, can thus be obtained.

From this, we would expect that Sandler's solution is a unique solution of the endochronic rod subjected to some disturbance. Since his solution is for an infinitely long rod, the disturbance can be applied to the rod only through the initial condition (i.e., inhomogeneous initial condition) or the external loading (i.e., inhomogeneous governing equation). If the solution by Sandler corresponds to both homogeneous initial condition and governing equation, then the endochronic rod indeed can not yield unique solution. Otherwise, the said solution is unique.

The governing equation for a rod vibrating axially is given by

$$C^2 \partial\epsilon/\partial x - \partial v/\partial t = p \tag{15}$$

where $\epsilon = \partial u/\partial x$ and $v = \partial u/\partial t$, u is the axial displacement, p is the external loading, x and t are the spatial and time coordinates, and $C = \sqrt{(d\sigma/d\epsilon)/\rho}$. The initial condition is

$$\begin{aligned} \sigma(x,0) &= f(x), \\ v(x,0) &= g(x) \end{aligned} \tag{16}$$

where $\sigma(x, t)$ and $v(x, t)$ are the stress and velocity at a point x for time t.

It is easily shown by substituting that $\sigma(x, t) = \sigma_0$ and $v(x, t) = 0$ satisfy equation (15) with $p = 0$ and $f(x) = \sigma_0$ and $g(x) = 0$. Hence, they represent a rod under static equilibrium subjected to the stress of σ_0. To show that the solution proposed by Sandler as shown in Fig. 9 is not a solution subjected to the same condition of $p = 0, f(x) = \sigma_0$ and $g(x) = 0$, we use the method of characteristics.

The characteristic lines and the associated characteristic conditions

for equation (15) can be derived following the usual method. They are as follows:

$$\pm\, \partial\sigma/\partial t \,|_{t1}^{t2} - \rho C(\partial v/\partial t)|_{t1}^{t2} = \int_{t_1}^{t_2} \rho C(\partial p/\partial t)dt \quad \text{along} \quad dx/dt = \pm C,$$

$$\pm\, \partial\sigma/\partial x \,|_{t1}^{t2} - \rho C(\partial v/\partial x)|_{t1}^{t2} = \int_{t_1}^{t_2} \rho C(\partial p/\partial x)dt \quad \text{along} \quad dx/dt = \pm C \tag{17}$$

where t_1 and t_2 are the projections on the time axis of two arbitrary points along the respective characteristic line. These characteristic conditions are derived by first taking partial spatial and time derivatives of equation (15). The loading p on any characteristic line can be obtained by

$$\int_{t_1}^{t_2} dp = \int_{t_1}^{t_2} (\partial p/\partial t)dt \pm \int_{t_1}^{t_2} C(\partial p/\partial x)dt \quad \text{along } dx/dt = \pm C. \tag{18}$$

Substituting equation (17) into equation (18), we have

$$p(t_2) - p(t_1) = \{[\pm\sigma_t(t_2) \mp \sigma_t(t_1)]/\rho C - [v_t(t_2) - v_t(t_1)]\} \pm C\{[\pm\, \sigma_x(t_2) \mp \sigma_x(t_1)]/\rho C - [v_x(t_2) - v_x(t_1)]\} \quad \text{along} \quad dx/dt = \pm C \tag{19}$$

where subscripts x and t denote partial derivatives with respective to x and t, respectively.

Equation (19) relates the external loading p to the solution σ_t, σ_x, v_t, and v_x along any characteristic line. The solution presented in Fig. 9 has zero values for these solutions except along four distinct characteristic lines, namely $dx/dt = \pm\, V_{LD}$ and $dx/dt = \pm\, V_{UN}$. Furthermore, these solutions are constants along these lines. Therefore equation (19) is satisfied for $p(t_2) = p(t_1) = 0$ for any t_2 and t_1 provided they are greater than time zero, say $t_2 > t_1 \geq 0^+$ where 0^+ is positive and infinitesimally close to zero. Hence, we see that whatever happened happens around time zero.

If we assume that the initial condition is homogeneous, i.e., $\sigma(x, t) = \sigma_0$, and $v(x, t) = 0$ for $t \leq 0$, then it follows that there exists a point $x = 0$ and $t = 0^-$ where 0^- is negative and infinitesimally close to zero such that $\sigma_t(0, 0^-) = \sigma_x(0, 0^-) = v_t(0, 0^-) = v_x(0, 0^-) = 0$. If we further assume that no load is applied before the time 0^- or $p(0, 0^-) = 0$; then by taking $t_2 = 0^+$ and $t_1 = 0^-$, equation (19) yields

$$\begin{aligned}
p(V_{UN}t, t) &= 2(\sigma_1 - \sigma_0)\delta(t)/\rho V_{UN} && \text{along} \quad dx/dt = V_{UN},\\
p(-V_{UN}, t) &= -2(\sigma_1 - \sigma_0)\delta(t)/\rho V_{UN} && \text{along} \quad dx/dt = -V_{UN},\\
p(V_{LD}t, t) &= 2(\sigma_2 - \sigma_1)\delta(t)/\rho V_{LD} && \text{along} \quad dx/dt = V_{LD},\\
p(-V_{LD}, t) &= -2(\sigma_2 - \sigma_1)\delta(t)/\sigma V_{LD} && \text{along} \quad dx/dt = -V_{LD}
\end{aligned} \tag{20}$$

where $\delta(t)$ is the delta function which has the property

$$\delta(t) = 0, \quad \text{if} \quad t \neq 0,$$

and

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

Equation (20) clearly indicates that the solution proposed by Sandler or Fig. 9 is a solution of a rod subjected to an external loading if we assume that the initial condition is homogeneous. Had we assumed that there is no external loading, we can also prove by using equation (19) that the initial condition cannot be also homogeneous as $\sigma(x, 0) = \sigma_0$ and $v(x, 0) = 0$. In this case, the point $x = 0$ and $t = 0$ is a singular point and multiple initial values must be allowed for this point in order to admit the original solution.

It should be noted that the delta functions in equation (20) are assumed to be pulses at time zero. This in term requires that the point $x = 0$ and $t = 0$ to permit multiple values. In this case, equation (20) can be interpreted as an explosion and an implosion occurring at $x = 0$, $t = 0$. Hence, Fig. 9 represents the solution of an endochronic rod initially at rest and under stress σ_0, and then is subjected to an explosion and an implosion at time zero at a generic point. However, it is not necessary to require that all the delta functions occur at one instant. The explosion and the implosion can occur in sequence in a very short time, and the order of this sequence depends on the sign of the initial stress σ_0.

This analysis is based on the characteristic conditions for the derivatives of the stress. This is important for problems that involve material which has different properties depending on the sign of $d\sigma = \sigma_t dt + \sigma_x dx$. Since an endochronic material does have this property it is concluded that an endochronic rod under initial static stress σ_0 will not go into motion unless it is disturbed either through boundary condition or external loading.

From the foregoing analysis, we have shown that an endochronic rod is stable under static stress. The solution presented by Sandler is indeed the response of the rod when subjected to some specific external loadings (or initial/boundary conditions). The magnitudes of these loadings are infinite as indicated by the delta functions of equation (20). This behavior of an endochronic material is not so surprising if we keep in mind that any deformation is treated as irreversible in this theory. As a matter of fact, similar situations could happen for a classical elastoplastic material if all other conditions are right. For instance, nontrivial solution can be obtained for a bilinearly elastoplastic rod which is subjected to a delta function of impulsive loading at one of its ends. The solution in this case, however, is not simple waves [10]. This is because the irreversible effect of the loading process, e.g., plastic strain, cannot be eliminated by the unloading process. Therefore, situations similar to Fig. 9 are not unique to the endochronic material.

The error analysis by Sandler which results in the conclusion that a numerical method applied to an endochronic material can not be accurate for dynamic problems is doubtful, since his analysis is based on the incorrect interpretation of Fig. 9. In other words, he misinterpreted Fig. 9 as the alternative solution of the rod subjected to homogeneous initial condition and no external loading. Therefore, he claims the numerical error corresponding to Fig. 9 will be everpresent in a model which uses endochronic material since the initialization of responses such as Fig. 9 does not need the presence of other numerical error (or external loading) at point 0. However, we believe this type of error will not occur in any numerical method even if Sandler's interpretation were correct. This is because the existence of solution such as Fig. 9 is caused by the cause/effect occurring in zero time, but any numerical method must use a finite time increment for integration. Hundreds of time steps are used to study the plate impact and solid penetration problems by using both endochronic and elastoplastic theories [8]. Reasonable agreements are obtained in all the problems for both theories. In the high velocity penetration of solids, it is also noted that the endochronic formulation requires less computation effort. Furthermore, the solution is much less sensitive to the time integration increment compared to the elastoplastic solution. It is observed that the use of endochronic material does not generate any additional numerical error.

We have shown in this analysis that Fig. 9 is a unique solution of an endochronic rod subjected to an external loading. This should not be interpreted as saying that the uniqueness theorem for a nonlinear equation for an endochronic rod is proven in this analysis.

Conclusion

The endochronic theory of describing material behavior is still in its infancy. One and multidimensional applications of this theory are examined by comparing the results to those of the classical theories. It is found that this theory may be characteristically different than the bilinearly elastoplastic theory in some situations when only one term is used in the endochronic relaxation function. In this case, it is suggested that the theory be used only to represent materials that do not possess pronounced yield points such as copper, aluminum, etc. For materials that do have pronounced yield points such as structural mild steel, it is more appropriate to use the classical elas-

toplastic theory or the more complicated rather than the simple endochronic model.

Material behaviors can be better described in the simple endochronic theory by taking more terms in its relaxation function. However, more endochronic material parameters will be generated when this is done. More research efforts are needed in relating these parameters to the experimental results such as tension tests.

The endochronic theory can be used with no less confidence than the classical elastoplastic theory in numerical applications. No special numerical difficulties are associated with this theory. In some cases, this theory is more efficient and time-integration-increment-independent than the corresponding elastoplastic theory in numerical applications for dynamic problems. The decision to use this theory or not should be determined by the nature of the problem and material.

Acknowledgment

This work was performed under the auspices of the U.S. Department of Energy, Division of Environment Control Technology.

References

1 Valanis, K. C., "A Theory of Viscoplasticity Without a Yield Surface," *Archives of Mechanics, Archiwum Mechaniki Stosowanej,* Vol. 23, No. 4, Warszawa, 1971, pp. 517–533.

2 Valanis, K. C., "A Theory of Viscoplasticity Without a Yield Surface," *Archives of Mechanics, Archiwum Mechaniki Stosowanej,* Vol. 23, No. 4, Warszawa, 1971, pp. 535–551.

3 Bazant, Z. P., and Krizek, R. J., "Endochronic Constitutive Law for Liquifaction of Sand," *Journal of Engineering Mechanics Division,* ASCE, Vol. 102, No. EM4, 1976, pp. 225–238.

4 Bazant, Z. P., and Bhat, P., "Endochronic Theory of Inelasticity and Failure of Concrete," *Journal of Engineering Mechanics Division,* ASCE, Vol. 102, No. EM4, 1976, pp. 701–722.

5 Lin, H. C., "Dynamic Plastic Deformation of Axisymmetric Circular Cylindrical Shells," *Nuclear Engineering and Design,* Vol. 35, No. 2, 1975, pp. 283–293.

6 Sandler, I. S., "On the Uniqueness and Stability of Endochronic Theories of Material Behavior," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, June 1978, pp. 263–266.

7 Hsieh, B. J., "The Use of Endochronic Plasticity for Multidimensional Small and Large Strain Problems," Argonne National Laboratory Report, ANL-CT-79-19.

8 Hsieh, B. J., "On the Uniqueness and Stability of Endochronic Plasticity Theory," Argonne National Laboratory Report, ANL-CT-78-51.

9 Bazant, Z. P., "Endochronic Inelasticity and Incremental Plasticity," *International Journal of Solids and Structures,* Vol. 14, No. 9, pp. 691–714.

10 Cristescu, N., *Dynamic Plasticity,* North Holland Publishing Co., 1967, pp. 63.

U. W. Cho
Postdoctoral Fellow.

W. N. Findley
Professor of Engineering,
Fellow ASME

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

Creep and Creep Recovery of 304 Stainless Steel Under Combined Stress With a Representation by a Viscous-Viscoelastic Model

Creep and creep-recovery data of 304 stainless steel are reported for experiments under constant combined tension and torsion at 593°C (1100°F). The data were represented by a viscous-viscoelastic model in which the strain was resolved into five components—elastic, plastic (time-independent), viscoelastic (time-dependent recoverable), and viscous (time-dependent nonrecoverable) which has separate positive and negative components. The data are well represented by a power function of time for each time-dependent strain. By applying superposition to the creep-recovery data, the recoverable creep strain was separated from the nonrecoverable. The form of stress-dependence associated with a third-order multiple integral representation was employed for each strain component. The time-dependent recoverable and nonrecoverable strains had different nonlinear stress dependence; but, the time-independent plastic strain and time-dependent nonrecoverable strain had similar stress-dependence. A limiting stress below which creep was very small or negligible was found for both recoverable and nonrecoverable components as well as a yield limit. The limit for recoverable creep was substantially less than the limits for nonrecoverable creep and yielding. The results showed that the model and equations used in the analysis described quite well the creep and creep-recovery under the stress states tested.

Introduction

Type 304 stainless steel is one of the principal materials being considered for design of critical parts of modern power plants subjected to complex stress states. Previous experimental work in this area was reviewed in [1]. Preliminary experiments on 304 stainless steel under combined tension and torsion at about 593°C (1100°F) in the present program were reported in [2] and analyzed to define a creep surface in [3]. In the present paper creep and recovery[1] data are reported for pure tension, pure torsion, and combined tension and torsion of 304 stainless steel at 593°C (1100°F). The present analysis of these data employs an approach for separation of creep strain into several components by using recovery data measured at zero stress after complete unloading following creep under constant states of stress. The main point of view in this analysis is to define the recoverable component of the time-dependent creep strain and to determine the relative significance of the recoverable strain compared to the nonrecoverable strain. The constitutive equation developed for constant stress will be extended to variable stress states in work to be reported later.

Material and Specimens

Type 304 stainless steel, heat No. 9T2796, was supplied by Oak Ridge National Laboratory. The chemical composition in percent was 0.059 C, 1.26 Mn, 0.44 Si, 18.6 Cr, 9.5 Ni, 0.033 P, 0.015 S, 0.35 Mo, and 0.25 Cu. Three 12-foot bars numbered 16, 17, 20, 4.88 cm (1.92 in.) in diameter have been used to date for test specimens. The specimens were machined into thin-walled tubes of which the nominal outside diameter, wall thickness, and gage length were 2.540, 0.1524, and 10.16 cm (1.000, 0.060, and 4.00 in.), respectively. The rough machined specimens were heat treated in an argon-filled vertical tube at 1107°C (2025°F). They were heated for 30 min, then cooled exponentially by lowering them in the tube from the heated zone down over a cold plug. The resulting minimum cooling rate at 538°C (1000°F) was 4°C per

[1] Recovery as used in this paper refers to creep recovery.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, November, 1979; final revision, March, 1980. Paper No. 80-WA/APM-10.

sec (467°F per min). For a description of the method see [2] (1972). The heat-treated specimens were finish-machined by removing about 0.1524 cm (0.060 in.) from the inside and outside of the test section. After heat treating the hardness was Rockwell A 37 to 40 and the average grain size was ASTM-3. The tensile yield stress of 304 stainless steel was reported [4] to be about 48 MPa (7 ksi) at 649°C (1200°F), 55 MPa (8 ksi) at 538°C (1000°F), and 97 MPa (14 ksi) at room temperature. The average yield stress in tension obtained from the present creep tests at 593°C (1100°F) was 57.2 MPa (8.3 ksi).

Experimental Apparatus and Procedure

The combined tension and torsion machine used for these experiments was described in detail previously [5]. The required load for one-step loading was applied by manual control of a jack to lower the weights at the end of the lever. Strain was measured by a mechanical extensometer [5] whose sensitivity was 2×10^{-6} for axial strain and 3×10^{-6} for engineering shear strain.

Some modifications for high temperature use were made to the extensometer. Ferro-Tec conical points were used to establish the gage length. These were mounted on leaf springs of Elgiloy which were fastened to stainless steel rings at each end of the gage length. Invar rods were used to transfer the motion of the two gage point rings to a location below the specimen where a differential transformer indicated the axial strain and a pointer, drum, and microscope were used to measure the angle of twist (and thus the shearing strain).

The specimen was heated from the inside by a quartz-tube infrared lamp about 2 in. longer than the specimen and mounted on the specimen axis. In addition, heaters, which consisted of about six turns of Nichrome wire insulated with ceramic beads and wrapped around the specimen, were used at each end to account for heat loss to the enlarged ends of the specimen. The lamp and heaters were controlled separately by two thermac controllers. The proportion of heat supplied to each end heater was adjusted by Variacs.

A reflecting shield 10.2 cm (4 in.) in diameter was placed around the specimen area and supported from the top specimen grip. Initially the invar rods were located inside the shield. After Test 38, the location of the extensometer rods was changed from the inside to the outside of the reflecting tube. This was accomplished by using extension arms connected to the inner ring at each gage point and extending through holes in the reflecting tube. The outer ring was not used. Small shields and insulation were used to minimize the disturbance of the holes. These modifications were made so that the temperature of the invar extensometer rods would be within the range for which their thermal expansion was negligible. Even though the temperature of the specimen was constant the fluctuation of thermal expansion of the rods was too great when located within the reflecting tube.

The thermocouples first used for control were chromel-alumel. This was later changed to Platinel II and more recently changed to platinum-platinum 10 percent rhodium. The beaded end of the control thermocouple was first attached to the inner surface of the specimen. But this usually caused +5.6°C (+10°F) temperature drift during 100 h of creep and recovery. Many different positions of the control thermocouple were tried. One control thermocouple attached to the outside surface of the specimen near the center for control of the lamp and another attached just below the top endheater for control of both endheaters resulted in the least temperature drift [within ±0.6°C (±1°F) through the whole test].

The temperature was measured by eight thermocouples of 0.0254 cm (0.01 in.) in diameter spot welded onto both sides of the specimen along the gage length, using a Doric digital temperature indicator with a sensitivity of 0.06°C (0.1°F). Early tests employed chromel-alumel thermocouples for measurement. This was later changed to chromel-constantan. The temperature drift of the chromel-constantan thermocouple itself was found to be less than ±0.6°C (±1°F) for 500 hr at 593°C (1100°F). The apparent temperature gradient along the gage length was less than ±2.8°C (±5°F).

The specimen was allowed to soak at the test temperature of 593°C (1100°F) for about 20 hr prior to testing.

A final strain reading just before loading was taken as the zero point. The required load was applied by means of dead weights. One-step loading and unloading was performed by manual control of a jack in less than 20 sec. The time at which the weight was fully applied or released was taken as zero time. For torsion the stress and strain were computed at the middle of the wall thickness.

Creep and Recovery Data

Each of the constant stress states employed was chosen to lie on a Mises ellipse having one of the following effective stresses: 47.8, 68.9, 86.2, 103.4 MPa (6.928, 10, 12.5, 15 ksi). Proportional loading (and unloading) was employed.

For a given Mises stress level the relation between σ and τ is as follows:

$$\sigma^2 + 3\tau^2 = (\sigma_{\text{eff}})^2, \tag{1}$$

where σ and τ are the applied tension and shear stress, respectively, and σ_{eff} is the effective Mises stress.

In the present paper, 13 creep and recovery tests, performed at a temperature of 593°C (1100°F), were included in the analysis of the data. Tests 9, 12, 20, and 22 were taken from [2], for which the creep periods were more than 500 hr. All tests employed are listed in Table 1. For the tests whose creep periods were more than 120 h, the data considered were limited to the first 120 h for creep in order to minimize the effect of apparent temperature drift on the creep data and also the effect of aging.

For all tests of recovery, only the first 250 h of data after unloading were considered.

The recovery data for creep tests lasting as long as 1000 h were considered as well as those lasting 100 h since the duration of the creep period did not affect the time exponent for recovery, as discussed later.

Analysis of Data—A Viscous-Viscoelastic Model

With the experimental observation that only part of the instantaneous response and part of the time-dependent creep strain were recovered on complete unloading, the creep strain was separated into the five components

$$\epsilon = \epsilon^E + \epsilon^P + \epsilon^{VE} + \epsilon^V(\text{pos}) + \epsilon^V(\text{neg}), \tag{2}$$

where ϵ^E is an elastic strain, ϵ^P is a plastic strain (time-independent), ϵ^{VE} is a viscoelastic strain (time-dependent recoverable), and ϵ^V is a viscous strain (time-dependent nonrecoverable). Separate positive and negative parts of ϵ^V allow for creep in either sense and reversal of stress. This kind of viscous-viscoelastic model was employed satisfactorily in the analysis of creep of an aluminum alloy by Findley and Lai [6–8].

The creep data on the 304 stainless steel considered was represented quite well by

$$\epsilon_{ij} = \epsilon_{ij}{}^0 + \epsilon_{ij}{}^+ t^N, \tag{3}$$

where $\epsilon_{ij}{}^0$, $\epsilon_{ij}{}^+$ are functions of stress and N is a constant at constant temperature.

In order to determine the recoverable part of the time-dependent creep strain, ϵ^{VE}, it was assumed that ϵ^{VE} may have the same form as (3) and that a modified superposition equation [9] was applicable for the recoverable creep strain and for the recovery strain. Thus

$$\epsilon_{ij} = A_{ij} + \epsilon_{ij}{}^{+VE}\{t^{n_1} - (t - t_1)^{n_1}\}, \tag{4}$$

where t_1 is the creep period and A_{ij} is the nonrecoverable strain accumulated up to t_1. By assuming a similar form of (3) for the nonrecoverable time dependent creep strain ϵ^V the creep data may be separated by means of (2) as follows:

$$\epsilon_{ij} = \epsilon_{ij}{}^E + \epsilon_{ij}{}^P + \epsilon_{ij}{}^{+VE} t^{n_1} + \epsilon_{ij}{}^{+V} t^{n_2}. \tag{5}$$

The analysis of the data which follows shows that the foregoing assumptions were reasonable.

The elastic strain was determined from (5) by setting $t = t_1$ and considering that there was no new plastic strain on unloading as follows:

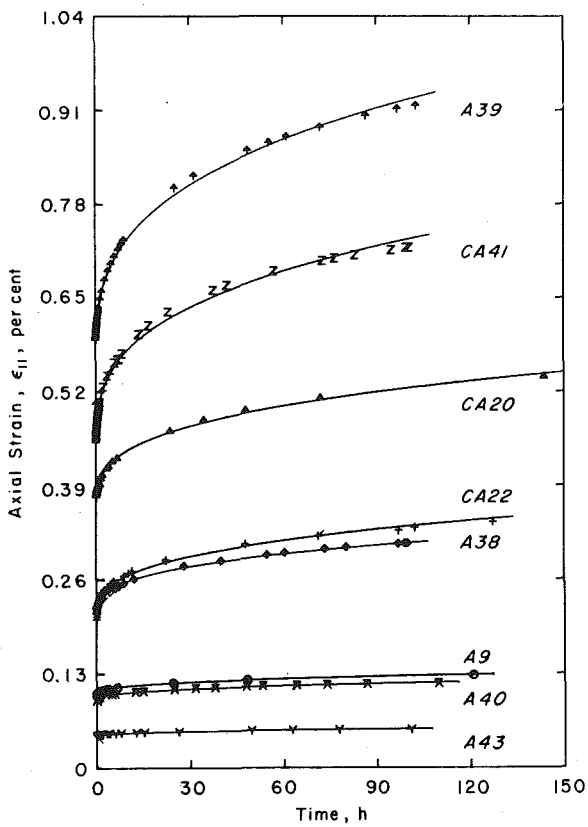


Fig. 1 Creep curves for axial strains under pure tension and combined tension and torsion. Numbers on curves indicate test numbers. Symbol *A* indicates tensile stress and *C* indicates combined tension and torsion. See Table 1 for stresses; solid lines are equation (5).

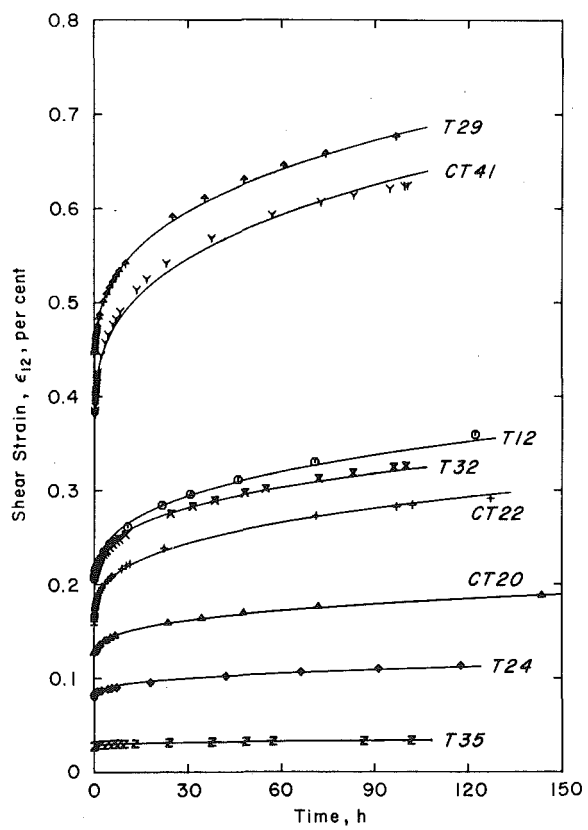


Fig. 2 Creep curves for shear strains under pure torsion and combined tension and torsion. Numbers on curves indicate test numbers. Symbol *T* indicates shearing stress and *C* indicates combined tension and torsion. See Table 1 for stresses; solid lines are equation (5).

Table 1 The values of N, n_1, and n_2 in equations (3), (4), and (7), the strain components for equation (5) with $N = n_1 = n_2 = 0.315$, and elastic moduli

Test	Tension, σ MPa	Tension, σ (ksi)	Torsion, τ MPa	Torsion, τ (ksi)	Duration of Creep t_1, h	Exponents N	Exponents n_1	Exponents n_2	Strain components with $N = n_1 = n_2 = 0.315$: ϵ^E, percent	ϵ^{+VE}, 10^{-2} percent/h^n	ϵ^{+V}, 10^{-2} percent/h^n	ϵ^P percent	$\epsilon^{+V}/\epsilon^{+VE}$	Elastic moduli 10^4MPa(10^6psi): From recovery test	From modulus test
A 9	68.9	(10)	0	(0)	989	0.296	0.548	0.200	0.0442	0.1983	0.4115	0.0561	2.1	15.6 (22.6)	* (*)
A 38	86.2	(12.5)	0	(0)	100	0.384	0.545	0.364	0.0596	0.2867	2.1026	0.1489	7.3	14.5 (21.0)	14.4 (20.9)
A 39	103.4	(15)	0	(0)	103	0.297	0.496	0.287	0.0693	0.4203	8.3686	0.4806	19.9	15.0 (21.7)	14.8 (21.4)
A 40	68.9	(10)	0	(0)	110	0.293	0.481	0.252	0.0456	0.1457	0.4400	0.0462	3.0	15.1 (21.9)	14.8 (21.5)
A 43	47.8	(6.928)	0	(0)	101	0.215	0.596	0.120	0.0325	0.0559	0.1795	0.0109	3.2	14.7 (21.3)	14.8 (21.5)
CA 20	79.4	(11.515)	19.4	(2.807)	912	0.337	0.488	0.322	0.0529	0.2972	3.5781	0.3064	12.0	15.2 (22.0)	14.8 (21.5)
CA 22	61.1	(8.857)	35.1	(5.093)	507	0.253	0.448	0.227	0.0409	0.2649	2.7920	0.1620	10.5	15.0 (21.7)	15.0 (21.7)
CA 41	73.3	(10.628)	42.1	(6.111)	100	0.277	0.525	0.266	0.0486	0.3297	6.6578	0.3847	20.2	15.1 (21.9)	* (*)
T 12	0	(0)	49.8	(7.217)	624	0.422	0.294	0.429	0.0415	0.2095	3.4050	0.1471	16.3	6.00(8.70)	5.41(7.85)
T 24	0	(0)	39.8	(5.773)	1010	0.415	0.317	0.447	0.0357	0.1390	0.5996	0.0427	4.3	5.58(8.09)	5.42(7.86)
T 29	0	(0)	59.7	(8.660)	100	0.344	0.362	0.346	0.0533	0.3136	5.9223	0.3614	18.9	5.60(8.12)	5.54(8.04)
T 32	0	(0)	49.8	(7.217)	100	0.430	0.375	0.439	0.0450	0.1844	2.8160	0.1490	15.3	5.53(8.02)	5.46(7.92)
T 35	0	(0)	27.6	(4.000)	102	0.301	0.165	0.301	0.0249	0.0296	0.1204	0.0023	4.1	5.54(8.03)	5.58(8.10)
CT 20	79.4	(11.515)	19.4	(2.807)	912	0.364	0.201	0.368	0.0178	0.0817	1.3524	0.1019	16.6	5.43(7.88)	5.40(7.83)
CT 22	61.1	(8.857)	35.1	(5.093)	507	0.279	0.327	0.278	0.0315	0.1645	2.9007	0.1229	17.6	5.57(8.08)	5.49(7.96)
CT 41	73.3	(10.628)	42.1	(6.111)	100	0.263	0.360	0.359	0.0390	0.2185	6.2618	0.3186	28.7	5.40(7.83)	* (*)

Note: * = no modulus test.

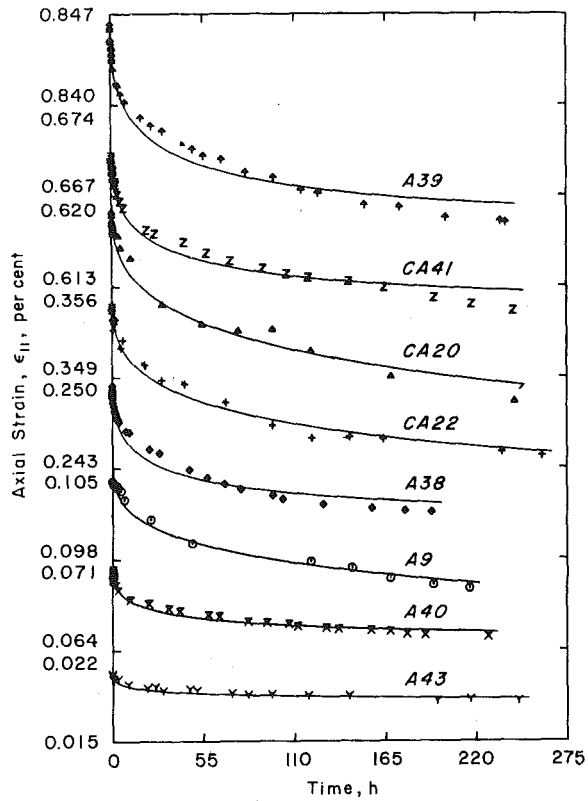


Fig. 3 Recovery curves for axial strains following pure tension and combined tension and torsion creep in Fig. 1. Dual strain values indicate that curves have been shifted. Numbers on curves indicate test numbers. Symbol *A* indicates tensile stress and *C* indicates combined tension and torsion. See Table 1 for stresses and duration of creep; solid lines are equation (4).

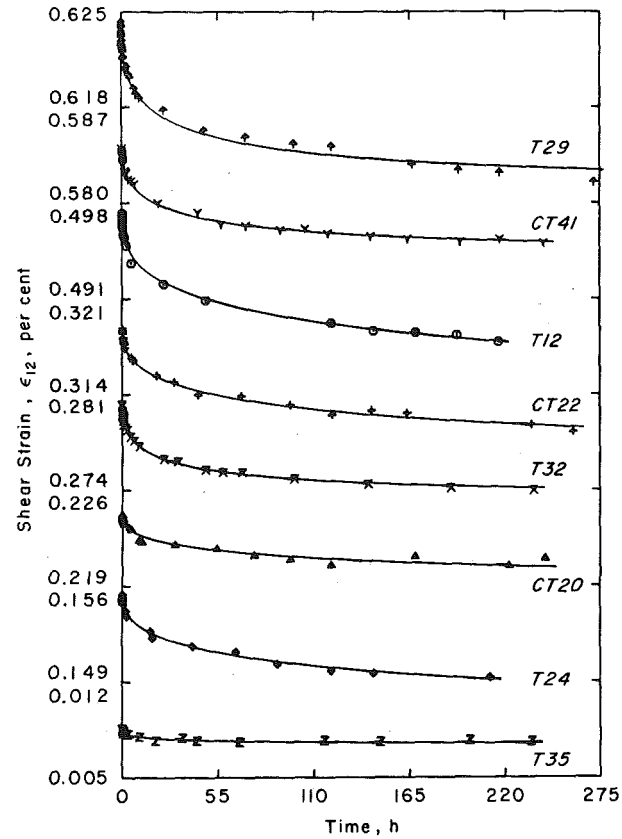


Fig. 4 Recovery curves for shear strains following pure torsion and combined tension and torsion creep in Fig. 2. Dual strain values indicate that curves have been shifted. Symbol *T* indicates shearing stress and *C* indicates combined tension and torsion. See Table 1 for stresses and duration of creep; solid lines indicate equation (4).

$$\epsilon_{ij}{}^{E} = \epsilon_{ij}|_{t=t_1} - A_{ij} - \epsilon_{ij}{}^{+VE} t_1{}^{n_1}, \tag{6}$$

where $A_{ij} = \epsilon_{ij}{}^{P} + \epsilon_{ij}{}^{+V} t_1{}^{n_2}$ were obtained from (4). Then ϵ^{P}, ϵ^{+V}, and n_2 were determined by use of the following rearrangement of (5) as applied to the creep data,

$$\epsilon_{ij} - (\epsilon_{ij}{}^{E} + \epsilon_{ij}{}^{+VE} t^{n_1}) = \epsilon_{ij}{}^{P} + \epsilon_{ij}{}^{+V} t^{n_2}. \tag{7}$$

The values of N, n_1, and n_2 were obtained from (3), (4), and (7), respectively, by applying least squares to the creep and recovery data. The results are shown in Table 1. The values of N, n_1, and n_2 for ϵ_{11} and ϵ_{12} were very close to each other, except that the value of n_1 for ϵ_{11} was larger than the other values of N, n_1, and n_2. This difference appears to be real material behavior, but no explanation has been found. Thus it seemed reasonable to consider $N = n_1 = n_2$.

The average ($n = 0.315$) of the N of ϵ_{11} and the N, n_1 of ϵ_{12} was used in the following analysis (with $N = n_1 = 0.315$ it follows that $n_2 = 0.315$). Using $n = 0.315$ the values of $\epsilon_{ij}{}^{E}$, $\epsilon_{ij}{}^{P}$, $\epsilon_{ij}{}^{+VE}$ and $\epsilon_{ij}{}^{+V}$ were recomputed from (4), (6), and (7) as shown in Table 1.

The creep and recovery curves calculated by using the values of the strain components given in Table 1 for $n = 0.315$ are shown in Figs. 1–4 by solid lines. The good agreement with the test data of Figs. 1, 2, and 4 indicates that the aforementioned assumptions were reasonable. For several recovery tests in tension the rate of recovery shown in Fig. 3 was more rapid than described by the theory because n_1 for tension was greater than 0.315.

The ratio $\epsilon^{+V}/\epsilon^{+VE}$ appeared to be independent of stress for an aluminum alloy [6]. As shown in Table 1 this ratio for 304 stainless steel was not independent of stress. It was nearly constant for the lower stress levels but depended strongly on the Mises stress level at the higher values.

Elastic Strain

The elastic strain was obtained from the recovery data by (6) and shown in Fig. 5 and Table 1. A modulus test was also performed by small incremental loadings and unloadings in the elastic range at 593°C (1100°F) just after completion of each recovery test. The elastic modulus was obtained from the incremental loading tests as the slope of a plot of stress versus strain.

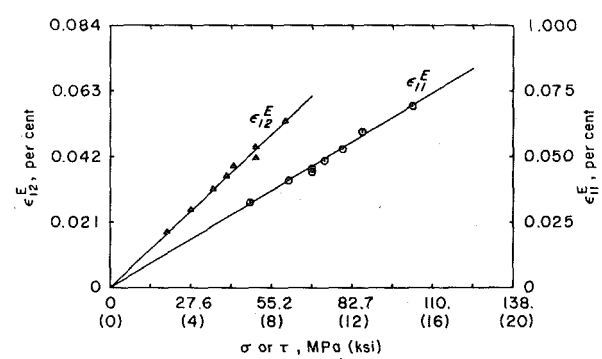


Fig. 5 Axial stress σ versus strain $\epsilon_{11}{}^{E}$ and shear stress τ versus strain $\epsilon_{12}{}^{E}$.

The moduli obtained by both methods are shown in Table 1 for tension E and shear G. It was observed that the results were quite consistent, but indicated a slightly smaller value for the moduli tests compared to those for the recovery tests in most instances. The fact that the modulus determined from both tests were so close together indicates that no plastic strain occurred on unloading. Poisson's ratio ν was calculated from the average value of all determinations of the elastic modulus E and the shear modulus G as

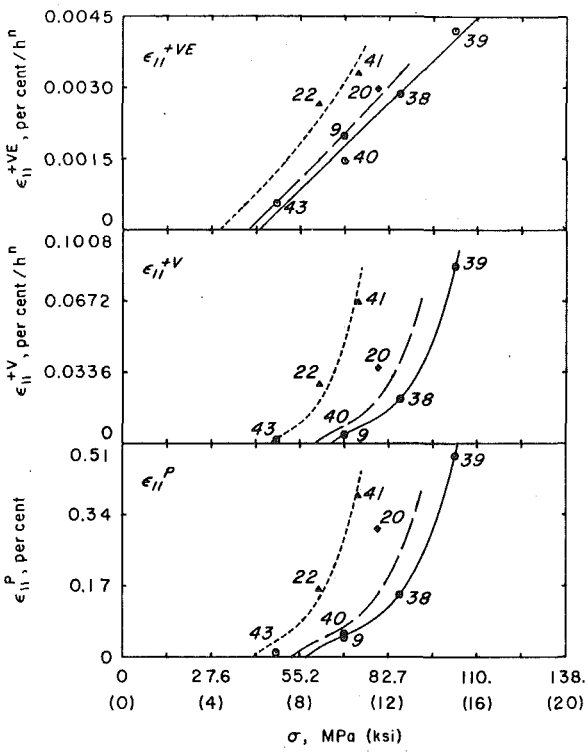


Fig. 6 Axial stress σ versus $\epsilon_{11}{}^{+VE}$, $\epsilon_{11}{}^{+V}$, and $\epsilon_{11}{}^{P}$. Numbers indicate test numbers in Table 1. Lines indicate equation (8) for $F_i{}^{VE}$, $F_i{}^{V}$, and $F_i{}^{P}$.

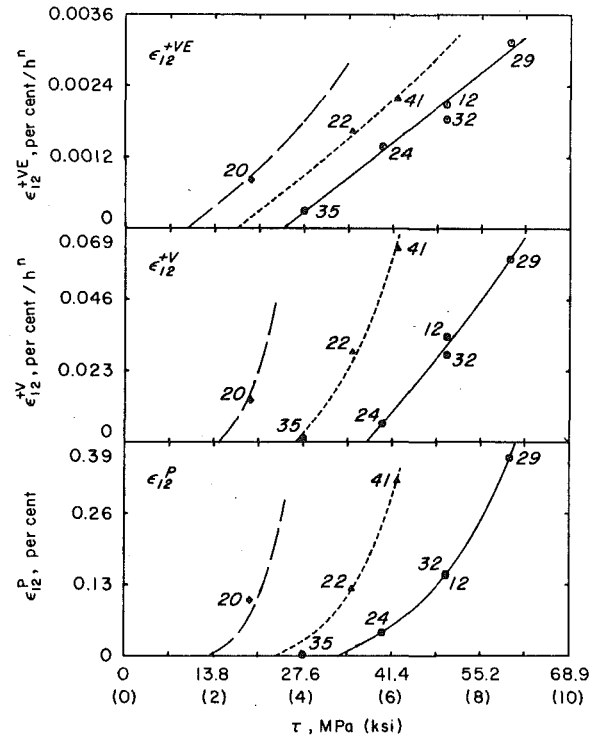


Fig. 7 Shear stress τ versus $\epsilon_{12}{}^{+VE}$, $\epsilon_{12}{}^{+V}$, and $\epsilon_{12}{}^{P}$. Numbers indicate test numbers in Table 1. Lines indicate equation (9) for $G_i{}^{VE}$, $G_i{}^{V}$, and $G_i{}^{P}$.

$$\nu = \frac{E}{2G} - 1 = 0.333,$$

where $E = 14.81 \times 10^4$ MPa (21.48×10^6 psi) and $G = 5.56 \times 10^4$ MPa (8.06×10^6 psi) at 593°C (1100°F). These values of E and G were used to construct the slopes of the straight lines in Fig. 5.

Effect of Stress—Multiple Integral Equation

In previous work of one of the authors, it was found that the nonlinear creep of viscoelastic materials under combined stress was well described by the first three terms of a multiple integral representation [10] with simplications to single integrals by means of the modified superposition principle [9] for varying stress states.

For constant stress the third-order representation yields the following expression for axial strain ϵ_{11} and shear strain ϵ_{12} under combined tension σ and torsion τ stresses,

$$\epsilon_{11} = F_1\sigma + F_2\sigma^2 + F_3\sigma^3 + F_4\sigma\tau^2 + F_5\tau^2, \tag{8}$$

$$\epsilon_{12} = G_1\tau + G_2\tau^3 + G_3\sigma\tau + G_4\sigma^2\tau, \tag{9}$$

where

$$F_i = F_i{}^0 + F_i(t), \tag{10}$$

$$G_i = G_i{}^0 + G_i(t), \tag{11}$$

and $F_i{}^0$ and $G_i{}^0$ are time-independent terms while $F_i(t)$ and $G_i(t)$ are time-dependent terms. $F_i(t)$ and $G_i(t)$ become $F_i{}^+t^n$ and $G_i{}^+t^n$ in accordance with (3) for the present material. As shown in the foregoing, n had no definite trend with stress (so was taken to be a constant) and $F_i{}^+$ and $G_i{}^+$ were functions of stress.

Creep Limit. As shown in Figs. 6 and 7, it appeared that there may exist limiting stresses (creep limits) below which the creep strain components, $\epsilon_{ij}{}^{+VE}$ and $\epsilon_{ij}{}^{+V}$ are zero or negligible. Incorporating the concept of a creep limit, (8) and (9) may be rewritten as

$$\epsilon_{11} = F_1(\sigma - \sigma') + F_2(\sigma - \sigma')^2 + F_3(\sigma - \sigma')^3 + F_4(\sigma - \sigma')(\tau - \tau')^2 + F_5(\tau - \tau')^2, \tag{12}$$

$$\epsilon_{12} = G_1(\tau - \tau') + G_2(\tau - \tau')^3 + G_3(\sigma - \sigma')(\tau - \tau') + G_4(\sigma - \sigma')^2(\tau - \tau'), \tag{13}$$

where σ', τ' are components of a creep limit under combined tension and torsion.

A Mises' relation for the creep limit results in,

$$(\sigma')^2 + 3(\tau')^2 = (\sigma^*)^2, \tag{14}$$

$$\sigma^* = \sqrt{3}\,\tau^*, \tag{15}$$

$$\sigma/\tau = \sigma'/\tau', \tag{16}$$

where σ^*, τ^* are creep limits for pure tension and pure torsion respectively. A Tresca creep limit is given by (14)–(16) by replacing 3 by 4 in (14) and (15).

Determination of Constants. The values of F_i, G_i, σ^*, and τ^* may be determined from the results of a minimum of four pure tension, three pure torsion, and two combined tension and torsion creep tests. Simultaneous equations are required except where special circumstances permit simplifications.

Viscoelastic Strain Component

The data of $\epsilon_{ij}{}^{+VE}$ for pure tension σ and pure torsion τ seemed to be best fitted to a straight line but synergistic in combined σ and τ as shown in Figs. 6 and 7. Following this observation, $G_2{}^{VE}$, $F_2{}^{VE}$, and $F_3{}^{VE}$ were taken to be zero in equations equivalent to (12) and (13). The VE superscripts were used to identify the constants for the strain components, $\epsilon_{ij}{}^{+VE}$.

The values of $F_1{}^{VE}$, σ^{*VE} and $G_1{}^{VE}$, τ^{*VE} were obtained by best fit from all pure tension $\epsilon_{11}{}^{+VE}$ and pure torsion $\epsilon_{12}{}^{+VE}$ test data in Figs. 6 and 7, respectively. The limit stresses σ^{*VE} and τ^{*VE} were the viscoelastic creep limits below which the viscoelastic strain appeared negligible. The ratio $\sigma^{*VE}/\tau^{*VE} = 5.917/3.532 = 1.675$ was close to $\sqrt{3}$ of the Mises relation. Thus the Mises relation was used in the following analysis. Since the torsion data was less influenced by temperature drift than tension, a new value of $\sigma^{*VE} = 6.113$ ksi was

Table 2 Constants of equations (13) and (14) for ϵ^{+VE}, ϵ^{+V}, and ϵ^{P}

Constants	ϵ^{+VE} Percent/h^n	ϵ^{+VE} 10^{-4}%/h^n	ϵ^{+V} Percent/h^n	ϵ^{+V} 10^{-4}%/h^n	ϵ^{P} Percent	ϵ^{P} 10^{-4} percent
F_1	6.574×10^{-5}/MPa	(4.533 /ksi)	1.201×10^{-5}/MPa	(82.80 /ksi)	6.398×10^{-3}/MPa	(441.1/ksi)
F_2	0	(0)	-5.406×10^{-5}/MPa^2	(−25.70 /ksi^2)	-2.621×10^{-4}/MPa^2	(−124.6/ksi^2)
F_3	0	(0)	2.040×10^{-6}/MPa^3	(6.685/ksi^3)	7.545×10^{-6}/MPa^3	(24.73/ksi^3)
F_4	2.815×10^{-8}/MPa^3	(0.09227 /ksi^3)	5.922×10^{-6}/MPa^3	(19.41 /ksi^3)	1.845×10^{-5}/MPa^3	(60.46/ksi^3)
F_5	0	(0)	0	(0)	0	(0)
G_1	8.548×10^{-5}/MPa	(5.894 /ksi)	2.461×10^{-3}/MPa	(169.7 /ksi)	5.800×10^{-3}/MPa	(399.9/ksi)
G_2	0	(0)	3.801×10^{-7}/MPa^3	(1.246/ksi^3)	1.088×10^{-5}/MPa^3	(35.65/ksi^3)
G_3	0	(0)	0	(0)	0	(0)
G_4	2.545×10^{-9}/MPa^3	(0.008343/ksi^3)	2.131×10^{-6}/MPa^3	(6.986/ksi^3)	6.880×10^{-6}/MPa^3	(22.55/ksi^3)
σ^*	42.18 MPa	(6.118 ksi)	64.73 MPa	(9.388 ksi)	57.21 MPa	(8.297 ksi)
τ^*	24.35 MPa	(3.532 ksi)	37.37 MPa	(5.420 ksi)	33.03 MPa	(4.790 ksi)

calculated using (15) and $\tau^{*VE} = 3.532$ ksi. Also F_1^{VE} was recalculated from the new σ^{*VE}. Since there occurred negligible axial creep strain during pure torsion creep tests, the constant F_5^{VE} of (12) was taken to be zero. Calculation of G_3^{VE} and G_4^{VE} for ϵ_{ij}^{+VE} and ϵ_{ij}^{+V} from two combined tests, C22 and C41, which had the same ratio of σ and τ, yielded a negative value of G_4^{VE}, which seemed physically unlikely. Then assumptions of $G_3^{VE} = 0$ or $G_4^{VE} = 0$ were tried. The results showed no significant difference between the two. So $G_3^{VE} = 0$ was used in subsequent analysis because negative axial loading (compression) was expected not to differ much from positive in its influence on shear creep strain within the small creep strain range. $G_3^{VE} \neq 0$ indicates an effect from the sign of the axial stress, where as $G_4^{VE} \neq 0$ has no effect of sign of axial stress. Test C20 was not used in calculating the constants because only two combined stress tests were required.

The components of the viscoelastic creep limits σ'^{VE} and τ'^{VE} for the combined stress tests were calculated by (14)–(16). All constants for ϵ^{VE} are shown in column 2 of Table 2. The calculated curves for ϵ_{11}^{+VE} and ϵ_{12}^{+VE} are shown in Figs. 6 and 7 as solid lines for pure tension or torsion, dotted lines for stress states having the same ratio σ/τ as tests C22 and C41 and dashed lines for stress states having the same ratio σ/τ as test C20.

From these results the values of ϵ_{11}^{+VE} and ϵ_{12}^{+VE} were calculated for combined stress test C20. The results compared very satisfactorily with the test data in Fig. 7; but less so in Fig. 6.

Effect of Creep Period on Recovery of Exponent n_1. The viscoelastic strain component ϵ_{ij}^{+VE} was obtained from the recovery data using (4), which is a function of the creep period t_1. Since the viscoelastic strain ϵ^{VE} during creep might be affected by aging, by strain hardening, by saturation of recoverable strain and by temperature drift the recovery following quite different creep periods might be different. A plot (not shown) of n_1 {obtained from (4) and shown in Table 1} versus the prior creep period t_1 showed no significant variation of n_1 over the range of t_1 from 100 to 1000 h. Thus n_1 was taken as independent of t_1 in the foregoing analysis.

Viscous Strain Component

The viscous strain component ϵ_{ij}^{+V} was calculated by (7) and shown in Table 1 and Figs. 6 and 7. The component ϵ_{11}^{+V} was nonlinear in stress as shown in Fig. 6. The plot of ϵ_{12}^{+V} versus τ for pure torsion was nearly a straight line as shown in Fig. 7. Equations (12) and (13) were employed to describe the data using V as a superscript to identify the viscous component. Data for tests A43 and T35 were omitted in calculating the constants because their strains were negligible and their stresses were below the viscous creep limits. Then the pure tension tests at three different stress levels were insufficient to determine the four constants F_1^V, F_2^V, F_3^V and σ^{*V}. So the values of G_1^V, G_2^V, and τ^{*V} were first computed from pure torsion data. Then σ^{*V} was calculated from (15) and F_1^V, F_2^V, F_3^V were calculated from pure tension creep data using the value of σ^{*V} obtained from (15).

Taking $F_5^V = G_3^V = 0$ the values of F_4^V and G_4^V were determined from tests C22 and C41 in the same way as for ϵ^{+VE}. Also σ'^V and τ'^V were calculated by (14)–(16). All the resulting constants are shown in column 3 of Table 2. The corresponding theory curves computed from (12) and (13) are shown in Figs. 6 and 7 as solid lines for pure tension or torsion, as dotted lines for stress states having the same ratio σ/τ as tests C22 and C41, and as dashed lines for stress states having the same ratio σ/τ as test C20.

The test C20 was not used in determining the constants. As shown in Fig. 6 the C20 data lies significantly above the theory curve for ϵ_{11}^{+}. However, in Fig. 7 the agreement between the data for C20 and the theory is very satisfactory.

In the present paper only positive viscous strain components ϵ_{ij}^{V} (pos) were considered because all stresses were positive.

Plastic Strain Component

The plastic strain was obtained from (7) as shown in Table 1. The plots ϵ_{11}^P versus σ and ϵ_{12}^P versus τ were very similar to those of ϵ_{ij}^{+V} as shown in Figs. 6 and 7, respectively. This observation suggested that the stress-dependence of plastic strain on initial loading might be described by the same form as (12) and (13). These equations do not describe changes in shape of the yield surface due to yielding or creep straining or strain aging anymore than (12) and (13) describe creep following changes in stress state. Thus (12) and (13) with superscripts P to identify the plastic strain component were also used to describe the plastic strain data. The computations followed the same process as employed for the viscous strain component ϵ^{+V} to determine the constants F_i^P, G_i^P, σ^{*P}, τ^{*P}, σ'^P, and τ'^P. The values determined are shown in column 4 of Table 2. The curves for ϵ_{11}^P and ϵ_{12}^P calculated from the constants in Table 2 are shown in Figs. 6 and 7 as solid lines for pure tension or torsion, as dotted lines for stress states having the same ratio σ/τ as Tests C22 and C41 and as dashed lines for stress states having the same ratio σ/τ as test C20. Again test C20 was not employed in the analysis. The agreement between the theory and test C20 was very good in Fig. 7, but poor in Fig. 6.

Discussion

As shown in the analysis the creep strain was separated into several components, assuming no interaction among them and independent controlling mechanisms for each component. But several questions remain. The plastic strain on loading might affect the creep strain by strain hardening and by anisotropy induced by straining [11]. At low temperature, creep and time-independent plastic deformation are usually considered as physically equivalent processes [12] which may be partly right even at high temperature. The similar stress dependence of ϵ_{11}^P and ϵ_{11}^V in Fig. 6 suggests this view. It was also noted that time-dependent strain may be affected by plastic strain such that the larger the plastic strain the smaller the time-dependent strain at the same stress level, though the available data are inconclusive.

The difference between the yield limit and the viscous creep limit is perhaps not significant. For instance if ϵ_{12}^P and ϵ_{12}^{+V} in Fig. 7 were represented by straight lines the limit stress would be about the same for each. On the other hand there is no doubt that the creep limit for the viscoelastic strain component ϵ^{+VE} is less than the limits for ϵ^P and ϵ^{+V}.

As shown in Figs. 6 and 7 the viscoelastic creep components $\epsilon_{11}{}^{+VE}$ and $\epsilon_{12}{}^{+VE}$ are linear in $\sigma - \sigma^{*VE}$ and $\tau - \tau^{*VE}$, respectively, for pure tension or pure torsion. However, they are also synergistic as shown by the fact that the values are larger when the other stress component was added to the pure stress component. Thus a linear theory such as linear viscoelasticity can not be used. Figs. 6 and 7 also show that ϵ^P and ϵ^{+V} are also strongly synergistic whether nearly linear as in $\epsilon_{12}{}^{+V}$ versus $\tau - \tau^{*V}$ or nonlinear as in $\epsilon_{11}{}^{+V}$ versus $\sigma - \sigma^{*V}$.

The available data indicate that there may be some nonrecoverable creep and some plastic strain even at stresses below the apparent creep limit or yield limit, respectively, see tests 43 and 35 in Figs. 6 and 7, respectively.

The values of the ratio $\epsilon^{+V}/\epsilon^{+VE}$ in Table 1 suggest that the ratio may be constant at low stresses {Mises stresses of 68.9 MPa (10 ksi) or less}. Above this stress level the ratio increases substantially with increase in stress level.

Work in Progress

Subsequent papers will include the following investigations of creep of 304 stainless steel at 593°C (1100°F): the basic information given in Table 1 will be used together with suitable constitutive equations to predict the creep behavior under combinations of tension and torsion with step changes in stress including stepup, stepdown, side steps (in which one stress component remains constant while the other increases or decreases); and stress reversal. These predictions will be compared with results of actual experiments. The effect of aging will also be reported.

Time permitting, compression creep, relaxation, and stresses below the apparent creep limits will also be investigated.

Conclusions

It was shown that creep strain could be separated explicitly into five components by a viscous-viscoelastic model employing a power function of time. The time-dependence of the recoverable and nonrecoverable components was found to be the same, except for recoverable tensile strain.

Limit stresses were found below which plastic and creep strains were negligible. While the limit stresses for plastic and nonrecoverable creep were nearly the same the limit stress for recoverable creep was significantly lower. The stress-dependence of the time-independent nonrecoverable (plastic) strain and time-dependent nonrecoverable (viscous) strain was similar, nonlinear, and synergistic. The time-dependent recoverable strain was nearly linear in the stress increment above the creep limit but was synergistic.

Using the same value of time exponent for all components of creep the model described the creep and recovery under tension, torsion and combined tension and torsion very well, except for recovery of tensile components, where the actual recovery was more rapid than described. A larger value of time exponent was indicated for this situation.

Acknowledgments

This work was supported by the National Science Foundation under Grant No. ENG76-02519. The authors are grateful to R. M. Reed for assistance in experiments and Ms. P. C. Orzechowski for typing the manuscript. The early experimental work on this project was supported and all material tested was supplied by the Oak Ridge National Laboratory of the U.S. Atomic Energy Commission.

References

1 Findley, W. N., Cho, U. W., and Ding, J. L., "Creep of Metals and Plastics Under Combined Stresses, a Review," ASME *Journal of Engineering Materials and Technology,* Vol. 101, Oct. 1979, pp. 365–368.

2 Findley, W. N., and Mark, R., "Multiaxial Creep Behavior of 304 Stainless Steel," Annual Report No. 1 (Brown University Report EMRL-55, Oct. 1972), No. 2 (EMRL-58, Nov. 1973), No. 3 (EMRL-62, Nov. 1974), No. 4 (EMRL-63, July 1975) to Oak Ridge National Laboratory.

3 Mark, R., and Findley, W. N., "Concerning a Creep Surface Derived From a Multiple Integral Representation for 304 Stainless Steel Under Combined Tension and Torsion," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, No. 4, Dec. 1978, pp. 773–780.

4 Liu, K. C., "Elastic-Plastic Behavior Studies," in High Temperature Structural Design Methods for LMFBR Components, Progress Report for Period of Oct. 1972 through Mar. 1973, by W. L. Greenstreet, J. M. Corum, and C. E. Pugh, USAEC Report ORNL-TM-4089, Oak Ridge National Laboratory, 1973.

5 Findley, W. N., and Gjelsvik, A., "A Biaxial Testing Machine for Plasticity, Creep, or Relaxation Under Variable Principal-Stress Ratios," *Proceedings, ASTM,* Vol. 62, 1962, pp. 1103–1118.

6 Findley, W. N., and Lai, J. S., "Creep and Recovery of 2618 Aluminum Alloy Under Combined Stress With a Representation by a Viscous-Viscoelastic Model," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, No. 3, Sept. 1978, pp. 507–513.

7 Lai, J. S., and Findley, W. N., "Creep of 2618 Aluminum Under Step Stress Changes Predicted by a Viscous-Viscoelastic Model," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, No. 1, Mar. 1980, pp. 21–26.

8 Findley, W. N., and Lai, J. S., "Creep of 2618 Aluminum Under Side Steps of Tension and Torsion and Stress Reversal Predicted by a Viscous-Viscoelastic Model," EMRL Report No. 72, Brown University, Aug. 1979.

9 Findley, W. N., and Lai, J. S. Y., "A Modified Superposition Principle Applied to Creep of Nonlinear Viscoelastic Material Under Abrupt Changes in State of Combined Stress," *Transactions of the Society of Rheology,* Vol. 11, Part 3, 1967, pp. 361–380.

10 Findley, W. N., Lai, J. S., and Onaran, K., *Nonlinear Creep and Relaxation of Viscoelastic Materials, With an Introduction to Linear Viscoelasticity,* North-Holland, Amsterdam, 1976.

11 Rabotnov, Yu. N., *Creep Problems in Structural Members,* North-Holland, Amsterdam, 1969.

12 Lagneborg, R., "Dislocation Mechanisms in Creep," *International Metallurgical Reviews,* Vol. 17, 1972, pp. 130–146.

R. M. Christensen

Department of Chemistry
and Materials Science,
Lawrence Livermore Laboratory,
L-338,
University of California,
Livermore, Calif. 94550
Fellow ASME

A Nonlinear Theory of Viscoelasticity for Application to Elastomers

A simple nonlinear theory of viscoelasticity has been developed for application to elastomers. The theory is the viscoelastic generalization of the kinetic theory of rubber elasticity, and it is used to model time and rate-dependent effects. The method of derivation reveals that the theory is applicable to stress-imposed rather than strain-imposed conditions. Thus the creep test provides the logical means for deducing the material properties of the theory, while the relaxation-test technique is not applicable. When the long time asymptotic state is at the same level of deformation for both tests, the sources and degree of nonlinearity in relaxation tests are shown to be more severe than those involved in creep tests. A simple means is deduced for obtaining the creep function from the nonlinear creep response of the elastomer.

Introduction

In the behavior of materials many sources for nonlinearities exist. In the case of polymers, nonlinear behavior is commonplace, even at small strains. However, a severe complication for polymers is the interaction of the inherent time-dependence and the sources of nonlinearity. A simple nonlinear theory for the viscoelastic behavior of elastomers is developed here, one which has a mathematically consistent treatment of the time-dependent effects and of the characterization of the nonlinearity.

Before considering the subject of nonlinear viscoelasticity, we must appraise the status of nonlinear elasticity. Any theory of nonlinear viscoelasticity must admit the limiting case of elastic behavior and, therefore, must include elasticity theory results. The general continuum theory of nonlinear elasticity, as developed by Cauchy, has been used by Rivlin [1] to obtain general solutions to boundary-value problems without specifying the strain-energy function. Within the continuum context there are no specific derived forms for the stress-strain relations. Often the stress-strain relations are taken as those resulting from a polynomial expansion of a strain-energy function in terms of the invariants of the deformation measure. However, Rivlin and Sawyers [2] show that expressing the strain-energy function as an algebraic function is not useful; it is more meaningful to simply utilize the experimentally derived dependence of the strain energy upon the invariants. The main observation to be made at this point is that from a continuum mechanics point of view there is no simple universal form for the stress-strain relations of nonlinear elasticity.

From a molecular point of view, in contrast to the situation just described for continuum mechanics, there is a very specific result. In particular, the kinetic theory of rubber elasticity predicts a specific form for the strain-energy function, namely, that the strain-energy depends linearly on the first invariant of the deformation measure. To be sure, this simple result does not model all experimental data. However, over a limited strain range for many elastomers, it does a reasonable job. The overwhelming significance of the kinetic theory of rubber elasticity is that the stress-strain relation is derived from a very simple set of reasonable physical assumptions (see Treloar [3] for the specifics of the derivation). In general terms, the use of Gaussian statistics leads to the "entropy-spring" formulation for the single molecule, which provides the general result when summed over all molecules. The kinetic theory of rubber elasticity is taken here as the guide for the development of a viscoelastic theory at the same level of applicability.

Our purpose is to derive the viscoelastic counterpart of the kinetic theory of rubber elasticity. This problem is not approached here by means of molecular theory but rather by seeking the answer in the continuum context. Obtaining the viscoelastic generalization of the kinetic theory of rubber elasticity is not simply a matter of reinterpreting functions as functionals. As we shall see, the problem is far more subtle than that. In fact, we must be very careful in implementing the statement that the viscoelastic theory should reduce to the kinetic elasticity theory. Under what conditions would we expect

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, February, 1980; final revision, April, 1980. Paper No. 80-WA/APM-18.

the viscoelastic theory to model the given elastic behavior? Precise statements of these requirements are given in the next section.

To be sure, there already exist many theories of nonlinear viscoelasticity, and we make no attempt to review them all. None of them correspond exactly to what is found here nor have any of them been posed in the context discussed in the foregoing. The existing nonlinear theories that have some relationship to the present work will be mentioned in the appropriate section. None of the existing theories provides a universal model of material behavior, when compared with actual data. There has been a tendency to add one more parameter or function with each succeeding theory to enable it to model yet one more effect. The approach here is just the opposite. In essence, the simplest, physically meaningful, nonlinear theory of viscoelasticity is sought and its range of applicability will be delineated through careful analysis of the underlying assumptions of the theory.

In addition, the applicability of the theory to specific material behavior will be examined. That is, we will specialize the general theory to model various specific problems and will consider the determination of the relevant material properties. Finally, some of the physical implications of the results with regard to sources of nonlinear behavior will be discussed.

One more restriction on the work should be discussed before proceeding to the general theory. Attention will be restricted to what are called stress-imposed problems rather than strain-imposed problems. In a strain-imposed problem, boundary displacements are the primary input or excitation variables, whereas in stress-imposed problems, boundary stresses are the primary variables. Consider, first, strain-imposed problems. In a practical context, if an elastomer is required to function in a strain-imposed context, generally speaking, the material is only required to "fill the gap" with no concern for its stiffness or response. Such materials are often referred to as fillers and potting compounds. For stress-imposed problems, the polymer is typically required to bear the load, and it deforms by whatever amount is necessary to sustain the load without failure. Examples of stress-imposed problems are the response of rubber tires and the behavior of bond lines in adhesive joints. Clearly the stress-imposed problems pose a greater challenge than do strain-imposed problems. In this paper, discussion of theoretical developments will be restricted to those for application to stress-imposed problems and also will be restricted to incompressible and isotropic materials, under quasi-static conditions.

A Special Nonlinear Theory

Several restrictions, in addition to incompressibility and isotropy, on the general theory must now be delineated. The material will be taken as behaving under isothermal conditions, at a temperature far above that of the glass transition. This requirement means that in the absence of history (time)-dependence, the material would respond in the rubbery range.

When time and rate-dependence are included, it is not sufficient to say the material is in the rubbery range of behavior. In fact, if the excitation is rapid enough, the material will respond in a manner dictated by a glassy-type elasticity effect, even though the temperature is far above the glass transition temperature. When including time and rate-dependence, to be precise, we must say that for a sufficiently slow process the material will respond in a rubbery manner. The complete requirements are stated in (*i*) and (*ii*).

(*i*) Under a sufficiently slow process, the viscoelastic theory must reduce to the kinetic theory of rubber elasticity, which is given by

$$\sigma_{ij} = x_{i,K}x_{j,L}\frac{\partial W}{\partial E_{KL}},$$

where $W = cI$, with I being the first invariant of the strain measure given by

$$2E_{KL} = C_{KL} - \delta_{KL},$$

where C_{KL} is the right Cauchy-Green tensor.

The foregoing well-known terminology is defined later in the viscoelastic context. Although it is intuitively obvious what is meant by a "slow process," this term will be given a mathematical characterization in the developments to follow. The second requirement is stated as

(*ii*) The viscoelastic theory will be applicable to stress-imposed rather than strain-imposed problems.

As discussed in the Introduction, a stress-imposed problem is one in which the primary variable of excitation is that of stress (or load) rather than strain (or displacement). We also must assume certain smoothness conditions. Specifically, in comparing stress-imposed with strain-imposed histories, it is necessary to assume the same degree of smoothness in each. Thus we compare step functions in strain, for relaxation tests, with step functions in stress, for creep tests. In the present context, it would not be proper to compare, say, a step function in stress with a much smoother history in strain. For the most part, in the following, when speaking of stress and strain-imposed problems, our specific interest will be in the respective creep and relaxation tests.

The starting point for the development of the stress-strain relation is the general form of the constitutive relation given by

$$\sigma_{ij} = -p\delta_{ij} + x_{i,K}x_{j,L}\,\underset{s=0}{\overset{\infty}{\phi_{KL}}}(E_{KL}(t-s), E_{KL}(t)), \tag{1}$$

where σ_{ij} is the Cauchy stress, p is the pressure, $x_{i,K}$ denotes the displacement gradients, $\underset{s=0}{\overset{\infty}{\phi_{KL}}}(\ )$ is a tensor-valued functional with a dependence not only on the strain history but also on the current strain, which is defined by

$$2E_{KL} = x_{i,K}x_{i,L} - \delta_{KL}. \tag{2}$$

Cartesian tensor notation is used, with x_i being the coordinates of the deformed configuration with X_K being the coordinates of the stress-free preferred configuration, which is assumed to exist, and the comma notation designates partial differentiation.

The next step is to provide a special representation for the functional shown in (1). To this end, the Green-Rivlin expansion method [4] will be followed and $\underset{s=0}{\overset{\infty}{\phi_{KL}}}$ will be taken as a polynomoial expansion in the linear function of the type

$$\int_0^t g(t-\tau)\frac{\partial E_{KL}(\tau)}{\partial \tau}d\tau.$$

Performing the expansion, beginning with the zero-order term, gives

$$\sigma_{ij} = -p\delta_{ij} + x_{i,K}x_{j,L}\left[g_0\delta_{KL} + \int_0^t g_1(t-\tau)\frac{\partial E_{KL}(\tau)}{\partial \tau}d\tau + \delta_{KL}\int_0^t g_2(t-\tau)\frac{\partial E_{JJ}(\tau)}{\partial \tau}d\tau + \ldots\right]. \tag{3}$$

For an incompressible material, Pipkin [5] has shown that the requirement of incompressibility gives

$$\mathrm{tr}\,\mathbf{E} = \mathrm{tr}\,(\mathbf{E}^2) - [\mathrm{tr}\,(\mathbf{E})]^2 + \ldots, \tag{4}$$

where only second and higher-order terms are involved on the right-hand side of (4). By the use of (4), the E_{JJ} term in (3) can be absorbed into the quadratic integral, and (3) can be written as

$$\sigma_{ij} = -p\delta_{ij} + x_{i,K}x_{j,L}\left[g_0\delta_{KL} + \int_0^t g_1(t-\tau)\frac{\partial E_{KL}(\tau)}{\partial \tau}d\tau + \ldots\right]. \tag{5}$$

The next term not explicitly shown in (5) is that of the quadratic functional. Note that the term involving g_0 in (5) corresponds to the kinetic theory of rubber elasticity.

To implement requirement (*i*), we must analytically specify what is meant by a slow process. To do this, we consider accelerated and retarded processes, as discussed by Gurtin and Herrera [6]. For a given deformation history, $x_{i,K}(t)$ and $E_{KL}(t)$, accelerated and retarded strain histories, respectively, are specified by $E_{KL}(\alpha t)$, where $\alpha > 1$ for the accelerated and $\alpha < 1$ for the retarded strain history.

The effect of accelerating the history is the same as compressing the time scale; thus we seek to determine the stress at reduced values of its argument as specified by

$$\sigma_{ij}\left(\frac{t}{\alpha}\right) = -p\delta_{ij} + x_{i,K}(t)\, x_{j,L}(t)\left[g_0\delta_{KL} + \int_0^{t/\alpha} g_1\left(\frac{t}{\alpha} - \tau\right)\frac{\partial E_{KL}(\alpha\tau)}{\partial\tau}\, d\tau + \ldots\right]. \quad (6)$$

With a change of variable, (6) can be written as

$$\sigma_{ij}\left(\frac{t}{\alpha}\right) = -p\delta_{ij} + x_{i,K}(t)\, x_{j,L}(t) \cdot\left[g_0\delta_{KL} + \int_0^{t} g_1\left(\frac{t-\eta}{\alpha}\right)\frac{\partial E_{KL}(\eta)}{\partial\eta}\, d\eta + \ldots\right]. \quad (7)$$

For infinitely accelerated and retarded histories, it is clear from (7) that

$$\sigma_{ij}\left(\frac{t}{\alpha}\right)_{\text{accelerated}} = -p\delta_{ij} + x_{i,K}(t)x_{j,L}(t) \cdot [g_0\delta_{KL} + g_1(0)E_{KL}(t) + \ldots]_{\alpha\to\infty}, \quad (8)$$

and

$$\sigma_{ij}\left(\frac{t}{\alpha}\right)_{\text{retarded}} = -p\delta_{ij} + x_{i,K}(t)x_{j,L}(t) \cdot [g_0\delta_{KL} + g_1(\infty)E_{KL}(t) + \ldots]_{\alpha\to\infty}. \quad (9)$$

Relations (8) and (9) express the well-known result that for very rapid processes only the initial, glassy value of the relaxation function is involved, as with $g_1(0)$ in (8) whereas for very slow processes only the long-time asymptote of the relaxation function is involved, as with $g_1(\infty)$ in (9).

Now consider requirement (*i*). For a very slow process, the viscoelastic theory is required to give the kinetic-theory of elasticity result. Since the kinetic theory is in fact embodied in relations (7)–(9) by presence of the term g_0, it follows from (9) that $g_1(\infty) \equiv 0$, and also the long-time asymptotes of the relaxation functions in all the higher-order terms in (7) must vanish. Thus requirement (*i*) gives

$$g_n(\infty) \equiv 0, \quad n = 1, 2, \ldots, \quad (10)$$

where $g_n(t)$ denotes the relaxation functions involved in (7). Relation (10) is one of the basic results sought. Even though the material is a solid, the relaxation functions decay to zero, as in a fluid. The preferred configuration of the solid is remembered by the rubber elasticity term g_0 in (7).

Before explicitly considering (*ii*), we must obtain some background results. We wish to examine the behavior of (7) for accelerated processes. To this end, we take the derivative of (7) to obtain

$$\frac{\partial\sigma_{ij}}{\partial\alpha} = x_{i,K}(t)x_{j,L}(t)\left[\int_0^t \frac{\partial g_1\left(\frac{t-\eta}{\alpha}\right)}{\partial\alpha}\frac{\partial E_{KL}(\eta)}{\partial\eta}\, d\eta + \ldots\right]. \quad (11)$$

For a positive, monotone-decreasing relaxation function $g_1(t)$ we have

$$\frac{\partial g_1\left(\frac{t-\eta}{\alpha}\right)}{\partial\alpha} > 0. \quad (12)$$

Rewriting (11), we obtain

$$\frac{\partial\sigma_{ij}}{\partial\alpha} = x_{i,K}(t)x_{j,L}(t)\left[\int_0^t H(t-\eta,\alpha)\frac{\partial E_{KL}(\eta)}{\partial\eta}\, d\eta + \ldots\right], \quad (13)$$

where

$$H(t-\eta,\alpha) = \frac{\partial g_1\left(\frac{t-\eta}{\alpha}\right)}{\partial\alpha}.$$

From (12) and (13) we see that as the deformation history is accelerated the stress amplitudes increase. That is, the magnitude of the stress response is increased as the deformation history is accelerated. In arriving at this general conclusion we assume that all the relaxation functions exhibit the type of behavior shown in (12).

Up to this point we have been working with a specified deformation history. Our interest in applying requirement (*ii*) concerns imposed stresses, not strain or deformation. It is not possible to explicitly invert (7) to express strain as functionals of stress, the functionals being determined in terms of those in (7). However, for present purposes, it is sufficient to consider a particular component of stress at a particular instant of time. Thus, for any one component of σ_{ij} in (7), we write

$$\sigma = f(\alpha,E), \quad (14)$$

where E is any convenient scaler measure of the amplitude of the given strain history, as specified by

$$E_{KL}(t) = E\,\hat{E}_{KL}(t). \quad (15)$$

With E taken as a positive, real number, it will later be used as a strain-amplitude scaling factor; we restrict the derivatives of σ in (14) as

$$\frac{\partial|\sigma|}{\partial E} > 0 \quad (16)$$

and

$$\frac{\partial|\sigma|}{\partial\alpha} > 0. \quad (17)$$

Requirement (16) simply expresses the observation that increasing strain amplitude increases stress magnitude while (17) follows from equations (12) and (13).

Under the conditions specified by requirement (*ii*), stress is taken to be the control variable, and we write

$$\frac{d\sigma}{d\alpha} = 0 \quad (18)$$

This result simply states that as the deformation history, specified by (15), is accelerated, the stress is required to remain unchanged. Using (14), we find (18) takes the form

$$\frac{\partial|f(\alpha,E)|}{\partial\alpha} + \frac{\partial|f(\alpha,E)|}{\partial E}\frac{dE}{d\alpha} = 0 \quad (19)$$

or

$$\frac{dE}{d\alpha} = \frac{-\dfrac{\partial|f(\alpha,E)|}{\partial\alpha}}{\dfrac{\partial|f(\alpha,E)|}{\partial E}} = \frac{-\dfrac{\partial|\sigma|}{\partial\alpha}}{\dfrac{\partial|\sigma|}{\partial E}}. \quad (20)$$

Relations (16) and (17) then allow (20) to be written as

$$\frac{dE}{d\alpha} < 0. \quad (21)$$

Relation (21) is a basic result we have been seeking; namely, that if stress amplitudes are controlled to remain unchanged, then accelerating a deformation history $\hat{E}_{KL}(t)$ in (15) must diminish the amplitude of the deformation history. This result is useful in truncating the expansion (5), which is considered next.

To collect our observations, for slow processes the integral terms in (5) are negligible compared with the g_0 term because of the requirement (10). The g_0 term is just the contribution of the kinetic theory of rubber elasticity. Therefore the viscoelastic theory must be subject to the same restrictions as the elasticity theory. In this latter case, for example, extension ratios greater than about 1.5 cannot be

accommodated in natural rubber. Indeed, in any material the strain range of applicability of the kinetic theory is limited because of the assumption of Gaussian statistics, as well as other possible complications. Thus the strain levels will be restricted to small strains in the sense of the possible ultimate strains in elastomers. This does not, however, mean that the strain level is infinitesimal; we retain a fully nonlinear treatment of the kinematics of deformation.

We next consider the effect of rapid processes. For a very rapid process (in stress-control problems), with the result (21) we can show that the strain amplitude must be small. With small strains the quadratic and higher functionals in (5) are of higher order than the linear functional in (5). Thus there is a physical justification for truncation of (5) under very fast process conditions. With such a justification for truncation under both very fast and very slow processes, we in fact truncate (5) explicitly to give

$$\sigma_{ij} = -p\delta_{ij} + x_{i,K}(t)x_{j,L}(t) \times \left[g_0\delta_{KL} + \int_0^t g_1(t-\tau)\frac{\partial E_{KL}(\tau)}{\partial \tau}d\tau\right]. \quad (22)$$

This result, subject to conditions in (10), is the form sought as the simplest, physically meaningful generalization of the kinetic theory of rubber elasticity to model viscoelastic effects. Note that the obtaining of this viscoelastic form was not simply a case of reinterpreting the elastic constant g_0, as a relaxation function and inserting a hereditary effect.

We do not expect relation (22) to be of great generality in modeling material behavior; nevertheless, it may be of use under conditions consistent with its derivation. This possibility will be investigated later.

It is of interest to examine the implications of the restriction of (22) to infinitesimal deformation conditions. Under this circumstance, we write

$$x_{i,K} = \delta_{iK} + u_{i,K}, \quad E_{KL} \cong \epsilon_{KL}, \quad (23)$$

where u_i is the displacement and ϵ_{KL} is the infintesimal strain tensor. Using (23) in (22) gives

$$\sigma_{ij} \cong (-p + g_0)\,\delta_{ij} + 2g_0\epsilon_{ij} + \int_0^t g_1(t-\tau)\frac{\partial E_{ij}(\tau)}{\partial \tau}d\tau. \quad (24)$$

Absorb g_0 into the pressure p and write (24) as

$$\sigma_{ij} \cong -p\delta_{ij} + 2\int_0^t \mu(t-\tau)\frac{\partial \epsilon_{ij}(\tau)}{\partial \tau}d\tau, \quad (25)$$

where

$$2\mu(t) = 2g_0 + g_1(t). \quad (26)$$

Relation (25) is the standard infinitesimal theory form. Comparing (25) and (26) with (22), we see that the nonlinear theory contains the same material properties as are involved in the infinitesimal theory. This result lends special utility to the nonlinear form (22). That is, the properties in the nonlinear theory are accessible through infinitesimal testing.

As a last general remark here, we consider the types of applications for which (22) might be expected to apply. Requirement (*ii*) was a central influence in the derivation. This requirement concerned the applicability to stress rather than strain-imposed problems. The standard creep test is certainly a stress-imposed problem, however the relaxation test is strain-imposed. Thus we expect, at the simplest level, that the result (22) of this derivation is more likely to successfully model creep conditions rather than relaxation conditions. This observation can be generalized to include any type of history, with the present results being more applicable to stress than to strain-imposed histories. We shall consider both types of histories in the following developments. Of course this distinction between strain and stress-imposed histories disappears as infinitesimal deformation conditions are approached.

Creep Inversion

With respect to stress-imposed problems, it is logical to invert the stress-strain relation (22) such that strain is the variable specified in terms of stress history. Contract (22) with $X_{M,i}X_{N,j}$ to obtain

$$s_{KL} + pX_{K,i}X_{L,i} - g_0\delta_{KL} = \int_0^t g_1(t-\tau)\frac{\partial E_{KL}(\tau)}{\partial \tau}d\tau, \quad (27)$$

where s_{KL} is the symmetric Piola stress,

$$s_{KL} = \sigma_{ij}x_{K,i}x_{L,j}.$$

The form (27) may be inverted to give

$$E_{KL}(t) = -J_1(t)\,g_0\delta_{KL} + \int_0^t J_1(t-\tau)\frac{\partial}{\partial \tau}\left[s_{KL}(\tau) + p(\tau)X_{K,i}(\tau)X_{L,i}(\tau)\right]d\tau, \quad (28)$$

where the creep function $J_1(t)$ is defined through the relation

$$\int_0^t J_1(t-\tau)\frac{dg_1(\tau)}{d\tau}d\tau = t. \quad (29)$$

While the integral equation (28) does involve stress history in the integrand, unfortunately the deformation gradient is also involved in the integrand. Thus it is not possible to specify stress, perform the integration, and thereby determine strain response. Rather, the integral equation must be evaluated indirectly, which, in general, is a very difficult matter. Thus, although the stress-strain relation (22) can be formally inverted to give relation (28), there is no practical utility in doing so. For application to creep-test conditions, some other means must be found to invert (22). This will be presented in later developments.

Simple Shear Deformation

The first evaluation of (22) will be given for the case of simple shear deformation. Simple shear deformation is specified by

$$\begin{aligned} x_1 &= X_1 + K(t)X_2, \\ x_2 &= X_2, \\ x_3 &= X_3, \end{aligned} \quad (30)$$

The deformation gradient is given by

$$x_{i,K} = \begin{bmatrix} 1 & K(t) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (31)$$

while the strain tensor is

$$[2E_{KL}] = \begin{bmatrix} 0 & K(t) & 0 \\ K(t) & K^2(t) & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (32)$$

The stresses, evaluated from (22) are

$$\begin{aligned} \sigma_{11} &= -p + g_0[1 + K^2(t)] + K(t)\int_0^t g_1(t-\tau)\frac{dK(\tau)}{d\tau}d\tau + \frac{K^2(t)}{2}\int_0^t g_1(t-\tau)\frac{dK^2(\tau)}{d\tau}d\tau, \\ \sigma_{22} &= -p + g_0 + \frac{1}{2}\int_0^t g_1(t-\tau)\frac{dK^2(\tau)}{d\tau}d\tau, \\ \sigma_{33} &= -p + g_0 \\ \sigma_{12} &= g_0K(t) + \frac{1}{2}\int_0^t g_1(t-\tau)\frac{dK(t)}{d\tau}d\tau + \frac{K(t)}{2}\int_0^t g_1(t-\tau)\frac{dK^2(\tau)}{d\tau}d\tau. \end{aligned} \quad (33)$$

Stress-relaxation conditions can be readily evaluated from (33). We do not do so, however, but instead evaluate (33) under conditions of harmonic oscillation, which corresponds to the common testing technique involving oscillatory shear. The question arises as to the relationship of this type of deformation to an imposed stress test. Under steady-state conditions, the distinction between imposed stress and strain is not of relevance, and either form will reveal the inherent nonlinearities.

Let

$$K(t) = \mathrm{Re}\,(K_0 e^{i\omega t}), \tag{34}$$

where K_0 is a real number corresponding to the amplitude of deformation, and ω is the frequency of excitation. Substituting (34) into the shear-stress term of (33) then gives

$$\begin{aligned}\sigma_{12} = &\left[g_0 K_0 + \frac{\omega K_0}{2}\int_0^\infty g_1(s)\sin\omega s\,ds\right]\cos\omega t\\ &-\left[\frac{\omega K_0}{2}\int_0^\infty g_1(s)\cos\omega s\,ds\right]\sin\omega t\\ &+\left[\omega K_0^3\int_0^\infty g_1(s)\sin 2\,\omega s\,ds\right]\cos 3\omega t\\ &-\left[\omega K_0^3\int_0^\infty g_1(s)\cos 2\,\omega s\,ds\right]\sin 3\omega t.\end{aligned} \tag{35}$$

The effect of the nonlinearity is to produce a higher harmonic term.

It is of practical use to evaluate the work done over a cycle. Since only the shear-stress term does work, the calculation can be written as

$$W = \int_0^{2\pi/\omega}\sigma_{12}(t)\dot{x}_{1,2}(t)\,dt. \tag{36}$$

Using (31) and (35), (36) becomes

$$W = \frac{\pi}{2}K_0^2\omega\int_0^\infty g_1(s)\cos\omega s\,ds. \tag{37}$$

Using conventional terminology, (37) can be written as

$$W = \frac{\pi}{2}K_0^2 g_1''(\omega), \tag{38}$$

where $g_1''(\omega)$ is just the imaginary part of the corresponding complex shear modulus. It is noteworthy that the work done/energy dissipated only depends quadratically on deformation amplitude, the same as in the corresponding linear theory result.

Simple Extension

Next, we consider simple extension conditions. The deformation is specified by

$$\begin{aligned}x_1 &= \lambda X_1\\ x_2 &= (1/\lambda)^{1/2}X_2.\\ x_3 &= (1/\lambda)^{1/2}X_3\end{aligned} \tag{39}$$

The strain tensor is given by

$$[2E_{KL}] = \begin{bmatrix}\lambda^2-1 & 0 & 0\\ 0 & (1/\lambda)-1 & 0\\ 0 & 0 & (1/\lambda)-1\end{bmatrix}. \tag{40}$$

Using the condition $\sigma_{22} = \sigma_{33} = 0$ to evaluate p, it is found from (22) that

$$\begin{aligned}\sigma_{11} = g_0\left(\lambda^2 - \frac{1}{\lambda}\right) - \frac{1}{2\lambda}\int_0^t g_1(t-\tau)\frac{d}{d\tau}\left(\frac{1}{\lambda}\right)d\tau\\ + \frac{\lambda^2}{2}\int_0^t g_1(t-\tau)\frac{d\lambda^2}{d\tau}d\tau.\end{aligned} \tag{41}$$

Relation (41) will be evaluated first for stress-relaxation conditions specified by

$$\lambda = (\lambda_0 - 1)h(t) + 1, \tag{42}$$

where $h(t)$ is the unit step function. Before we substitute (42) into (41), it is useful to integrate (41) by parts. Hence it is found that

$$\sigma_{11} = \left(\lambda_0^2 - \frac{1}{\lambda_0}\right)g_0 + \frac{1}{2}\left(\lambda_0^4 - \lambda_0^2 + \frac{1}{\lambda_0} - \frac{1}{\lambda_0^2}\right)g_1(t). \tag{43}$$

This result will be compared with experimental results, but first the corresponding creep test case will be considered.

The difficulties in dealing with stress-imposed problems was discussed earlier. Here we will determine a general method to deal with creep-imposed conditions, but only in the context of simple extension, which is by far the most important test state. In pursuing this objective of inverting (22) for creep conditions, we shall restrict the results to second-order deformation conditions. That is, we write

$$\lambda = 1 + \epsilon, \tag{44}$$

where ϵ is the strain, and we retain only first and second-order effects in ϵ. This is not a serious restriction in the strain range of intended application. Using (44) in (41) and neglecting third and higher-order terms gives

$$\sigma_{11} = 3g_0\epsilon(t) + \frac{3}{2}[1 + \epsilon(t)]\int_0^t g_1(t-\tau)\frac{d\epsilon(\tau)}{d\tau}d\tau. \tag{45}$$

It is interesting to note there is no second-order contribution to the g_0 term of rubber elasticity.

Common creep tests use imposed load rather than imposed stress. To deal with this case we introduce the engineering stress p_{11} defined by

$$p_{11} = \frac{\sigma_{11}}{\lambda}, \tag{46}$$

which is the stress per unit initial area. Substituting (46) into (45) by using (44) and retaining up to second-order terms gives

$$p_{11} = 3g_0\epsilon(t)[1 - \epsilon(t)] + \frac{3}{2}\int_0^t g_1(t-\tau)\frac{d\epsilon(\tau)}{d\tau}d\tau. \tag{47}$$

Let

$$3g_0 + \frac{3}{2}g_1(t) = E(t). \tag{48}$$

Then with (48), equation (47) becomes

$$p_{11} + 3g_0\epsilon^2 = \int_0^t E(t-\tau)\frac{d\epsilon(\tau)}{d\tau}d\tau. \tag{49}$$

The form (49) can be inverted directly to give

$$\epsilon(t) = \int_0^t J(t-\tau)\frac{d}{d\tau}[p_{11}(\tau) + 3g_0\epsilon^2(\tau)]\,d\tau. \tag{50}$$

We are finally to the point of being able to specify the conditions of a creep test involving a step change in load given by

$$p_{11} = Th(t). \tag{51}$$

Inserting (51) into (50) gives

$$\epsilon(t) = J(t)T + 3g_0\int_0^t J(t-\tau)\frac{d\epsilon^2(\tau)}{d\tau}d\tau. \tag{52}$$

This form is still an integral equation and must be solved for the creep function $J(t)$. To obtain the solution, let

$$J(t) = \frac{\epsilon(t)}{T} + \frac{f(t)}{T}, \tag{53}$$

where $J(t) = \epsilon(t)/T$ is the linear theory result and the term involving $f(t)$ represents the correction to the linear result so that nonlinear effects can be modeled. Substituting (53) into (52) gives

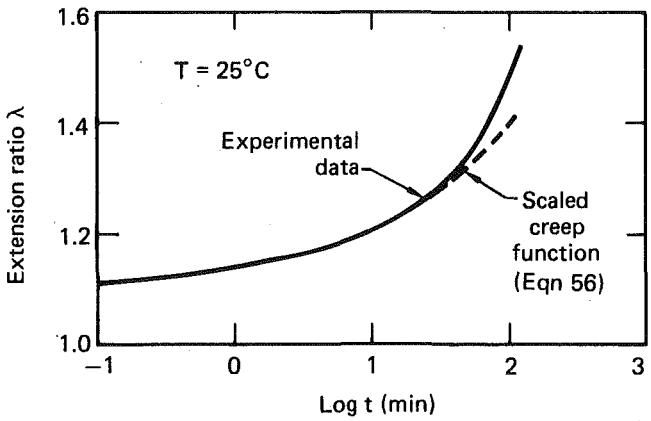


Fig. 1 Creep response of polyisobutylene

$$f(t) + \frac{3g_0}{T}\int_0^t f(t-\tau)\frac{d\epsilon^2(\tau)}{d\tau}d\tau = \frac{-3g_0}{T}\int_0^t \epsilon(t-\tau)\frac{d\epsilon^2(\tau)}{d\tau}d\tau. \quad (54)$$

Note that if the second term on the left-hand side of (54) is neglected, then $f(t)$ is of order $O(\epsilon^3)$ in $\epsilon(t)$. If $f(t)$ is taken to be $O(\epsilon^3)$, then the term that was neglected is of $O(\epsilon^5)$ order and thus is of higher order than the remaining terms. Therefore, it is legitimate to neglect the integral term involving $f(t)$ in (54) provided the coefficient $3g_0/T$ is less than $O(\epsilon^{-1})$. It will be assumed that the coefficient $3g_0/T$ is such that this term can be neglected, and it will be necessary to check this assumption in particular applications. With this order, assessment (54) reduces to

$$f(t) \cong -\frac{3g_0}{T}\int_0^t \epsilon(t-\tau)\frac{d\epsilon^2(\tau)}{d\tau}d\tau. \quad (55)$$

Now (53) can be written as

$$J(t) = \frac{\epsilon(t)}{T} - \frac{3g_0}{T^2}\int_0^t \epsilon(t-\tau)\frac{d\epsilon^2(\tau)}{d\tau}d\tau. \quad (56)$$

Equation (56) is the final form to be used to deduce the creep function from the nonlinear creep test. The last term in (56) represents the nonlinear correction to the linear theory result $\epsilon(t)/T$. It is fortunate that the present theory adapts so easily to the specification of constant load rather than constant Cauchy stress, since the constant load case is the expedient test method.

Next, the result (56) will be evaluated from the creep data of Zapas and Craft [7] on polyisobutylene. The level of the test is given by the stress per unit initial area of

$$T = 0.46 \times 10^6 \text{ dyn/cm}^2.$$

The data of reference [7] are used to estimate g_0 as

$$3g_0 = 0.37 \times 10^6 \text{ dyn/cm}^2.$$

From these two results, the coefficient $3g_0/T$ in (54) is given by

$$3g_0/T = 0.81.$$

Thus the term neglected in obtaining (55) from (54) is established to be of higher order than the terms retained. The creep-response curve from Zapas and Craft [7] is shown in Fig. 1. Simple integration has been performed to estimate the last term in (56). The scaled creep function deduced in this manner is also shown in Fig. 1. At $t = 10$ min, the creep function curve deviates from the response curve by about 4 percent while at $t = 100$ min, the deviation involves about a 20 percent correction.

Some uncertainty is necessarily associated with the calculation of $J(t)$ from (56) (as shown in Fig. 1). This is because reference [7] does not contain an explicit evaluation of the constant g_0, and it was necessary to estimate it indirectly from relaxation data in reference [7].

Finally, it should be mentioned that the general theory derived herein was applied to model the relaxation data on polyisobutylene [7]. It is not necessary to display the data but is sufficient to note that results were not well modeled by (22). This appears to be consistent with the observations found herein; namely, that nonlinearities induced in relaxation tests are generally more severe and more difficult to model than those induced in creep tests. This point will be discussed further in the next section.

Discussion

The relationship of the present work to some other theories should be mentioned. For example, the BKZ theory [8] and finite linear viscoelasticity theory [9] have terms that are similar to those in (22). In fact, the BKZ theory for solids has some terms that are identical to those in (22), and similar comments apply to the finite linear viscoelasticity application of McGuirt and Lianis [9]. However, different expansion techniques were used in these two theories than were employed here. Furthermore, the emphasis in these latter works was on taking enough parameters and functions to fit a wide variety of data. The emphasis herein is to find the viscoelastic generalization of the kinetic theory of elasticity and to deduce the physical limitations on its range of applicability. Indeed, we have done this and have found that creep testing is the perferred mode as we shall elaborate below. Furthermore, the theory is put into a form that allows convenient data reduction in the case of creep in simple extension. Our initial inquiry led to the viscoelastic generalization of the kinetic theory of rubber elasticity. While this result, relation (22), could not be expected to have great generality, it nevertheless appears to be the simplest, completely nonlinear theory of viscoelastic solids. The assumptions under which the theory is derived lead to a new differentiation on sources of nonlinearity in polymer behavior. Specifically, important differences have been found to exist in stress-imposed problems and strain-imposed problems. This difference is most clearly illustrated in the comparison of relaxation and creep tests.

The derivation of the general relation (22) was based on conditions inherent in stress-imposed problems. In the previous section, the attempted application of (22) to relaxation conditions, did not meet with success. The degree of nonlinearity involved in the relaxation test appears to be far different from that accommodated by (22). The critical observation uncovered by this work is that the type and degree of nonlinearity is very different under relaxation conditions from the type and degree involved during creep conditions. This result is consistent with the general characteristics of relaxation and creep tests. For illustrative purposes, consider for a moment that the rubber elasticity term g_0 in (22) models perfectly the long-time, slow-process behavior of a particular material. Now under relaxation conditions, the stress has an initially high value but decays monotonically to the long-time asymptote. On the other hand, the creep test has an initially low value but increases monotonically to the long-time asymptote. Clearly, at short times, the relaxation test is imposing a far more stringent condition on the material than is the creep test. The magnitude of this effect can be established very easily, as follows:

Let stress and strain amplitudes be related by the long-time rubbery asymptote of the relaxation function, as

$$\sigma_0 = E(\infty)\,\epsilon_0.$$

Take the strain amplitude in a relaxation test as ϵ_0 and the stress amplitude in a creep test to be given by σ_0. Using linear theory relations in the case of relaxation, we have

$$\sigma_R(t) = E(t)\epsilon_0;$$

and in the case of creep, we have

$$\epsilon_C(t) = J(t)\sigma_0.$$

Now considering initial values, the last two relations become

$$\sigma_R(0) = E(0)\epsilon_0;$$

and

$$\epsilon_C(0) = J(0)\sigma_0.$$

Taking the ratio of these last two forms gives

$$\frac{\sigma_R(0)}{\epsilon_C(0)} = \frac{E(0)}{J(0)}\frac{\epsilon_0}{\sigma_0}.$$

Using $J(0) = 1/E(0)$ and the first relation $\sigma_0 = E(\infty)\epsilon_0$, we find

$$\sigma_R(0) = \left[\frac{E(0)}{E(\infty)}\right] E(0)\epsilon_C(0).$$

This very simple relation reveals the inclination of the relaxation test to amplify initial-time nonlinearities over that of the creep test by the factor $[E(0)/E(\infty)]$. To say this another way, in a circumstance where the long-time stress in a relaxation test and the long-time strain in a creep test are adjusted to give the same state of deformation (or stress) in the material, then the initial deformation (or stress) in the relaxation test is amplified over that in the creep test by the factor $[(E(0)/E(\infty)]$. Thus the relaxation test at short times provides a far greater penetration of a possible nonlinear range of behavior than does a creep test.

In the past, it appears there has been no concern for the distinction between nonlinear effects in creep and relaxation tests or, more generally, in stress-imposed versus strain-imposed tests. The results here show that it is overwhelmingly important to make this distinction.

The present theory and results are limited to nonlinearities inherent in the nonlinear kinematics of the problem. There are many other sources of nonlinearity in polymeric-material behavior. The role of void nucleation and flaw growth in polymers is now appreciated as being a major source of nonlinearity. Continuum mechanical theories are often applied to model testing results in which these types of nonlinearities are present, often with rather poor results.

The present approach suggests that continuum theories that do not explicitly account for flaw growth cannot be expected to have great generality. This consideration is expected to be especially important in highly filled polymer systems where very rigid particles cause local strain-concentration effects, giving local regions of highly nonlinear deformation with accompanying flaw growth.

Acknowledgments

The author is grateful to Prof. P. M. Naghdi of the University of California, Berkeley, for several very helpful discussions. This work was performed under the auspices of the U. S. Department of Energy by Lawrence Livermore Laboratory under Contract No. W-7405-Eng-48.

References

1 Rivlin, R. S., "Large Elastic Deformations of Isotropic Materials IV. Further Developments of the General Theory," *Philosophical Transactions of the Royal Society*, London, Series A, Vol. 241, 1948, p. 379.

2 Rivlin, R. S., and Sawyers, K. N., "The Strain-Energy Function for Elastomers," *Transactions of The Society of Rheology*, Vol. 20, 1976, p. 545.

3 Treloar, L. R. G., *The Physics of Rubber Elasticity*, 3rd ed., Oxford University Press, 1975.

4 Green, A. E., and Rivlin, R. S., "The Mechanics of Nonlinear Materials With Memory; Part I," *Archive for Rational Mechanics and Analysis*, Vol. 1, 1957, p. 1.

5 Pipkin, A. C., "Small Finite Deformation of Viscoelastic Solids," *Reviews of Modern Physics*, Vol. 36, 1964, p. 1034.

6 Gurtin, M. E., and Herrera, I., "On Dissipation Inequalities and Linear Viscoelasticity," *Quarterly of Applied Mathematics*, Vol. 23, 1965, p. 235.

7 Zapas, L. J., and Craft, T., "Correlation of Large Longitudinal Deformations With Different Strain Histories," *Journal of Research of the National Bureau of Standards*, Section A, Vol. 69, 1965, p. 541.

8 Bernstein, B., Kearsley, E. A., and Zapas, L. J., "A Study of Stress Relaxation With Finite Strain," *Transactions of the Society of Rheology*, Vol. 7, 1963, p. 391.

9 McGuirt, C. W., and Lianis, G., "Constitutive Equations for Viscoelastic Solids Under Finite Uniaxial and Biaxial Deformations," *Transactions of the Society of Rheology*, Vol. 14, 1970, p. 117.

S. R. Swanson

Research Associate Professor,
Department of Mechanical and
Industrial Engineering,
The University of Utah,
Salt Lake City, Utah 84112
Assoc. Mem. ASME

Approximate Laplace Transform Inversion in Dynamic Viscoelasticity

Laplace transform techniques greatly simplify many problems in linear viscoelasticity. However, if realistic material property representations are used, inversion of the resulting transforms can be difficult. Although approximate transform inversion methods have been widely used in quasi-static viscoelastic problems, the application of these techniques to wave propagation problems has been less successful. Inaccuracy of the transform inversion has been noted previously in the literature. The present work shows that one of the numerical Laplace transform inversion techniques of Bellman can successfully be applied to dynamic viscoelasticity. Comparisons with literature solutions and exact functions indicate accuracies to within ±1 percent can be obtained.

Introduction

Approximate Laplace transform inversion has played a useful role in quasi-static linear viscoelastic stress analysis. The motivation for transform techniques stems from the well-known elastic-viscoelastic correspondence principle, as discussed by Christensen, for example, [1], by which the transform (Laplace or Fourier) of a viscoelastic solution can be obtained from an elastic solution. However, if realistic viscoelastic material properties are used, it has often been found difficult to carry out the transform inversion. As indicated by Cost [2], a procedure suggested by Schapery [3] has been found particularly useful. In this procedure the transform of an exponential series is fitted to the transform of the problem solution, by collocation for real, positive values of the Laplace parameter. The inversion of this exponential series transform then is used as an approximate solution of the original problem. According to Schapery, the procedure is based in part on the fact that many viscoelastic solutions change rather slowly with respect to time.

Attempts by Arenz [4, 5] to use this approximate inversion technique in dynamic viscoelastic problems have met with criticism by Knauss [6] and Sackman and Kaya [7]. The difficulty has been associated with accuracy of the approximate inversion in the vicinity of the wave front. Arenz found oscillations at the wave front and attributed them to material dispersion. However, using an alternate numerical technique Knauss [6] showed that the oscillations were ficticious. Thus the accuracy of approximate transform inversion in dynamic viscoelasticity was somewhat in doubt.

However other numerical Laplace transform techniques have been presented by Bellman [8], which display good accuracy for a wide range of problems. In fact one of these techniques has been employed by Tsay [9] in solving a problem in spherical wave propagation in a viscoelastic material, although no basis for judging the accuracy of the solution was established. In view of the generality and ease of obtaining many solutions in the Laplace plane, it was thus of interest to further investigate the accuracy and applicability of approximate Laplace transform inversion in dynamic viscoelasticity.

In the following a brief review will be given of one of Bellman's approximate inversion techniques. This technique will then be applied to the problem of uniaxial wave propagation in a viscoelastic rod, and comparisons with the results of Knauss [6] will be made. A comparison with some simple exact functions is used to illustrate the accuracy of the method. Finally, the half-space problem previously investigated by Arenz [5] is used to illustrate the advantages of the present approach.

Bellman's Technique

Although Bellman [8] actually presents a number of techniques, only one, perhaps the simplest, will be employed in the present work. In this technique, the usual Laplace transformation integral is used with a substituion of the variable $X = \exp(-t)$ so that

$$\overline{F}(s) = \int_0^\infty \exp(-st)F(t)dt = \int_0^1 X^{s-1}F(-\ln X)dX \qquad (1)$$

The integral is then approximated by the sum, as in a standard Gauss quadrature,

$$\overline{F}(s) = \sum_{i=1}^{n} W_i X_i^{s-1} F(-\ln X_i) \qquad (2)$$

where the X_i are the roots of the shifted Legendre polynomials of order n and the W_i are weight functions. Choosing various values for s (taken as positive integers) leads to a system of equations that can be explicitly solved as

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, June, 1980. Paper No. 80-WA/APM-29.

$$F(-\ln X_i) = \sum_{j=1}^{n} C_n(i,j)\overline{F}(S_j) \tag{3}$$

where s_j stands for integer values of s from 1 to n. Thus the desired inverse function F is obtained at each of the time values $t_i = -\ln X_i$, where the X_i are the roots of the nth-order shifted Legendre polynomial. These values of F are obtained by evaluating the function $\overline{F}(s)$ at integer values of the Laplace parameter s from 1 to n, multiplying by the tabulated coefficients $C_n(i, j)$, and summing. Bellman has tabulated the coefficients $C_n(i, j)$ for the shifted Legendre polynomials of order $n = 3$ to 15, as well as the roots.

Two other properties of the Laplace transform are necessary to make the foregoing procedure useful. The first property is related to change of scale. Using the usual relationship

$$L[F(At)] = \frac{1}{A}\overline{F}\left(\frac{s}{A}\right) \tag{4}$$

where L denotes the Laplace transformation, equation (3) becomes

$$F(-A \ln X_i) = \sum_{j=1}^{n} C_n(i,j)\overline{F}(s_j/A)/A \tag{5}$$

where the coefficients $C_n(i, j)$ are the same as used in equation (3).

The second useful property of the Laplace transform is exploited when the function $F(t)$ is known over part of the time range $0 < t < t_0$. Using the change of interval,

$$\overline{F}(s) = \int_0^{t_0} \exp(-st)F(t)dt + \int_{t_0}^{\infty} \exp(-st)F(t)dt \tag{6}$$

or

$$\overline{F}(s) = \int_0^{t_0} \exp(-st)F(t)dt + \exp(-st_0)\int_0^{\infty} \exp(-s\tau)F(\tau + t_0)d\tau \tag{7}$$

The first integral in equations (6) and (7) is assumed to be known. For example, in wave propagation problems this may be the case before the arrival of the first wave.

Combining the relationships given in equations (4) and (7) gives finally

$$F(t_0 - A \ln X_i) = \sum_{j=1}^{n} C_n(i,j) \exp(s\, t_0/A)\{\overline{F}(s_j/A)/A - I_0\} \tag{8}$$

where

$$I_0 = (1/A)\int_0^{t_0} \exp(-st/A)F(t)dt \tag{9}$$

Equation (8) then gives an approximate inverse of the Laplace transform, based on the evaluation of the transformed function $\overline{F}(s_j/A)$ for integer values of s_j. This expression can be evaluated with the aid of a hand calculator for moderately low values of n, and surprising accuracy is obtained for $n = 3$ or 4, for example. However it can also be easily programmed. Because the coefficients $C_n(i, j)$ become large as n increases and alternating in sign, double precision is recommended. In the present work on viscoelastic wave propagation, the highest accuracy was obtained in every case with $n = 15$, the highest order polynomial for which the coefficients are given by Bellman.

In practice it is useful to vary the scale factor A so as to achieve increased accuracy and function definition in a selected time region. Further discussion will be given on this in conjunction with the example problems, but as a starting point if t_c is a time value approximately at the center of the regime of interest, a reasonable value of A is given by

$$A \approx t_c - t_0 \tag{10}$$

Examples

The procedure previously given will be applied to several problems to illustrate the accuracy possible, and the ease of application. These problems are wave propagation in a viscoelastic rod, some simple "wave-like" functions, and a moving pressure load on a viscoelastic half space.

Table 1 Prony series coefficients for stress relaxation modulus of HYSOL 8705, T = 291°K, from [10]

$$E_{rel}(t) = 630 + \sum_{i=1}^{11} E_i \exp(-a_i t)$$

i	E_i (psi)	a_i (min^{-1})
1	6250	1. E+13
2	11950	1 E+12
3	56000	1. E+11
4	44100	1. E+10
5	39400	1. E+9
6	11890	1. E+8
7	4020	1. E+7
8	368	1. E+6
9	279	1. E+5
10	70.7	1. E+4
11	53.4	1. E+3

Wave Propagation in a Viscoelastic Rod. The rod problem considered previously by Arenz [4] and Knauss [6] will again be studied here. Under the usual assumptions the differential equation for wave propagation in an elastic rod is given by

$$E\frac{\partial^2 u}{\partial x^2} = \rho\frac{\partial^2 u}{\partial t^2} \tag{11}$$

where u is displacement in the one-dimensional x-direction, and E and ρ are tensile modulus and mass density. Taking Laplace transforms the displacement in the semi-infinite rod ($x \geq 0$) for a step input displacement u_0 at $x = 0$ is given by

$$\overline{u}(x, s)/u_0 = (1/s)\exp(-xs\sqrt{\rho}/\sqrt{\overline{E}}) \tag{12}$$

Using the elastic-viscoelastic correspondence principle [1], the elastic modulus E is replaced by $s\overline{E}_{rel}$, where $\overline{E}_{rel}$ is the transform of the viscoelastic tensile stress relaxation modulus. The material to be employed is Hysol 8705, used by Knauss [6]. The tensile stress relaxation modulus for this material has been presented in reference [10] in the form of a Prony or Dirichlet series as

$$E_{rel}(t) = E_r + \sum_{i=1}^{n} E_i \exp(-a_i t) \tag{13}$$

The coefficients E_r, E_i, and a_i from reference [10] are given in Table 1. As emphasized in [10], this series provides an excellent fit to measured properties spanning 10 decades of time. Further, the required transform is easily obtained as

$$s\overline{E}_{rel}(s) = E_r + \sum_{i=1}^{n} E_i s/(s + a_i) \tag{14}$$

which can then be substituted for E in equation (12).

The inverse of equation (12) after substitution of equation (14) is then given in Figs. 1–3. Fig. 1 and in more detail Fig. 2 show comparisons with the results of Knauss [6]. As can be seen, the agreement is within plotting accuracy. Fig. 3 shows another detailed comparison with Knauss and also with Arenz taken from [4]. Again the agreement with Knauss is excellent. The spurious oscillation and rounding of the wave form in Arenz's curve can be seen in this comparison.

It should be pointed out that the solutions presented by Knauss [6] and Arenz [4] cannot be directly compared, at least in detail. The material properties used by Arenz differ slightly from those used by Knauss, and are referenced to a temperature of 14.2°C while Knauss uses a reference temperature of 20°C. Further, Arenz plots the displacement versus t', which is defined as time minus the time of arrival

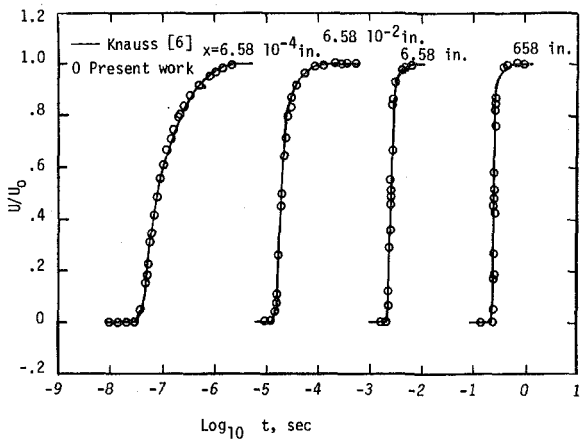


Fig. 1 Wave propagation in a viscoelastic rod

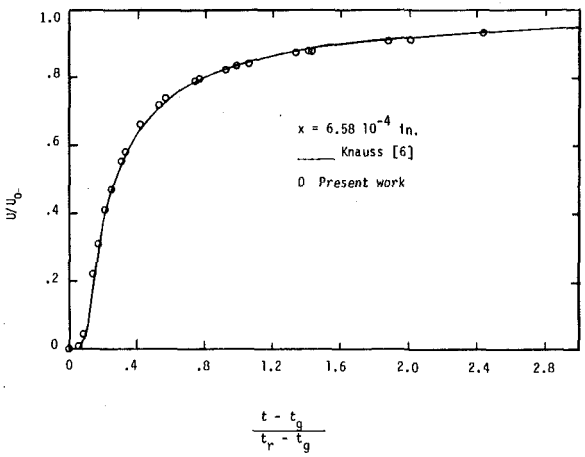


Fig. 2 Displacement history in a viscoelastic rod, normalized time scale

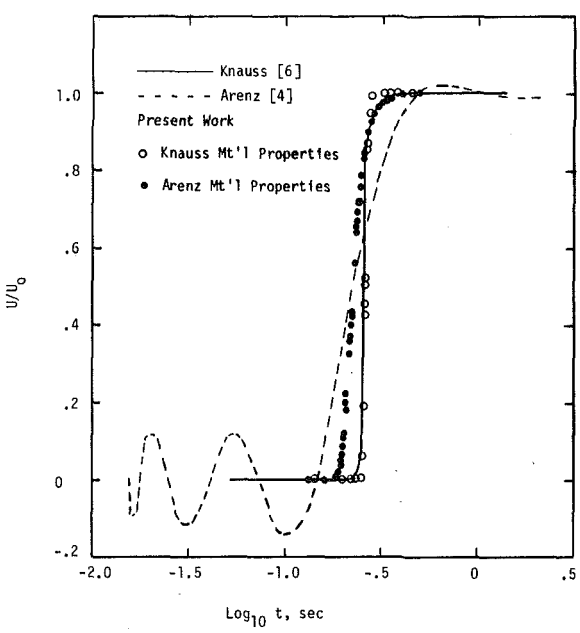


Fig. 3 Comparison of results of Arenz [4], Knauss [6], and present work for wave propagation in a viscoelastic rod

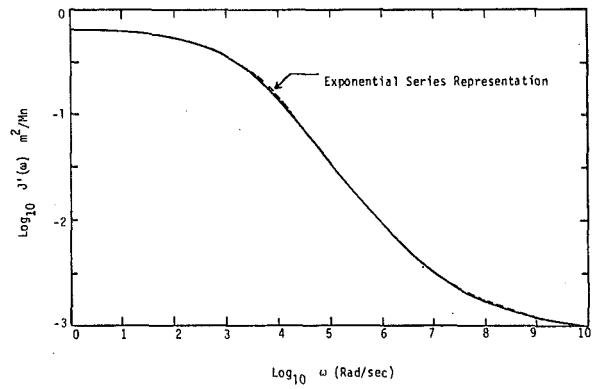


Fig. 4 Real part of the complex shear compliance for polyurethane material, from Arenz [4, 5]

Table 2 Prony series coefficients for shear creep compliance of polyurethane material of [5]

$$J_{crp}(t) = 1.\ E\text{-}3 + \sum_{i=1}^{9} J_i[1 - \exp(-a_i\, t)]$$

i	J_i (m^2/Mn)	a_i (sec^{-1})
1	.454122 E-1	1. E+1
2	.120042	1. E+2
3	.224348	1. E+3
4	.190546	1. E+4
5	.357213 E-1	1. E+5
6	.119715 E-1	1. E+6
7	.174820 E-2	1. E+7
8	.877782 E-3	1. E+8
9	.319831 E-3	1. E+9

of the fastest (glassy) wave. To enable a more direct comparison to be made with Arenz's solution, the rod problem was also solved with the material properties used by Arenz.

Arenz gives the real part of the complex shear compliance, shown reproduced in Fig. 4. It is easy to show that the coefficients of the exponential series representation of the creep compliance

$$J_{\text{crp}}(t) = J_g + \sum_{i=1}^{n} J_i[1 - \exp(-a_i t)] \tag{15}$$

are related to the real part of the complex shear compliance by

$$J'(\omega) = J_g + \Sigma J_i/(1 + \omega^2/a_i^2) \tag{16}$$

Thus the creep compliance coefficients can be obtained by fitting equation (16) to Fig. 4. The coefficients so determined by a least-squares method are given in Table 2 for $n = 9$, and the fit to the original $J'(\omega)$ curve is also shown in Fig. 4. As a side comment it may be remarked that a least-squares series fitting is superior to a collocation scheme as smoother fits are usually obtained. Finally, following Arenz [4], the approximation $D_{\text{crp}} \approx J_{\text{crp}}/3$ was used, where D_{crp} is the tensile creep compliance.

The solution obtained with the foregoing properties, and plotted versus t', is also shown in Fig. 3. While not differing significantly from that obtained using Knauss's properties, it does show that Arenz's method does predict the arrival time and displacement away from the immediate vicinity of the wave front reasonably well.

The procedure followed in obtaining these results is as follows. A

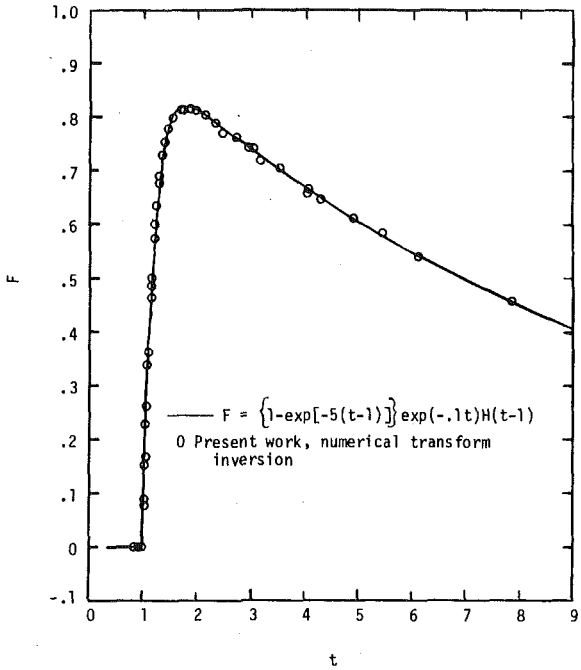


Fig. 5 Comparison of numerical and exact Laplace transform inversion

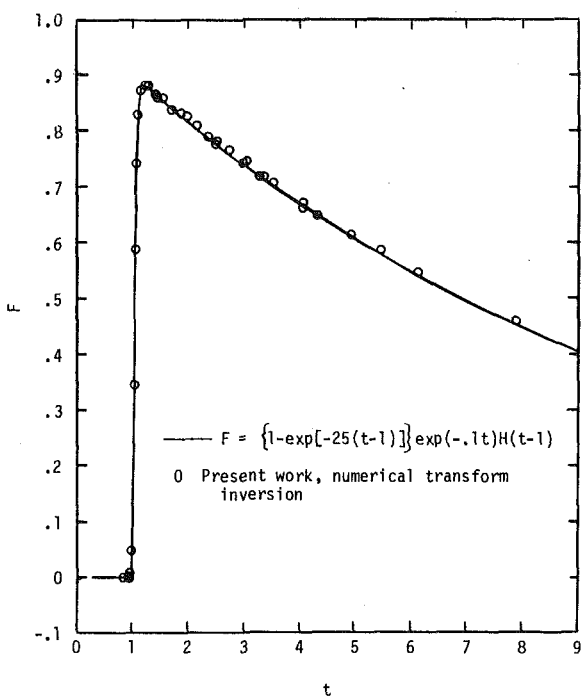


Fig. 6 Comparison of numerical and exact Laplace transform inversion

computation was made for the location x of interest, using $t_0 = 0$ and a range of A-values. This computation served to identify the arrival of the wave. Alternatively, the glassy modulus could have been used for this purpose. The computation was then repeated with t_0 established so that some of the time values of the roots (say 2–4) occurred before the arrival of the wave front. The A-values were then selected so that the time values of the later roots occurred "after" the arrival of the wave front. Since viscoelastic waves can be quite dispersed, the term after can be taken somewhat arbitrarily. The accuracy of the solution is not significantly affected by a wide range of A-values. The

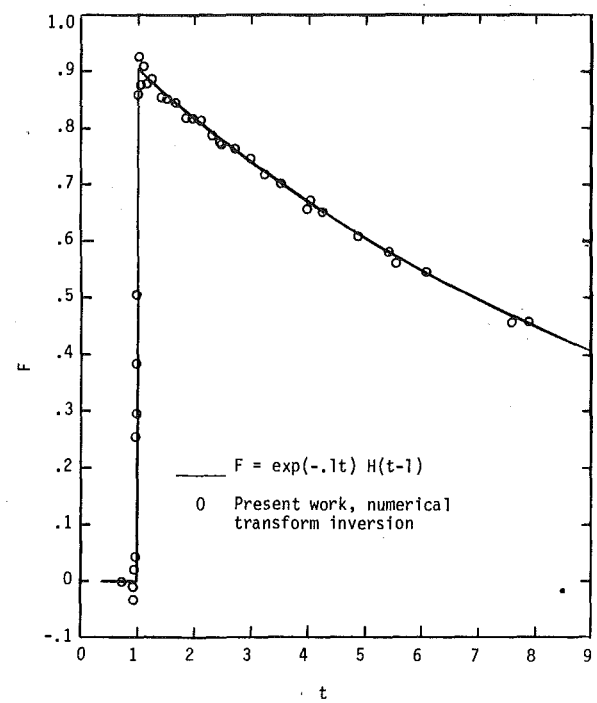


Fig. 7 Comparison of numerical and exact Laplace transform inversion

value $I_0 = 0$ is used in equation (8), as the wave propagates into material at rest. If the value of t_0 is taken too large, i.e., if the displacement is specified as zero longer than it should be, the solution will diverge in an obvious manner.

Exact Functions. Although the comparisons just shown are certainly encouraging, it is difficult to make statements about the accuracy of the present approximate Laplace transform inversion technique. This question was addressed more directly by considering the functions

$$F_1(t) = \{1 - \exp[-b(t-1)]\}\{\exp(-0.1t)\}H(t-1) \tag{17}$$

$$\bar{F}_1(s) = \exp[-(s+0.1)]\{b/[(s+0.1)(s+b+0.1)]\} \tag{18}$$

for various values of b, and in the limit as $b \to \infty$

$$F_2(t) = \exp(0.1t)H(t-1) \tag{19}$$

$$\bar{F}_2(s) = \exp[-(s+0.1)]/(s+0.1) \tag{20}$$

This function has some of the character of viscoelastic stress and strain response curves. A comparison of numerical and exact transform inversion for $b = 5$, 25, and $b \to \infty$ is shown in Figs. 5–7. It might be expected that the more rapidly the function changes with respect to time, the more error will occur in an approximate inversion. This in fact appears to be the case. However even for $b = 25$ the approximation reproduces the exact function to within ±1 percent. However, for the step function $F_2(t)$, more deviation does occur in the immediate vicinity of the step function. The deviation is systematic, giving first a deviation away from zero and then overshooting the function, both by a value of approximately 5 percent. The remainder of the function is then reproduced within about ±2 percent.

It would appear that the dispersion characteristic of viscoelastic waves will enhance the accuracy of the approximate Laplace transform inversion. However in any case the accuracy would seem to be acceptable for many purposes.

Moving Pressure on Viscoelastic Half Space. The final example to be considered is that of a pressure load on a viscoelastic half space, previously examined by Arenz [5]. The problem is that of normal pressure loading advancing over the half space at a constant velocity V_0 under steady-state conditions, with V_0 restricted to be

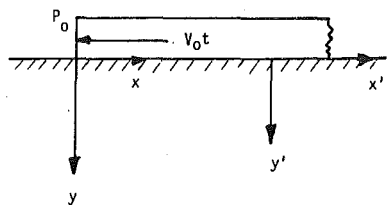


Fig. 8 Coordinate systems for half space with steadily moving step pressure load

greater than the speed of irrotational waves with modulii equal to the glassy or short-time viscoelastic modulii. Plane strain is assumed. Following Arenz [5] and other elastic solutions [11], we establish a fixed coordinate system x', y', and t' and a coordinate system x, y, t moving with the front of the advancing load, as illustrated in Fig. 8. The coordinates are then related by

$$x = x' + V_0 t'$$
$$y = y'$$
$$t = t' \tag{21}$$

The elastic solution can be obtained by means of the Helmholtz resolution, by which the displacements u and v in x', and y'-directions are given by

$$u = \frac{-\partial\phi}{\partial x'} + \frac{\partial\psi}{\partial y'}$$
$$v = \frac{-\partial\phi}{\partial y'} \frac{-\partial\psi}{\partial x'} \tag{22}$$

where ϕ and ψ are displacement potentials, which when substituted into the equations of motion give

$$(K + 4/3\, G)\nabla^2\phi = \rho \frac{\partial^2\phi}{\partial t'^2} \tag{23}$$

$$G\nabla^2\psi = \rho \frac{\partial^2\psi}{\partial t'^2} \tag{24}$$

The coordinate transformation (21) is then used, and steady-state conditions are obtained by requiring $\partial/\partial_t = 0$. Thus

$$(K + 4/3\, G)\nabla^2\phi = \rho V_0^2 \frac{\partial^2\phi}{\partial x^2} \tag{25}$$

$$G\nabla^2\psi = \rho V_0^2 \frac{\partial^2\psi}{\partial x^2} \tag{26}$$

with boundary conditions

$$\sigma_{yy}(x, 0) = -P_0 H(x) \tag{27}$$

$$\sigma_{xy}(x, 0) = 0 \quad \text{all} \quad x \tag{28}$$

Making the further substitution $\xi = x/V_0$ (for later convenience) and taking Laplace transforms with respect to ξ (note that $\phi = \psi = 0$ for $\xi < 0$) gives

$$\frac{d^2\bar{\phi}}{dy^2} - \frac{s^2}{V_0^2} m_1^2 \bar{\phi} = 0 \tag{29}$$

$$\frac{d^2\bar{\psi}}{dy^2} - \frac{s^2}{V_0^2} m_2^2 \bar{\psi} = 0 \tag{30}$$

where the definitions

$$C_1^2 = (K + 4/3\, G)/\rho \tag{31}$$

$$C_2^2 = G/\rho \tag{32}$$

$$m_1^2 = V_0^2/C_1^2 - 1 \tag{33}$$

$$m_2^2 = V_0^2/C_2^2 - 1 \tag{34}$$

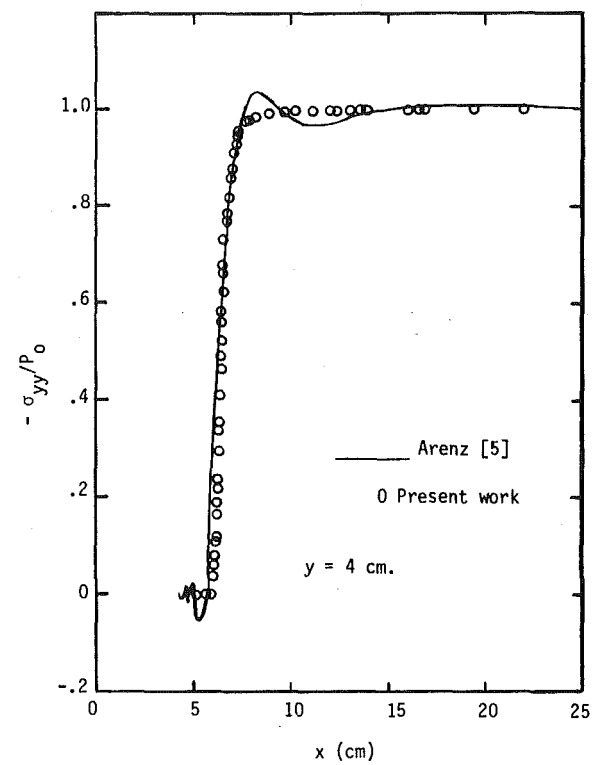


Fig. 9 Comparison of calculations for normal stress response in viscoelastic half space

have been used. Solving these equations, using the radiation conditions, the transform of the boundary conditions gives finally

$$\bar{\phi} = \frac{P_0 V_0^2}{G s^3} \frac{T}{R} \exp(-m_1 s y/V_0) \tag{35}$$

$$\bar{\psi} = \frac{-P_0 V_0^2}{G s^3} \frac{2m_1}{R} \exp(-m_2 s y/V_0) \tag{36}$$

where

$$T \equiv m_2^2 - 1 \tag{37}$$

$$R \equiv T^2 + 4 m_1 m_2 \tag{38}$$

The stresses are related to the displacement potentials through the displacement-potential, strain-displacement, and stress-strain relations. Thus, for example,

$$\sigma_{yy} = \frac{-P_0}{s}\left\{\frac{T^2}{R} \exp(-m_1 s y/V_0) + 4\frac{m_1 m_2}{R} \exp(-m_2 s y/V_0)\right\} \tag{39}$$

The transform of a viscoelastic solution can then be obtained by the usual material property substitution.

The viscoelastic material property substitution is made as follows. The elastic shear modulus G is replaced by the inverse of the elastic shear compliance $G = 1/J$. J is then replaced by $s\,\bar{J}_{\text{crp}}(s)$ where $\bar{J}_{\text{crp}(s)}$ is the transform of equation (15), given by

$$\bar{J}_{\text{crp}}(s) = J_g/s + \sum_{i=1}^{n} J_i a_i/[s(s + a_i)] \tag{40}$$

The bulk modulus K was taken as a constant, $K = 2.07$ Mn/m^2, and $\rho = 1077$ K$_g$/m^3, in order to match the solution of Arenz.

Arenz presents a detailed plot of the normal stress σ_{yy} due to the dilatational wave, which corresponds to the inverse of the first term of equation (39). Thus the first term of equation (39) has been numerically inverted, using techniques previously described. The comparison with Arenz's results is shown in Fig. 9.

Perhaps the most noticeable feature of the comparison of results shown in Fig. 9 is the good agreement of the two methods. In detail, however, Arenz's results show some of the oscillation about the wave front seen previously in the rod wave problems, although not nearly to the same degree. This oscillation is completely lacking in the present results.

Discussion

The preceding examples have attempted to show that numerical inversion of the Laplace transform in problems of dynamic viscoelasticity is a simple and reasonably accurate procedure. The accuracy has been assessed primarily by means of comparison with the numerical viscoelastic rod wave propagation solutions of Knauss [6], and some simple exact functions that have properties at least somewhat similar to that of viscoelastic waves.

It is worthwhile to repeat the argument given earlier by Arenz [4] to the effect that simple models cannot represent realistic viscoelastic materials. However with the techniques discussed in the present work there is no difficulty in obtaining accurate material property representations just by using a sufficient number of terms in an exponential series, and inverting the Laplace transforms numerically. In fact the broad transition range characteristic of many viscoelastic materials appears to enhance the accuracy of transform inversion in dynamic problems, by smoothing the waves through material dispersion.

It would seem that considerable extension of the present technique for transform inversion would be possible, and work in this direction has been presented by Bellman [8]. Perhaps what is most surprising is the accuracy that is apparently obtained by using, say, a 15 term polynomial series to represent an arbitrary function. However it should be noted that while the problems considered here had solutions much less smooth than is characteristic of many quasi-static problems, and thus presumably place more stringent demands on the numerical transform inversion, certainly less smooth problems may yet be envisioned. Thus some modification of present procedures may be necessary. But for problems of the type considered here, numerical Laplace transform inversion would seem to be little more difficult than say numerical quadrature.

Summary and Conclusions

Laplace transform techniques greatly simplify many problems in linear viscoelasticity, reducing the difficulty to that of an elastic problem. However, adequate viscoelastic material property representations make transform inversion difficult. Although approximate inversion techniques have been widely employed in quasi-static viscoelasticity problems, inaccuracy and distortion of the solution have been noted in applications of these techniques to wave propagation problems in viscoelastic materials.

The present work has attempted to show that one of the numerical transform inversion techniques of Bellman [8] can be used in dynamic viscoelasticity. The procedure presented here is simple to use and can be readily implemented, and gives good accuracy. Comparisons with the work of Knauss [6] for wave propagation in a viscoelastic rod appear to be within plotting accuracy. Comparisons of numerical and exact inversion of some simple functions indicate accuracy of about ± 1 percent, and probably this accuracy can be achieved in viscoelastic wave propagation problems. Undershoot and overshoot of a step function is about 5 percent.

References

1 Christensen, R. M., *Theory of Viscoelasticity,* Academic Press 1971.

2 Cost, T. L., "Approximate Laplace Transform Inversions in Viscoelastic Stress Analysis," *AIAA Journal,* Vol 2, 1964, pp. 2157–2166.

3 Schapery, R. A., "Approximate Methods of Transform Inversion for Viscoelastic Stress Analysis," *Proceedings of the 4th U. S. National Congress on Applied Mechanics,* Vol. 2, 1962, pp. 1075–1085.

4 Arenz, R. J., "Uniaxial Wave Propagation in Realistic Viscoelastic Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, 1964, pp. 17–21.

5 Arenz, R. J., "Two-Dimensional Wave Propagation in Realistic Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, 1965, pp. 303–314.

6 Knauss, W. G., "Uniaxial Wave Propagation in a Viscoelastic Material Using Measured Material Properties," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, 1968, pp. 449–453.

7 Sackman, J. L., and Kaya, I., "On the Propagation of Transient Pulses in Linearly Viscoelastic Media," *Journal of the Mechanics and Physics of Solids,* Vol. 16, 1968, pp. 349–356.

8 Bellman, R., Kalaba, R. E., and Lockett, J., *Numerical Inversion of the Laplace Transform,* Amer. Elsevier Publ. Co., 1966.

9 Tsay, L. J., "Numerical Computation of Spherical Waves in Two Viscoelastic Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, June 1972, pp. 401–406.

10 Arenz, R. J., et al., The Mechanical and Optical Characterization of HYSOL 8705 With Application to Photoviscoelastic Analysis, California Institute Technical Report WL TDR-64-4, June 1964.

11 Cole, J., and Huth, J., "Stresses Produced in a Half Plane by Moving Loads," ASME JOURNAL OF APPLIED MECHANICS, Vol. 25, 1958, pp. 433–436.

W. C. Johnson
Graduate Research Assistant,
Department of Metallurgical Engineering,
Michigan Technological University,
Houghton, Mich. 49931

Y. Y. Earmme
Assistant Professor,
Department of Mechanical Science,
Korea Advanced Institute of Science,
P.O. Box 150,
Cheong Ryang Ri,
Seoul, Korea

J. K. Lee
Associate Professor,
Department of Metallurgical Engineering,
Michigan Technological University,
Houghton, Mich. 49931

Approximation of the Strain Field Associated With an Inhomogeneous Precipitate

Part 1: Theory

Two independent methods are derived for the calculation of the elastic strain field associated with a coherent precipitate of arbitrary morphology that has undergone a stress-free transformation strain. Both methods are formulated in their entirety for an isotropic system. The first method is predicated upon the derivation of an integral equation from consideration of the equations of equilibrium. A Taylor series expansion about the origin is employed in solution of the integral equation. However, an inherently more accurate means is also developed based upon a Taylor expansion about the point of which the strain is to be calculated. Employing the technique of Moschovidis and Mura, the second method extends Eshelby's equivalency condition to the more general precipitate shape where the constrained strain is now a function of position within the precipitate. An approximate solution to the resultant system of equations is obtained through representation of the equivalent stress-free transformation strain by a polynomial series. For a given order of approximation, both methods reduce to the determination of the biharmonic potential functions and their derivatives.

Introduction

Understanding of the elastic strain energy associated with an ellipsoidal precipitate has progressed rapidly during the past two decades, principally due to the classical works of Eshelby [1–3]. He considered an ellipsoidal precipitate which undergoes a uniform stress-free transformation strain. Both precipitate and matrix were taken to be elastically isotropic. Subsequently, Willis [4], Kinoshita and Mura [5], and Asaro and Barnett [6] proved or extended to the anisotropic case Eshelby's theorem that the stress field inside an ellipsoidal precipitate is constant when the stress-free transformation strain is constant. Robinson [7], Barnett, et al. [8], and Shibata and Ono [9] calculated strain energies in the isotropic case, while Lee, et al. [10], obtained anisotropic strain energies of ellipsoidal precipitates.

Although the ellipsoid is a versatile shape, other precipitate shapes such as the cube and the rectangular parallelepiped are often observed in solid-solid transformations [11], thus requiring evaluation of the stress fields and strain energies of these mathematically more complicated morphologies. The solutions for an inhomogeneous ellipsoidal precipitate, i.e., an ellipsoidal inclusion whose elastic constants are different from those of the matrix, are based on the fact that the stress in the ellipsoidal precipitate is uniform. However, the stress field inside a nonellipsoidal precipitate is not uniform even if the stress-free transformation is a pure dilatation. Because of this complexity, no exact solution has been obtained for the strain field associated with a nonellipsoidal inhomogeneous precipitate.

In this study we will formulate two distinct methods for calculation of the strain field produced by a coherent inhomogeneous precipitate of arbitrary shape that has undergone a stress-free transformation strain. First, the technique of Moschovidis and Mura [12] will be used to extend Eshelby's ellipsoidal equivalency condition to the general inhomogeneity. The second method will be based upon derivation of an integral equation in the manner of Chen and Young [13] who first employed a similar technique in arriving at the strain field associated with an inhomogeneity embedded in a uniform strain field. In the subsequent paper, Part 2, we will apply these two techniques to the case of a cuboidal inhomogeneity within an isotropic matrix. The several modes of solution developed in Part 1 will be contrasted in Part 2 on the basis of ease and accuracy of solution.

The Modified Equivalency Method

Theory. Since the interior strain field associated with an ellipsoidal precipitate is constant for a constant stress-free transformation strain, Eshelby [3] was able to arrive at an equation that allowed an inhomogeneous ellipsoidal precipitate that had undergone a constant stress-free transformation strain, e_{ij}^{T*}, to be represented as an

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, May, 1980.

equivalent inclusion. The equivalent inclusion would result in the exact strain field as did the original inhomogeneity, yet would possess the same elastic constants as the matrix phase. Such representation permits direct computation of the strain field through the homogeneous Green's function. The equivalency condition for ellipsoids, as formulated by Eshelby is

$$C^*_{ijkl}(e^c_{kl} - e^{T*}_{kl}) = C_{ijkl}(e^c_{kl} - e^T_{kl}) \tag{1}$$

where e^T_{ij} is the equivalent stress-free transformation strain, e^c_{ij} is the constrained strain and the C_{ijkl} represent the elastic constants of the matrix and precipitate (the asterisk denotes precipitate) phases. The usual suffix notation is also employed. Repeated indices are to be summed over the values 1, 2, and 3 and suffixes following a comma denote differentiation with respect to the Cartesian coordinates x_1, x_2, and x_3, i.e., $u_{i,j} = \partial u_i/\partial x_j$.

For an arbitrary precipitate within an infinite matrix, the constrained strain within the precipitate is not a constant. Hence, equation (1) becomes intractable in its present form. Equation (1) can be modified for use with the general precipitate morphology by employing a technique first developed by Moschovidis and Mura [12] in their work with ellipsoids immersed in a uniform stress field. Adopting some of their notation, allow the equivalent stress-free transformation strain of an arbitrarily shaped precipitate to be represented as a polynomial, $\beta_{ij}(\mathbf{x})$, where

$$\beta_{ij}(\mathbf{x}) = B_{ij} + B_{ijk}x_k + B_{ijkl}x_kx_l + \ldots. \tag{2}$$

The coefficients, $B_{ij\ldots m}$, are symmetric with respect to i and j since the strain is symmetric in i and j. By simple permutation they are also symmetric in $k, l \ldots m$.

The constrained strain would likewise become a function of the position coordinate. Realizing this, Equation (1) can be rewritten for the general precipitate as

$$C^*_{ijkl}[e^c_{kl}(\mathbf{x}) - e^{T*}_{kl}] = C_{ijkl}[e^c_{kl}(\mathbf{x}) - \beta_{kl}(\mathbf{x})] \tag{3}$$

where equation (3) must be satisfied for *all* points within the precipitate. This modified equivalency condition cannot be solved directly, for equation (3) produces only nine independent equations and yet, due to $\beta_{kl}(\mathbf{x})$, consists of an infinite number of unknowns. Such difficulty necessitates formulation of an approximation scheme, which allows the coefficients of $\beta_{kl}(\mathbf{x})$ to be calculated to the desired degree of accuracy. In order to proceed further along such lines, it is first necessary to express the constrained strain in terms of the general stress-free transformation strain, $\beta_{ij}(\mathbf{x})$.

Determination of $e^c_{ij}(\mathbf{x})$. Assume an infinite matrix containing an arbitrarily shaped coherent inclusion that has undergone a stress-free transformation strain given by $\beta_{ij}(\mathbf{x})$, where $\mathbf{x}$ is the position vector with respect to a Cartesian coordinate system whose origin is located within the precipitate. The constrained displacement field would then be given by [1]

$$u^c_i(\mathbf{x}) = -\int\int_V\int C_{jlmn}\beta_{mn}(\mathbf{x}')G_{ij,l}(\mathbf{x}-\mathbf{x}')dV(\mathbf{x}') \tag{4}$$

where the integral is taken over the entire precipitate and $G_{ij}(\mathbf{x}-\mathbf{x}')$ is the elastic Green's function. If we assume an isotropic matrix, the Green's function can be expressed as

$$G_{ij}(\mathbf{x}-\mathbf{x}') = \frac{1}{8\pi\mu}\left\{\delta_{ij}\nabla^2 - \frac{1}{2(1-\nu)}\frac{\partial^2}{\partial x_i\partial x_j}\right\}|\mathbf{x}-\mathbf{x}'| \tag{5}$$

where μ and ν are shear modulus and Poisson's ratio, respectively, for the matrix phase, δ_{ij} is the Kronecker delta function and ∇^2 is the Laplacian. Substitution of equation (5) into equation (4), realizing that $\beta_{ij}(\mathbf{x}) = \beta_{ji}(\mathbf{x})$ and $C_{jlmn} = \lambda\delta_{jl}\delta_{mn} + \mu(\delta_{jm}\delta_{ln} + \delta_{jn}\delta_{lm})$ gives

$$u^c_i(\mathbf{x}) = -\frac{1}{8\pi}\left\{\frac{\nu}{(1-\nu)}\Psi_{mm,aai}(\mathbf{x}) - \frac{1}{(1-\nu)}\Psi_{jl,ijl}(\mathbf{x}) + \Psi_{il,aal}(\mathbf{x}) + \Psi_{li,aal}(\mathbf{x})\right\} \tag{6}$$

where

$$\Psi_{ij}(\mathbf{x}) = \int\int_V\int |\mathbf{x}-\mathbf{x}'|\beta_{ij}(\mathbf{x}')dV(\mathbf{x}'). \tag{7}$$

Differentiating equation (6) with respect to x_j, the constrained strain expressed in terms of $\beta_{kl}(\mathbf{x})$ *is*

$$e^c_{ij}(\mathbf{x}) = \frac{-1}{8\pi}\left\{\frac{\nu}{(1-\nu)}\delta_{kl}\Psi_{kl,aaij}(\mathbf{x}) - \frac{1}{(1-\nu)}\Psi_{kl,ijkl}(\mathbf{x}) + \frac{1}{2}(\delta_{kj}\Psi_{kl,aail}(\mathbf{x}) + \delta_{jl}\Psi_{kl,aaki}(\mathbf{x}) + \delta_{ik}\Psi_{kl,aalj}(\mathbf{x}) + \delta_{il}\Psi_{kl,aakj}(\mathbf{x}))\right\}. \tag{8}$$

Equation (8) presents a means by which the constrained strain can be expressed in terms of the equivalent stress-free transformation strain. It may be simplified by the introduction of a coupling matrix [14] between $e^c_{ij}(\mathbf{x})$ and $\beta_{ij}(\mathbf{x})$, that is by expressing $e^c_{ij}(\mathbf{x})$ as

$$e^c_{ij}(\mathbf{x}) = D_{ijkl}(\mathbf{x})B_{kl} + D_{ijklm}(\mathbf{x})B_{klm} + D_{ijklmn}(\mathbf{x})B_{klmn} + \ldots \tag{9}$$

The respective $D_{ij\ldots k}$ can be determined simply by substitution of equations (2) and (7) into equation (8). If the resulting terms are separated on the basis of the $B_{kl\ldots m}$ and the following definitions are adopted concerning the biharmonic potential functions:

$$\psi_{ij\ldots k} \equiv \psi_{ij\ldots k}(\mathbf{x}) = \int\int_V\int x'_ix'_j\ldots x'_k|\mathbf{x}-\mathbf{x}'|dV(\mathbf{x}') \tag{10}$$

then

$$D_{ijkl}(\mathbf{x}) = -\frac{1}{8\pi}\left\{\frac{\nu}{(1-\nu)}\delta_{kl}\psi_{,aaij} - \frac{1}{(1-\nu)}\psi_{,ijkl} + \frac{1}{2}[\delta_{jk}\psi_{,aail} + \delta_{jl}\psi_{,aaki} + \delta_{ik}\psi_{,aalj} + \delta_{il}\psi_{,aakj}]\right\}$$

$$D_{ijklm}(\mathbf{x}) = -\frac{1}{8\pi}\left\{\frac{\nu}{(1-\nu)}\delta_{kl}\psi_{m,aaij} - \frac{1}{(1-\nu)}\psi_{m,ijkl} + \frac{1}{2}[\delta_{jk}\psi_{m,aail} + \delta_{jl}\psi_{m,aaki} + \delta_{ik}\psi_{m,aalj} + \delta_{il}\psi_{m,aakj}]\right\}, \text{ etc.} \tag{11}$$

Notice that the D's are symmetric in i and j and in k and l. Also, as originally shown by Eshelby [1], the D's for an ellipsoid would be constant. With equations (9) and (11) the constrained strain can be represented in terms of the stress-free transformation strain and can now be utilized in the solution of the modified equivalency condition.

Solution of the Modified Equivalency Condition. To solve the modified equivalency condition, it has been shown that it is necessary to arrive at some sort of approximation scheme [12]. One straight forward method is through application of a Taylor series. First, rewrite equation (3) as

$$\Delta C_{ijkl}e^c_{kl}(\mathbf{x}) + C_{ijkl}\beta_{kl}(\mathbf{x}) = C^*_{ijkl}e^{T*}_{kl} \tag{12}$$

where

$$\Delta C_{ijkl} = C^*_{ijkl} - C_{ijkl}.$$

The constrained strain can be expressed in a Taylor series expanded about the origin as

$$\begin{aligned} e^c_{ij}(\mathbf{x}) = {} & D^\circ_{ijkl}B_{kl} + D^\circ_{ijklm}B_{klm} + D^\circ_{ijklmn}B_{klmn} + \ldots \\ & + x_p\{D^\circ_{ijkl,p}B_{kl} + D^\circ_{ijklm,p}B_{klm} + D^\circ_{ijklmn,p}B_{klmn} + \ldots\} \\ & + \tfrac{1}{2}x_px_q\{D^\circ_{ijkl,pq}B_{kl} + D^\circ_{ijklm,pq}B_{klm} \\ & + D^\circ_{ijklmn,pq}B_{klmn} + \ldots\} + \ldots \end{aligned} \tag{13}$$

where the superscript zero implies evaluation at the origin.

Now substitute equation (13) along with the expanded version of $\beta_{kl}(\mathbf{x})$ back into equation (12). Equating like powers of x_i and comparing coefficients, thereby requiring that the equivalency condition is satisfied for *all* points within the precipitate, the following system of equations is realized:

$$(a) \quad \Delta C_{ijkl}\{D^{\circ}_{klmn}B_{mn} + D^{\circ}_{klmnp}B_{mnp} + D^{\circ}_{klmnpq}B_{mnpq} + \ldots\} + C_{ijkl}B_{kl} = C^{*}_{ijkl}e^{T*}_{kl}$$

$$(b) \quad \Delta C_{ijkl}\{D^{\circ}_{klmn,s}B_{mn} + D^{\circ}_{klmnp,s}B_{mnp} + D^{\circ}_{klmnpq,s}B_{mnpq} + \ldots\} + C_{ijkl}B_{kls} = 0$$

$$(c) \quad \frac{1}{2}\Delta C_{ijkl}\{D^{\circ}_{klmn,st}B_{mn} + D^{\circ}_{klmnp,st}B_{mnp} + D^{\circ}_{klmnpq,st}B_{mnpq} + \ldots\} + C_{ijkl}B_{klst} = 0, \text{ etc.} \quad (14)$$

Equation (14) refers only to a constant stress-free transformation strain but the more general case of a nonuniform $e^{T*}_{ij}(\mathbf{x})$ can be handled by expressing $e^{T*}_{ij}(\mathbf{x})$ in terms of a polynomial and then equating like powers of x_i. It should be reiterated that the aforementioned method is applicable to any precipitate morphology provided the biharmonic potential functions and their derivatives are known, that is, the $D_{ij\ldots k}(\mathbf{x})$ can be calculated. In principle this can always be achieved, if not analytically, then through numerical techniques. The method is also employable in an anisotropic matrix through the expressions for the $D_{ij\ldots k}(\mathbf{x})$ would be much more complex.

In solving equation (14) it becomes necessary to determine the degree of precision that is required. As an illustration assume that a second-order approximation is desired, that is one must find the B_{kl}, B_{klm}, and B_{klmn} for a parabolic representation of the equivalent stress-free transformation strain, all other higher-order coefficients being taken as zero. Such a solution entails solving for 117 coefficients. The corresponding system of independent equations is obtainable from equation (14). Part (a) would generate nine independent equations, part (b) gives 27 equations with the final 81 independent equations being derived from part (c). Taken collectively this gives the necessary 117 simultaneous equations for the determination of the parabolic coefficients. Any order of approximation can likewise be determined though for each increase in the degree of precision the number of independent equations and the corresponding potential functions needed increases tremendously. Once $\beta_{kl}(\mathbf{x})$ has been approximated, the constrained strain of the equivalent inclusion corresponding to the inhomogeneity in question can be calculated by employing equation (9).

The Integral Equation Method

Theory. Another means of calculating the strain field of an inhomogeneous precipitate as engendered by a stress-free transformation strain, is through the derivation of an integral equation. A similar method has been employed by Chen and Young [13] in their study of the strain field associated with an inhomogenity embedded in an applied field. The integral equation method is distinct from the modified equivalency condition and hence the two methods may be used as checks upon one another. The starting point for the integral equation is the equation of equilibrium which, for the matrix and precipitate, can be represented as

$$\sigma^{c}_{kl,k}(\mathbf{x}) = 0 \quad \text{if} \quad \mathbf{x} \text{ in matrix}$$

$$\sigma^{c}_{kl,k}(\mathbf{x}) - \sigma^{T*}_{kl,k} = 0 \quad \text{if} \quad \mathbf{x} \text{ in precipitate.} \quad (15)$$

Since concern is focused on the case of a constant stress-free transformation strain, the two equations of equation (15) may be combined to give

$$C_{ijkl}u^{c}_{i,jk}(\mathbf{x}) = C_{ijkl}u^{c}_{i,jk}(\mathbf{x}) - C_{ijkl}(\mathbf{x})u^{c}_{i,jk}(\mathbf{x}) \quad (16)$$

where

$$C_{ijkl}(\mathbf{x}) = C_{ijkl} \quad \text{if} \quad \mathbf{x} \text{ in matrix}$$

$$C_{ijkl}(\mathbf{x}) = C^{*}_{ijkl} \quad \text{if} \quad \mathbf{x} \text{ in precipitate.}$$

Now define

$$D_{li} = C_{ijkl}\frac{\partial^2}{\partial x_k \partial x_j}$$

and write equation (16) as

$$D_{li}u^{c}_{i} = [C_{ijkl} - C_{ijkl}(\mathbf{x})]\partial_k\partial_j u^{c}_{i}, \quad (17)$$

where henceforth

$$\partial_k\partial_j \equiv \frac{\partial}{\partial x_k}\frac{\partial}{\partial x_j}, \quad \partial'_k \equiv \frac{\partial}{\partial x'_k} \quad (18)$$

and a comma still denotes differentiation with respect to x_i. The Green's function, $G_{im}(\mathbf{x} - \mathbf{x}')$ is defined as follows:

$$D_{li}G_{im}(\mathbf{x} - \mathbf{x}') + \delta_{lm}\delta(\mathbf{x} - \mathbf{x}') = 0$$

where $\delta(\mathbf{x})$ is the Dirac delta function. As pointed out by Barnett [15] an expression for the constrained displacement can be obtained by multiplying equation (17) by $G_{im}(\mathbf{x} - \mathbf{x}')$, multiplying equation (18) by u^{c}_{l}, subtracting equation (18) from equation (17) and then integrating the result over all space yielding

$$u^{c}_{m}(\mathbf{x}') = \iiint_{\infty} [G_{ml}(\mathbf{x} - \mathbf{x}')D_{li}u^{c}_{i} - u^{c}_{l}D_{li}G_{im}(\mathbf{x} - \mathbf{x}')]dV(\mathbf{x}) + \iiint_{\infty} [C_{ijkl}(\mathbf{x}) - C_{ijkl}]G_{ml}(\mathbf{x} - \mathbf{x}')u^{c}_{i,kj}dV(\mathbf{x}) \quad (19)$$

Since in the matrix $C_{ijkl}(\mathbf{x}) = C_{ijkl}$, the second integral in the foregoing reduces to an integration over the precipitate volume, V. The first integral of equation (19) can be simplified by considering the following two relations:

$$G_{ml}(\mathbf{x} - \mathbf{x}')u^{c}_{i,kj} - u^{c}_{l}G_{im,kj}(\mathbf{x} - \mathbf{x}') = \partial_k[G_{ml}(\mathbf{x} - \mathbf{x}')u^{c}_{i,j} - u^{c}_{l}G_{im,j}(\mathbf{x} - \mathbf{x}')] - G_{ml,k}(\mathbf{x} - \mathbf{x}')u^{c}_{i,j} + u^{c}_{l,k}G_{im,j}(\mathbf{x} - \mathbf{x}')$$

and

$$C_{ijkl}G_{ml,k}(\mathbf{x} - \mathbf{x}')u^{c}_{i,j} = C_{lkji}G_{mi,j}(\mathbf{x} - \mathbf{x}')u^{c}_{l,k}. \quad (20)$$

Substituting equations (20) into equation (19) and realizing that $G_{im}(\mathbf{x} - \mathbf{x}') = G_{mi}(\mathbf{x} - \mathbf{x}')$, the constrained displacement can be written

$$u^{c}_{m}(\mathbf{x}') = \iiint_{V} \Delta C_{ijkl}G_{ml}(\mathbf{x} - \mathbf{x}')u^{c}_{i,kj}dV(\mathbf{x}) + C_{ijkl}\iiint_{\infty} \partial_k[G_{ml}(\mathbf{x} - \mathbf{x}')u^{c}_{i,j} - u^{c}_{l}G_{im,j}(\mathbf{x} - \mathbf{x}')]dV(\mathbf{x}). \quad (21)$$

Invoking the divergence theorm and assuming the displacement at inifnity is zero, the second integral of equation (21) becomes

$$C_{ijkl}\iiint_{\infty} \partial_k[G_{ml}(\mathbf{x} - \mathbf{x}')u^{c}_{i,j} - u^{c}_{l}G_{im,j}(\mathbf{x} - \mathbf{x}')]dV(\mathbf{x}) = C_{ijkl}\iint_{S} [G_{ml}(\mathbf{x} - \mathbf{x}')\hat{u}^{c}_{i,j} - \hat{u}^{c}_{l}G_{im,j}(\mathbf{x} - \mathbf{x}')]n_k dS - C_{ijkl}\iint_{S} [G_{ml}(\mathbf{x} - \mathbf{x}')\hat{u}^{c}_{i,j} - \hat{u}^{c}_{l}G_{im,j}(\mathbf{x} - \mathbf{x}')]n_k dS \quad (22)$$

where n_k is the unit vector pointing outward from the precipitate, $\hat{u}^{c}_{l}$ and $\hat{\hat{u}}^{c}_{l}$ are the displacement fields on the matrix and precipitate sides, respectively, of the precipitate-matrix boundary, S. There must be a continuity of the constrained displacement across the precipitate-matrix interface hence, $\hat{u}^{c}_{i} = \hat{\hat{u}}^{c}_{i}$. The traction at any point can be given by

$$\hat{\sigma}_{kl}n_k = C_{ijkl}\hat{u}^{c}_{i,j}n_k$$

and

$$\hat{\hat{\sigma}}^{c}_{kl}n_k - \sigma^{T*}_{kl}n_k = C^{*}_{ijkl}\hat{\hat{u}}^{c}_{i,j}n_k - \sigma^{T*}_{kl}n_k \quad (23)$$

where $\hat{\sigma}_{kl}$ is the stress in the matrix and $\hat{\hat{\sigma}}^{c}_{kl}$ is the constrained stress within the precipitate. Since the traction must be continuous across the precipitate-matrix boundary, equations (23) must be equal on S. Substituting equations (22) and (23) back into equation (21) yields

$$u^{c}_{m}(\mathbf{x}') = \iiint_{V} \Delta C_{ijkl}G_{ml}(\mathbf{x} - \mathbf{x}')u^{c}_{i,jk}dV(\mathbf{x}) - \iint_{S} \Delta C_{ijkl}\hat{u}^{c}_{i,j}G_{ml}(\mathbf{x} - \mathbf{x}')n_k dS \quad (24)$$

$$+ \sigma_{kl}^{T*} \iint_S G_{ml}(\mathbf{x}-\mathbf{x}')n_k dS. \quad (24)$$

(*Cont.*)

Using the divergence theorm on the second integral of equation (24) gives

$$\Delta C_{ijkl} \iint_S \hat{u}_{i,j}^c G_{ml}(\mathbf{x}-\mathbf{x}')n_k dS$$

$$= \Delta C_{ijkl} \iiint_V \partial_k[u_{i,j}^c G_{ml}(\mathbf{x}-\mathbf{x}')]dV(\mathbf{x})$$

$$= \Delta C_{ijkl} \iiint_V [G_{ml,k}u_{i,j}^c + u_{i,jk}^c G_{ml}(\mathbf{x}-\mathbf{x}')]dV(\mathbf{x}). \quad (25)$$

Substituting equation (25) into equation (24) and using the divergence theorm on the last integral of equation (24) yields for the displacement

$$u_m^c(\mathbf{x}') = -\Delta C_{ijkl} \iiint_V G_{ml,k}(\mathbf{x}-\mathbf{x}')u_{i,j}^c dV(\mathbf{x})$$

$$+ \sigma_{kl}^{T*} \iiint_V G_{ml,k}(\mathbf{x}-\mathbf{x}')dV(\mathbf{x}) \quad (26)$$

since

$$\partial_n' G_{im}(\mathbf{x}-\mathbf{x}') = -\partial_n G_{im}(\mathbf{x}-\mathbf{x}').$$

If the integration is taken over the primed coordinate system, equation (26) becomes

$$u_i^c(\mathbf{x}) = \Delta C_{ijkl} \iiint_V G_{ml,k}(\mathbf{x}-\mathbf{x}')\partial_j' u_i^c(\mathbf{x}')dV(\mathbf{x}')$$

$$- \sigma_{kl}^{T*} \iiint_V G_{ml,k}(\mathbf{x}-\mathbf{x}')dV(\mathbf{x}'). \quad (27)$$

The integral equation of equation (27) gives us a method for determining the constrained displacement associated with a precipitate of general morphology in arbitrary anisotropy

Solution of the Integral Equation. Of the several means available for numerically solving the integral equation [13] perhaps one of the easiest and most direct is through the use of a Taylor series expansion of the constrained displacement about the origin, that is, let

$$u_i^c(\mathbf{x}) = \sum_{l=0}^{\infty} \frac{1}{l!} U_{i,p_1p_2\ldots p_l}^{\circ} x_{p_1}x_{p_2}\ldots x_{p_l} \quad (28)$$

where

$$U_{i,p_1p_2\ldots p_l}^{\circ} = \partial_{p_1}\partial_{p_2}\ldots\partial_{p_l}u_i^c(\mathbf{x})|_{\mathbf{x}=0}. \quad (29)$$

The strain at an arbitrary point $\mathbf{x}$ within the precipitate would then be given by

$$e_{ij}^c(\mathbf{x}) = \frac{1}{2}\sum_{n=0}^{\infty}\{U_{i,jp_1p_2\ldots p_n}^{\circ} + U_{j,ip_1p_2\ldots p_n}^{\circ}\}x_{p_1}x_{p_2}\ldots x_{p_n}. \quad (30)$$

An origin expansion would be incapable of giving the strain at a point outside of the precipitate. Such results are understandable in that the constrained strain is discontinuous at the interface. Equation (30) then necessitates determination of the $U_{i,jp_1\ldots p_n}^{\circ}$, which can be accomplished by the following procedure. Differentiate equation (27) l times with respect to x_p and evaluate the results at the origin;

$$U_{i,p_1p_2\ldots p_l}^{\circ} = \Delta C_{kjmn}\partial_{p_1}\partial_{p_2}\ldots\partial_{p_l}$$

$$\times \iiint_V G_{im,n}(\mathbf{x}-\mathbf{x}')\partial_j' u_k^c(\mathbf{x}')dV(\mathbf{x}')\Big|_{\mathbf{x}=0}$$

$$- \sigma_{nm}^{T*} \iiint_V G_{im,np_1p_2\ldots p_l}(\mathbf{x}-\mathbf{x}')dV(x')\Big|_{\mathbf{x}=0}. \quad (31)$$

Using the elastic Green's function for an isotropic system given by equation (5), the second integral on the right-hand side of equation (31) can be simplified to

$$- \sigma_{nm}^{T*} \iiint_V G_{im,np_1p_2\ldots p_l}(\mathbf{x}-\mathbf{x}')dV(\mathbf{x}')\Big|_{\mathbf{x}=0} = \sigma_{nm}^{T*}\alpha_{imnp_1p_2\ldots p_l}^{\circ} \quad (32)$$

where

$$\alpha_{imnp_1p_2\ldots p_l}^{\circ} = \frac{-1}{8\pi\mu}\left\{\delta_{im}\psi_{,aanp_1p_2\ldots p_l}^{\circ} - \frac{1}{2(1-\nu)}\psi_{,imnp_1p_2\ldots p_l}^{\circ}\right\} \quad (33)$$

and ψ° refers to the biharmonic function, ψ, defined by equation (10) and evaluated at the origin.

Now express $\partial_j' u_k^c(\mathbf{x}')$ in the first integral on the right-hand side of equation (31) in a Taylor series expanded about the origin,

$$\partial_j' u_k^c(\mathbf{x}') = \sum_{s=0}^{\infty}\frac{1}{s!}U_{k,jq_1q_2\ldots q_s}^{\circ} x_{q_1}' x_{q_2}'\ldots x_{q_s}'. \quad (34)$$

Again, the U_k° implies that the coefficients are evaluated at the origin. Substituting equations (32) and (34) back into equation (31) and realizing that the derivatives of U_k° are just constants to be evaluated yields,

$$U_{i,p_1\ldots p_l}^{\circ} = \Delta C_{kjnm}$$

$$\times \iiint_V \partial_n\partial_{p_1}\ldots\partial_{p_l}G_{im}(\mathbf{x}-\mathbf{x}')\Big|_{\mathbf{x}=0}\left\{\sum_{s=0}^{\infty}\frac{1}{s!}U_{k,jq_1\ldots q_s}^{\circ}\right.$$

$$\left. x_{q_1}'\ldots x_{q_s}'\right\}dV(\mathbf{x}') + \sigma_{nm}^{T*}\alpha_{imnp_1\ldots p_l}^{\circ} \quad (35)$$

or upon rearrangement and introduction of a new tensor, T, gives

$$U_{i,p_1\ldots p_l}^{\circ} = \Delta C_{kjnm}\sum_{s=0}^{\infty}\frac{1}{s!}U_{k,jq_1q_2\ldots q_s}^{\circ}T_{im,np_1p_2\ldots p_lq_1q_2\ldots q_s}^{\circ}$$

$$+ \sigma_{nm}^{T*}\alpha_{imnp_1\ldots p_l}^{\circ} \quad (36)$$

where

$$T_{im,np_1p_2\ldots p_lq_1q_2\ldots q_s}^{\circ}$$

$$= \iiint_V \partial_n\partial_{p_1}\ldots\partial_{p_l}G_{im}(\mathbf{x}-\mathbf{x}')x_{q_1}'x_{q_2}'\ldots x_{q_s}'dV(\mathbf{x}')\Big|_{\mathbf{x}=0}. \quad (37)$$

Upon substitution of the isotropic elastic Green's function given by equation (5) into equation (37) the expression for the T's becomes

$$T_{im,np_1\ldots p_lq_1\ldots q_s}^{\circ} = \frac{1}{8\pi\mu}\left\{\delta_{im}\psi_{q_1\ldots q_s,aanp_1\ldots p_l}^{\circ}\right.$$

$$\left. - \frac{1}{2(1-\nu)}\psi_{q_1\ldots q_s,imnp_1\ldots p_l}^{\circ}\right\} \quad (38)$$

where the ψ's are given by equation (10) and the superscript zero refers to evaluation at the origin. Note that here, as in the modified equivalency condition, the solution of the integral equation reduces to the determination of the biharmonic potential functions and their derivatives for an isotropic system and for the precipitate morphology under question.

Equations (36) and (38) now supply the means for arriving at the coefficients employed in the Taylor series expansion. Two approaches immediately present themselves. For the first approach assume that a second-order or parabolic approximation is desired; that is it is necessary to calculate the $U_{k,j}^{\circ}$, $U_{k,jm}^{\circ}$, and $U_{k,jmn}^{\circ}$. The simplest means of doing this is to calculate the different orders of coefficients independently. Returning to equation (36), taking l equal to one and s equal to zero yields the following system of equations,

$$U_{i,p}^{\circ} = \Delta C_{kjnm}U_{k,j}^{\circ}T_{im,np}^{\circ} + \sigma_{nm}^{T*}\alpha_{imnp}^{\circ} \quad (39)$$

which can be solved simultaneously for the $U_{i,p}^{\circ}$. Similarly, l could be taken equal to two and s equal to one in equation (36) and the resultant system of 27 equations solved simultaneously for the $U_{i,pq}^{\circ}$ provided that the $U_{i,p}^{\circ}$ values appearing are taken as those that would be calculated from equation (39). Subsequent to the determination of the $U_{i,p}^{\circ}$ and $U_{i,pq}^{\circ}$ take l equal to three and s equal to two to calculate the $U_{i,pqr}^{\circ}$ in the same manner as just described.

A second and intuitively more accurate process involves a simul-

taneous solution of all 117 coefficients that comprise the second-order approximation, a technique similar to that used in the modified equivalency condition. The system of equations can be generated by taking l equal to one and s equal to two in equation (36) yielding nine independent equations, l equal to two and s equal to two giving 27 equations with the final 81 of the total 117 independent equations arising from the case of l equal to three and s equal to two. Both of the foregoing approximation schemes can be directly extended to encompass any degree of accuracy required.

Taylor Expansion About an Arbitrary Point. To enhance the accuracy of the integral equation method for the solution of the constrained displacement without increasing the number of potential functions needed, a Taylor series expansion can be taken about the point of which the strain is to be calculated. Such a procedure would help eliminate the difficulty of representing a rapidly changing strain field with a finite number of terms by a Taylor series. Following the same format as when computing the displacement with an expansion about the origin we allow

$$u_i^c(\mathbf{x}) = \sum_{l=0}^{\infty} \frac{1}{l!} U^*_{i,p_1p_2\ldots p_l}(x_{p_1} - y^*_{p_1})(x_{p_2} - y^*_{p_2}) \ldots (x_{p_l} - y^*_{p_l}) \quad (40)$$

where $\mathbf{x}$ is the point at which the displacement is to be calculated, $\mathbf{y}^*$ is the point about which the expansion is taken and the $U^*_{i,p_1\ldots p_2}$ are the Taylor coefficients evaluated at the point $\mathbf{y}^*$ which need to be determined. Again differentiate equation (27) with respect to x_p l times. For an expansion about an arbitrary point the constrained displacement $u^c_{k,j}(\mathbf{x})$, would be given as

$$u^c_{k,j}(\mathbf{x}) = \sum_{s=0}^{\infty} \frac{1}{s!} U^*_{k,jq_1q_2\ldots q_s}(x_{q_1} - y^*_{q_1}) \ldots (x_{q_s} - y^*_{q_s}). \quad (41)$$

This allows equation (27), when $u_i^c(\mathbf{x})$ is expanded about an arbitrary point $\mathbf{y}^*$ to be expressed as

$$U^*_{i,p_1p_2\ldots p_l} = \Delta C_{kjnm} \iiint_V G_{im,np_1p_2\ldots p_l}(\mathbf{x}^* - \mathbf{x}')\Big|_{\mathbf{x}=\mathbf{y}^*} \times \left[\sum_{s=0} \frac{1}{s!} U^*_{k,jq_1\ldots q_s}(x'_{q_1} - y^*_{q_1}) \ldots (x'_{q_s} - y^*_{q_s})\right] dV(\mathbf{x}') + \sigma^{T*}_{nm}\alpha^*_{imnp_1p_2\ldots p_l} \quad (42)$$

where α^* has the same definition as α° of equation (32) only that the potential functions are evaluated at the position coordinate $\mathbf{y}^*$ rather than at the origin. Upon rearrangement of equation (42)

$$U^*_{i,p_1\ldots p_l} = \Delta C_{kjnm} \sum_{s=0}^{\infty} \frac{1}{s!} U^*_{k,jq_1\ldots q_s} \times \iiint_V G_{im,np_1\ldots p_l}(\mathbf{x} - \mathbf{x}')\Big|_{\mathbf{x}=\mathbf{y}^*} (x'_{q_1} - y^*_{q_1}) \ldots (x'_{q_s} - y^*_{q_s}) dV(\mathbf{x}') + \sigma^{T*}_{nm}\alpha^*_{imnp_1\ldots p_l} \quad (43)$$

or

$$U^*_{i,p_1\ldots p_l} = \Delta C_{kjnm} \sum_{s=0}^{\infty} \frac{1}{s!} U^*_{k,jq_1\ldots q_s} T^*_{im,np_1\ldots p_lq_1\ldots q_s} + \sigma^{T*}_{nm}\alpha^*_{imnp_1\ldots p_l} \quad (44)$$

where

$$T^*_{im,np_1\ldots p_lq_1\ldots q_s} = \iiint_V \left\{\left[\frac{\delta_{im}}{8\pi\mu}\partial_a\partial_a\partial_n\partial_{p_1}\ldots\partial_{p_l} - \frac{1}{16\pi\mu(1-\nu)}\partial_i\partial_n\partial_m\partial_{p_1}\ldots\partial_{p_l}\right] \Big|\mathbf{x} - \mathbf{x}'\Big|_{\mathbf{x}=\mathbf{y}^*}\right\} (x'_{q_1} - y^*_{q_1}) \ldots (x'_{q_s} - y^*_{q_s}) dV(\mathbf{x}'). \quad (45)$$

To determine the Taylor coefficients, $U^*_{i,p_1\ldots p_l}$, we need only proceed precisely as before to generate 117 equations for the 117 unknowns accruing from a second-order approximation. First, in equation (44), take l equal to one and s equal to two to generate nine independent equations, take l equal to two and s equal to two to yield 27 equations with the final 81 independent equations being determined from the case of l equal to three and s equal to two. The 117 independent equations are listed in tensor form as follows:

$$U^*_{i,p} = \Delta C_{kjnm}\{U^*_{k,j}T^*_{im,np} + U^*_{k,js}T^*_{im,np,s} + \frac{1}{2}U^*_{k,jst}T^*_{im,np,st}\}$$

$$U^*_{i,pq} = \Delta C_{kjnm}\{U^*_{k,j}T^*_{im,npq} + U^*_{k,js}T^*_{im,npq,s} + \frac{1}{2}U^*_{k,jst}T^*_{im,npq,st}\}$$

$$U^*_{i,pqr} = \Delta C_{kjnm}\{U^*_{k,j}T^*_{im,npqr} + U^*_{k,js}T^*_{im,npqr,s} + \frac{1}{2}U^*_{k,jst}T^*_{im,npqr,st}\}. \quad (46)$$

Summary

Two distinct methods have been derived for the calculation of the strain field associated with a coherent inhomogeneity embedded in an isotropic matrix that has undergone a stress-free transformation strain. The first method employs the technique of Moschovidis and Mura [12] in an extension of Eshelby's ellipsoidal equivalency condition to the general precipitate morphology. This is achieved by representing the inhomogeneity as an equivalent inclusion which would produce the identical strain field as did the original precipitate and yet not require knowledge of an inhomogeneous Green's function. The equivalent stress-free transformation strain is represented as a polynomial, the order of which is determined by the desired degree of accuracy. The results are developed completely for an isotropic system and the method reduced to the determination of the biharmonic potential functions and their derivatives for the precipitate shape under study. For a parabolic representation of the equivalent stress-free transformation strain, it is in general necessary to solve 117 simultaneous equations for the corresponding polynomial coefficient.

Inspired by the recent work of Chen and Young [13] in their consideration of inclusions subjected to an applied stress, we have developed a second approach to the misfitting inclusion problem. This method is based upon derivation of an integral equation which arises from consideration of the equations of equilibrium associated with the precipitate-matrix system. Solution of the integral equation was predicated first upon a Taylor series expansion of the constrained strain about the origin. In general, it is again required that a system of 117 independent equations be solved for a parabolic representation of the constrained strain about the origin. Invoking symmetry arguments may, however, substantially reduce such a system of equations. In an effort to circumvent any shortcomings of an origin expansion accruing from its inability to follow a rapidly changing strain field with a finite number of terms, a Taylor expansion centered on the point of interest was also developed. The integral equation and modified equivalency methods and their techniques of solution are contrasted for a cuboidal precipitate in an isotropic matrix in Part 2.

Acknowledgments

The authors are greatly indebted to Prof. D. M. Barnett of Stanford University, who corrected an error in the derivation of the integral equation which appeared in the original manuscript. They also wish to thank Prof. T. Mura for his helpful discussion of this work. This work was supported by the Division of Materials Research of the National Science Foundation under Grants DMR76-06855 A01 and 79-01693 for which much appreciation is expressed.

References

1 Eshelby, J. D., "Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proceedings of the Royal Society*, Vol. 241, Series A, 1957, pp. 376–396.

2 Eshelby, J. D., "The Elastic Field Outside an Ellipsoidal Inclusion," *Proceedings of Royal Society*, Vol. 252, Series A, 1959, pp. 561–569.

3 Eshelby, J. D., "Elastic Inclusions and Inhomogeneities," *Progress in Solid Mechanics*, Vol. 2, 1961, pp. 89–140.

4 Willis, J. R., "The Solution of Asymmetric Problems of Elasticity of Fourier Transforms," Adams Prize Essay, University of Cambridge, 1970, unpublished.

5 Kinoshita, N., and Mura, T., "Elastic Fields of Inclusions in Anisotropic Media," *Phys. Stat. Sol.* Series A, Vol. 5, 1971, pp. 759–768.

6 Asaro, R. J., and Barnett, D. M., "The Nonuniform Transformation Strain Problem for an Anisotropic Ellipsoidal Inclusion" *Journal of the Mechanics and Physics of Solids,* Vol. 23, 1975, pp. 77–83.

7 Robinson, K., "Elastic Energy of an Ellipsoidal Inclusion in an Infinite Solid," *Journal Applied Physics,* Vol. 22, 1951, pp. 1045–1054.

8 Barnett, D. M., et al., "The Strain Energy of a Coherent Ellipsoidal Precipitate," *Scripta Met.,* Vol. 8, 1974, pp. 1447–1450.

9 Shibata, M., and Ono, K., "The Strain Energy of a Spheroidal Inclusion and Its Application to B.C.C.-H.C.P. Martensite Transformation," *Acta Met.,* Vol. 23, 1975, pp. 587–597.

10 Lee, J. K., Barnett, D. M., and Aaronson, H. I., "The Elastic Strain Energy of Coherent Ellipsoidal Precipitates in Anisotropic Crystalline Solids," *Met. Trans.,* Vol. 8A, 1977, pp. 963–970.

11 Ardell, A. J. and Nicholson, R. B., with Appendix by Eshelby, J. D., "On the Modulated Structure of Aged Ni-Al Alloys," *Acta Met.,* Vol. 14, 1966, pp. 1295–1309.

12 Moschovidis, Z. A., and Mura, T., "Two Ellipsoidal Inhomogeneities by the Equivalent Inclusion Method," ASME JOURNAL OF APPLIED MECHANICS, 42, 1975, pp. 847–851.

13 Chen, F. C., and Young, K., "Inclusions of Arbitrary Shape in an Elastic Medium," *Journal of Mathematics and Physics,* Vol. 12, 1977, pp. 1412–1416.

14 Sendeckyj, G., PhD Thesis, Northern University, 1967.

15 Barnett, D. M., Stanford University, Private Communication 1980.

W. C. Johnson
Graduate Research Assistant,
Department of Metallurgical Engineering,
Michigan Technological University,
Houghton, Mich. 49931

Y. Y. Earmme
Assistant Professor,
Department of Mechanical Sciences,
Korea Advanced Institute of Science,
P. O. Box 150,
Cheong Ryang Ri,
Seoul, Korea

J. K. Lee
Associate Professor,
Department of Metallurgical Engineering,
Michigan Technological University,
Houghton, Mich. 49931

Approximation of the Strain Field Associated With an Inhomogeneous Precipitate

Part 2: The Cuboidal Inhomogeneity

The modified equivalency and integral equation methods for determination of the constrained strain field associated with a precipitate that has undergone a dilatational stress-free transformation strain as developed in Part 1, are applied to the case of a cuboidal inhomogeneity within an isotropic matrix. Agreement between the two methods is good for small and moderate differences in the shear moduli between precipitate and matrix. For large differences in the shear moduli, some divergence is observed in that fluctuations in the constrained strain field become quite pronounced near the cube edge and corner when considering the integral equation method. Although some error is inevitable due to the cutoff of higher-order terms in the Taylor series expansion, the modified equivalency method yields fair results under such circumstances. With the latter method, the constrained strain field of a cuboid is analyzed as a function of position and orientation. Although the strain field behaves as expected in the central regions of the cube in that the harder the precipitate the larger the constrained strain, its behavior becomes complicated as the precipitate-matrix interface is approached, demonstrating a strong dependency on precipitate rigidity. As a result, the dilatation in the inhomogeneous cuboidal precipitate is found not to be a constant as contrasted with the homogeneous case.

Introduction

In the previous paper [1], Part 1, two methods are derived for the calculation of the strain field associated with a general inhomogeneous precipitate embedded in an infinite matrix. Due to the inherent difficulties encountered with the anisotropic Green's function, the two methods are formulated only for an isotropic system. Unlike the case of a homogeneous inclusion, an analytical solution is unobtainable and it is necessary to invoke an approximation scheme. In employing the modified equivalency condition it is necessary to express the equivalent stress-free transformation strain as a polynomial in the position coordinates. The coefficients engendered by the polynomial are determined by the solution of a system of equations, the size of which depends upon the degree of accuracy desired. (For example, nine equations for a zeroth-order expansion compared to 117 equations for a parabolic representation.) Once the polynomial coefficients are computed, it is necessary only to calculate a limited number of the biharmonic potential functions at the point of which the strain is to be determined.

The integral equation also presents several different modes of solution. Here the constrained strain is expressed in terms of a polynomial. Different orders of expansion for the constrained strain can be taken about either the origin or about the point of which the strain is to be calculated. Ideally, an origin expansion of sufficient accuracy could be employed once the Taylor coefficients are calculated and the strain then determined at the point of interest through a simple polynomial. Otherwise, in the presence perhaps of a highly changing strain field, the Taylor series could be taken about the individual points of which the strain is to be calculated. In calculation of such strain fields, the relative computation time involved in the different techniques varies drastically with the order of approximation. Hence it becomes desirable to optimize both computation time and accuracy.

The inhomogeneous ellipsoidal transformation problem has been solved for both the isotropic [2–6] and anisotropic [7–10] systems.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, January, 1980.

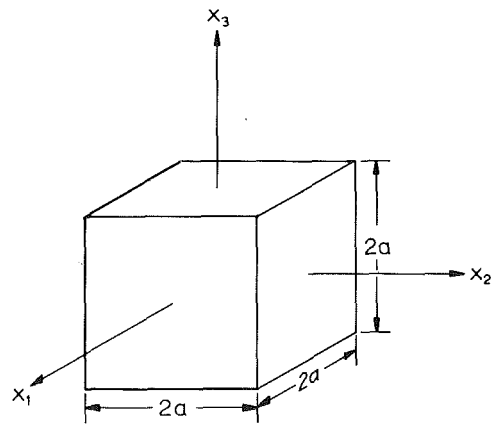


Fig. 1 Schematic drawing of a cuboidal precipitate showing coordinate axes and dimensions

However, precipitates and voids are often found to possess cuboidal [11] or rectangular [12] morphologies. It is the intent of this paper, Part 2, to examine the strain field associated with a cuboidal inhomogeneity that undergoes a purely dilatational stress-free transformation strain while immersed in an isotropic matrix. A comparison of the two techniques is also employed in gauging their relative accuracies. For small changes in the precipitate shear modulus, it is seen that a very simple zeroth-order approximation of the integral equation can be employed to calculate the strain field while large differences in the precipitate and matrix shear moduli require use of the modified equivalency method or higher-order approximations of the integral equation.

Symmetry Properties

A cuboidal precipitate embedded in an isotropic matrix and coupled with a stress-free transformation strain that is a *pure dilatation,* presents a highly symmetric system. It is advantageous to consider such symmetry arguments before directly applying either the modified equivalency or integral equation methods to the cuboid.

As shown in Fig. 1, allow the origin of a Cartesian coordinate system to coincide with the center of the cube with axes perpendicular to the cube faces and edges of length $2a$. Now assume that the constrained displacement arising from the foregoing system can be expressed in terms of a Taylor series expansion about the origin of the coordinate system, that is,

$$u_i^c(\mathbf{x}) = \sum_{l=0}^{\infty} \frac{1}{l!} U^{\circ}_{i,p_1p_2\ldots p_l} x_{p_1} x_{p_2} \ldots x_{p_l} \tag{1}$$

where $u_i^c(\mathbf{x})$ is the constrained displacement and $U^{\circ}_{i,p\ldots q}$ refers to the Taylor coefficients, the superscript zero denoting evaluation at the origin. From the symmetry of the system a large number of the Taylor coefficients are zero. If we employ a symmetry analysis similar to that of Chen and Young [13], relations between the Taylor coefficients can be deduced from consideration of the displacement of an arbitrary point and the displacements at points corresponding to a transformation of that point through planes and positions of symmetry. In Fig. 1, choose an arbitrary point, $\mathbf{x} = (x_1, x_2, x_3)$. By symmetry it follows that the constrained displacement in the x_1-direction,

$$\begin{aligned} u_1^c(x_1, x_2, x_3) &= u_1^c(x_1, -x_2, x_3) \\ &= u_1^c(x_1, x_2, -x_3) = u_1^c(x_1, -x_2, -x_3) \\ &= -u_1^c(-x_1, x_2, x_3) = -u_1^c(-x_1, -x_2, x_3) \\ &= -u_1^c(-x_1, x_2, -x_3) = -u_1^c(-x_1, -x_2, -x_3). \end{aligned}$$

In general, the relationship between the displacement at a point and corresponding points taken through planes of symmetry can be given as

$$\begin{aligned} u_i^c(\mathbf{x}) &= u_i^c(\mathbf{x}') \quad \text{if} \quad x_i = x_i' \\ u_i^c(\mathbf{x}) &= -u_i^c(\mathbf{x}') \quad \text{if} \quad x_i = -x_i' \end{aligned} \tag{2}$$

where

$$|x_i| = |x_i'|, \quad i = 1, 2, 3. \tag{3}$$

The derivative of equations (2) with respect to x_j under the restriction of equation (3) can be written formally as

$$\begin{aligned} u_{i,j}^c(\mathbf{x}) &= u_{i,j}^c(\mathbf{x}') \quad \text{if} \quad i = j \\ u_{i,j}^c(\mathbf{x}) &= u_{i,j}^c(\mathbf{x}') \quad \text{if} \quad i \neq j \quad \text{and} \quad x_i x_j = x_i' x_j' \\ u_{i,j}^c(\mathbf{x}) &= -u_{i,j}^c(\mathbf{x}') \quad \text{if} \quad i \neq j \quad \text{and} \quad x_i x_j = -x_i' x_j' \end{aligned} \tag{4}$$

Take the derivative of equation (1) with respect to x_j.

$$u_{i,j}^c(\mathbf{x}) = \sum_{l=1}^{\infty} \frac{1}{(l-1)!} U^{\circ}_{i,jp_1p_2\ldots p_{l-1}} x_{p_1} x_{p_2} \ldots x_{p_{l-1}} \tag{5}$$

This relationship must hold for any point within the cube. If we choose a point $\mathbf{x}$ and a symmetry point $\mathbf{x}'$ under the restriction of equation (3), then we can relate the derivatives of the displacements of these two points through equations (4) and (5). Since there must be term by term equality in the Taylor expansion of equation (5) for each point $\mathbf{x}$ and $\mathbf{x}'$, it becomes apparent that the following Taylor coefficients of equation (1) become identically zero:

$$U^{\circ}_{i,jpqr\ldots s} = 0 \tag{6}$$

if an odd number of indices are the same. By this we mean that as long as all indices do not appear an even number of times, the coefficient must be zero, for example,

$$U^{\circ}_{1,12} = U^{\circ}_{2,3311} = U^{\circ}_{2,111} = U^{\circ}_{1,2} = 0.$$

Further simplification can be achieved by invoking the argument that the cube remains invariant for any 90° rotation about a coordinate axis. This allows the Taylor coefficients when expanded about the origin to be represented as

$$\begin{aligned} U^{\circ}_{k,l} &= A\delta_{kl} \\ U^{\circ}_{k,lm} &= 0 \\ U^{\circ}_{k,lmn} &= C(\delta_{kl}\delta_{mn} + \delta_{km}\delta_{ln} + \delta_{kn}\delta_{lm}) + (B - 3C)\delta_{klmn}, \text{ etc.} \end{aligned} \tag{7}$$

where

$$A \equiv U^{\circ}_{1,1}, \quad B \equiv U^{\circ}_{1,111}, \quad C \equiv U^{\circ}_{1,122},$$

δ_{kl} is the Kronecker delta function and δ_{klmn} equals one if $k = l = m = n$ and is zero otherwise. For the parabolic approximation, there are only three independent coefficients that evolve due to the high symmetry of the system. If these constants can be determined, the constrained displacement and hence strain can be calculated at any point within the cuboid via equation (1). It can be shown that the constants B and C are not totally independent.

This is accomplished by applying the equilibrium equations of elasticity to the stress field within the cuboidal precipitate giving

$$\sigma^{c*}_{kl,l}(\mathbf{x}) = [C^*_{klij} e^c_{ij}(\mathbf{x})]_{,l} = 0 \tag{8}$$

where $\sigma^{c*}_{ij}(\mathbf{x})$ is the constrained stress, $e^c_{ij}(\mathbf{x})$ is the constrained strain and C^*_{ijkl} are the elastic constants of the precipitate phase. Substituting equation (7) into the expression for the constrained strain as calculated from equation (1), retaining only up to the parabolic terms and expanding gives for equation (8),

$$[(4C + 2B)\lambda + (8C + 4B)\mu]x_i = 0.$$

Since both λ and μ are independent, it is required that each of their respective coefficients go to zero, i.e., $B = -2C$ or $U^{\circ}_{1,111} = -2U^{\circ}_{1,122}$.

Integral Equation

Equation (7) permits solution of the system of equations given by

equation (36) in Part 1. All that need be determined are $U^{\circ}_{1,1}$ and $U^{\circ}_{1,111}$ for a second-order (parabolic) approximation to the constrained displacement field. With two unknowns, two independent equations are required. The first such equation can be derived from equation (36) of Part 1 by taking l equal to one and s equal to two yielding,

$$U^{\circ}_{i,p} = \Delta C_{kjnm} U^{\circ}_{k,j} T^{\circ}_{im,np} + \tfrac{1}{2}\Delta C_{kjnm} T^{\circ}_{im,np,st} U^{\circ}_{k,jst} + \sigma^{T*}_{nm} \alpha^{\circ}_{imnp}. \tag{9}$$

Substituting equation (33) of Part 1 and equation (7) into equation (9), setting $i = p$ and summing give for the first independent equation

$$A(1 - \Delta K T^{\circ}_{im,mi}) - C\frac{\Delta\mu}{3}\{T^{\circ}_{im,mi,ss} + 2T^{\circ}_{is,ti,st} - 5T^{\circ}_{im,mi,mm}\} = \frac{K^*\varepsilon}{2\mu}\frac{(1-2\nu)}{(1-\nu)} \tag{10}$$

where K is the bulk modulus, $\Delta K = K^* - K$, $\Delta\mu = \mu^* - \mu$, ν is Poisson's ratio and ε is the misfit, that is $e^{T*}_{ij} = \varepsilon\delta_{ij}$. The following notation is adopted concerning a repeated indice appearing four times:

$$T_{iijjjj} = T_{ii1111} + T_{ii2222} + T_{ii3333}.$$

The second independent equation is obtained in the same manner by equating l to three and s to two in equation (36) of Part 1:

$$U^{\circ}_{i,pqr} = \Delta C_{kjnm} U^{\circ}_{k,j} T^{\circ}_{im,npqr} + \tfrac{1}{2}\Delta C_{kjnm} U^{\circ}_{k,jst} T^{\circ}_{im,npqr,st} + \sigma^{T*}_{nm}\alpha^{\circ}_{imnpqr}. \tag{11}$$

By setting $i = p = q = r$, summing and realizing that $U^{\circ}_{i,iii} = -6C$, equation (11) reduces to

$$-3\Delta K A T^{\circ}_{im,miii} = C\{6 + \Delta\mu[T^{\circ}_{im,miii,ss} + 2T^{\circ}_{is,tiii,st} - 5T^{\circ}_{im,miii,mm}]\} - \frac{3K^*\varepsilon}{16\pi\mu}\frac{(1-2\nu)}{(1-\nu)}\psi^{\circ}_{,mmiiii}. \tag{12}$$

Simultaneous solution of equations (10) and (12) yields for the origin Taylor coefficients A and C,

$$C = U^{\circ}_{1,122} = \frac{\dfrac{3K^*\varepsilon(1-2\nu)X}{16\pi\mu(1-\nu)}\psi^{\circ}_{,mmiiii} - \dfrac{K^*\varepsilon(1-2\nu)Z}{2\mu(1-\nu)}}{XW + YZ}$$

$$A = U^{\circ}_{1,1} = \frac{\dfrac{K^*\varepsilon(1-2\nu)}{2\mu(1-\nu)} + CY}{X} \tag{13}$$

where

$$W = 6 + \Delta\mu[T^{\circ}_{im,miii,ss} + 2T^{\circ}_{is,tiii,st} - 5T^{\circ}_{im,miii,mm}]$$

$$X = 1 - \Delta K T^{\circ}_{im,mi}$$

$$Y = \frac{\Delta\mu}{3}\{T^{\circ}_{im,mi,ss} + 2T^{\circ}_{is,ti,st} - 5T^{\circ}_{im,mi,mm}\}$$

$$Z = 3\Delta K T^{\circ}_{im,miii}.$$

The values for the tensors T° are given in Appendix A.

Equation (13) gives the coefficients to the parabolic representation for the constrained strain when the constrained displacement is expanded in a Taylor series about the origin. Both constants now can be determined solely from the elastic constants of the precipitate and matrix. A first-order approximation may be quickly obtained from equation (13) by setting C equal to zero and then solving for A. For an expansion about the point of which the strain is to be calculated, no such simplification is possible and the entire system of equations, as given by equation (46) of Part 1, must be solved simultaneously at every point. Note that from equations (44) and (45) in Part 1 a more direct zeroth-order approximation may be attempted. Instead of solving the system of 117 equations for a parabolic expression at every point of interest, l could be taken equal to one and s equal to zero and the resultant system of nine simultaneous equations solved for the $U^{\circ}_{i,p}$. This is similar to the first method employed in solving the integral equation discussed in Part 1 when an origin expansion was considered.

Modified Equivalency Method

The symmetry arguments leading to equation (7) also allow for a reduction in the complexity of the system of equations associated with the modified equivalency method (see equation (14) in Part 1). The modified equivalency method is based upon expressing the strain field of the inhomogeneity by an equivalent inclusion which engenders the identical strain field as does the original precipitate. Such representation requires determination of an equivalent stress-free transformation strain, $\beta_{ij}(\mathbf{x})$, which can be expressed in terms of a polynomial in the position coordinates (equation (2) in Part 1).

The coefficients comprising $\beta_{ij}(\mathbf{x})$ can be determined through the following analysis. From equation (5) obtain the constrained strain and equate this with the expression for the constrained strain as given by equation (13) in Part 1. A term-by-term comparison requires that

$$D^{\circ}_{ijkl}B_{kl} + D^{\circ}_{ijklm}B_{klm} + D^{\circ}_{ijklmn}B_{klmn} + \ldots = M\delta_{ij} \tag{14}$$

$$D^{\circ}_{ijkl,p}B_{kl} + D^{\circ}_{ijklm,p}B_{klm} + D^{\circ}_{ijklmn,p}B_{klmn} + \ldots = 0 \tag{15}$$

$$D^{\circ}_{ijkl,pq}B_{kl} + D^{\circ}_{ijklm,pq}B_{klm} + D^{\circ}_{ijklmn,pq}B_{klmn} + \ldots = N\{\delta_{ij}\delta_{pq} + \delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp} - 5\delta_{ijpq}\}, \text{ etc.} \tag{16}$$

for the constant through parabolic terms, where M and N are constants. Now assume, as in the integral equation, that a parabolic or second-order approximation is desired.

The numerical values of the coupling matrix, $D^{\circ}_{ij\ldots k}$, depend upon the biharmonic potential functions and their derivatives. These expressions can be simplified through the following considerations. Remembering that the integral of an odd function over a region symmetric with respect to the origin is equal to zero, it becomes apparent that any biharmonic potential function (for a cube evaluated at the origin) with an odd number of subscripts is identically zero. For example,

$$\psi^{\circ}_{,ijklm} = \psi^{\circ}_{i,jklm} = \psi^{\circ}_{ij,klmnp} = 0.$$

A direct consequence is that any $D^{\circ}_{ij\ldots k}$ with an odd number of subscripts is also identically zero. Realizing this yields immediately for equation (15)

$$D^{\circ}_{ijklm,p}B_{klm} = 0. \tag{17}$$

By definition, both $D^{\circ}_{ijklm,p}$ and B_{klm} are symmetric with respect to k, l, and m and since the $D^{\circ}_{ijklm,p}$ are not all equal to zero, it follows that (remembering that i, j, and p are free indices) $B_{klm} = 0$. This leaves only the B_{kl} and B_{klmn} to determine, the nature of which can be extracted from equations (14) and (16) with the following considerations. It has been shown that for an invariant transformation of a cube no contribution from the permutation tnesor occurs [14]. The nonzero $D^{\circ}_{i\ldots j}$ values of equations (14) and (16) display invariance under a cube symmetry rotation by virtue that the constituent potential functions are cube invariant when evaluated at the origin. Since the right-hand sides of the equations are cube invariant [14] and each addition factor on the left-hand side undergoes a tensor contraction to a cube invariant function, it can be shown to follow that the B_{kl} and B_{klmn} must also be cube invariant. In general, the B's must be expressible as

$$B_{kl} = d\delta_{kl}$$

$$B_{klmn} = e\delta_{kl}\delta_{mn} + f(\delta_{km}\delta_{ln} + \delta_{kn}\delta_{lm}) + (h - e - 2f)\delta_{klmn} \tag{18}$$

where d, e, f, and h are constants. Hence there are at most four independent values comprising $\beta_{ij}(\mathbf{x})$ when a parabolic approximation is considered.

The numerical values of d, e, f, and h can be determined by considering the equations of equilibrium for the inhomogeneity and equivalent inclusion, that is

$$C_{ijkl}[\beta_{kl}(\mathbf{x})]_{,j} = 0. \tag{19}$$

Substituting equation (18) into equation (19) and expanding gives

$$[(h + 2e)\lambda + (4h + 8f)\mu]x_i = 0. \tag{20}$$

Again, since both λ and μ are independent, $e = f$ and $h = -2e$. Hence,

$$B_{kl} = d\delta_{kl}$$
$$B_{klm} = 0 \tag{21}$$
$$B_{klmn} = e\{\delta_{kl}\delta_{mn} + \delta_{km}\delta_{ln} + \delta_{kn}\delta_{lm} - 5\delta_{klmn}\}.$$

Therefore, for a cuboidal precipitate that undergoes a stress-free transformation strain of a pure dilatation, there are only three unique values appearing in $\beta_{ij}(\mathbf{x})$. This assumes that only a parabolic approximation is being treated. The three independent values are exemplified by B_{11}, B_{1111}, and $B_{1122} = B_{1212}$ and their permutations where $B_{1111} = -2B_{1122}$.

The two unknowns d and e, can be determined by substitution of equation (30) of Part 1 into the system of equations given by equation (14) in Part 1. Only two independent equations are needed. For the first equation use equation (14a) of Part 1 and let $i = j$. In the second independent equation, let $i = j = s = t$ in equation (14c) of Part 1 and then sum. The two simultaneous equations are then expressible as

$$d\Delta K D^{\circ}_{kkmm} + e\Delta K\{D^{\circ}_{kkmmpp} + 2D^{\circ}_{kkpmpm} - 5D^{\circ}_{kkmmmm}\} + 3dK = 3\varepsilon K^* \tag{22}$$

and

$$d\Delta\lambda D^{\circ}_{kkmm,ii} + 2d\Delta\mu D^{\circ}_{iimm,ii} + e\{\Delta\lambda D^{\circ}_{kkmmpp,ii} + 2\Delta\lambda D^{\circ}_{kkpmpm,ii} - 5\Delta\lambda D^{\circ}_{kkmmmm,ii} + 2\Delta\mu D^{\circ}_{iimmpp,ii} + 4\Delta\mu D^{\circ}_{iipmpm,ii} - 10\Delta\mu D^{\circ}_{iimmmm,ii} - 24\mu\} = 0. \tag{23}$$

Solving equations (22) and (23) simultaneously for d and e gives

$$e = \frac{3\varepsilon K^* Q}{SQ - RZ} \quad d = \frac{3\varepsilon K^* - eS}{R} \tag{24}$$

where

$$R = 3K + \Delta K D^{\circ}_{kkmm}$$
$$S = \Delta K(D^{\circ}_{kkmmpp} + 2D^{\circ}_{kkpmpm} - 5D^{\circ}_{kkmmmm})$$
$$Q = \Delta\lambda D^{\circ}_{kkmm,ii} + 2\Delta\mu D^{\circ}_{iimm,ii}$$
$$Z = \Delta\lambda D^{\circ}_{kkmmpp,ii} + 2\Delta\lambda D^{\circ}_{kkpmpm,ii} - 5\Delta\lambda D^{\circ}_{kkmmmm,ii} + 2\Delta\mu D^{\circ}_{iimmpp,ii} + 4\Delta\mu D^{\circ}_{iipmpm,ii} - 10\Delta\mu D^{\circ}_{iimmmm,ii} - 24\mu.$$

Note that a lower-order approximation may be obtained directly from equation (24) by setting $e = 0$. Such a representation approximates the equivalent stress-free transformation strain within the cube as a constant, the value of which is given by $d = 3\varepsilon K^*/R$. Once the B_{kl} and B_{klmn} have been determined, the constrained strain can be calculated from equation (9) in Part 1. The values of the D° are given in Appendix B.

The Strain Field

In the section, "Symmetry Properties," symmetry arguments associated with the cuboidal morphology are coupled with a uniform stress-free transformation strain to yield a simplified relationship between the Taylor coefficients of the constrained strain when the expansion is performed about the origin. These results can then be applied directly to the integral equation method and the 117 simultaneous equations associated with a second-order or parabolic expansion can be reduced to three independent equations. Further consideration of the equations of equilibrium requires that $U_{1,111} = -2U_{1,122}$; allowing the Taylor coefficients of the origin expansion to be representable by two constants whose values are dictated by the elastic constants of the precipitate and matrix phases. A more accurate solution requires expansion of the constrained strain about the point of which the strain is to be calculated and the corresponding solution of the resultant 117 simultaneous equations that would accrue from the parabolic approximation (see equation (46) in Part 1). A less complicated though somewhat less accurate version is the zeroth-order approximation which requires solution of only nine independent equations and yet encompasses a correspondingly tremendous decrease in the number of potential functions that need be calculated.

The cuboidal symmetry arguments can also be applied directly to the modified equivalency method which allows the equivalent stress-free transformation strain to be expressed in terms of three coefficients. Application of the equations of equilibrium again requires the relationship $B_{1111} = -2B_{1122}$ for a parabolic approximation about the origin. The values of these two coefficients are dependent upon the elastic constants of the precipitate and matrix phases.

With the plethora of solution techniques varying drastically in complexity and computational time, it becomes necessary to determine which solution mode optimizes desired quantities for a given situation. It is also necessary to determine the reliability of the individual techniques, a process which is facilitated in that the integral equation and modified equivalency methods are based upon two disparate approaches. With the intent of examining the strain field of a cuboidal inhomogeneity while also substantiating the reliability of the two methods, we will begin a systematic approach to examining the correspondence between the two techniques and their modes of solution. Homogeneous and other degenerate cases will be examined first, progressing from very small to very large differences in the elastic constants between the precipitate and matrix phases.

Degenerate Cases. The homogeneous case for a cuboidal inclusion and a pure dilatational stress-free transformation strain has been treated by Faivre [15]. For any degree of reliability, the modified equivalency condition and integral equation methods should degenerate into the homogeneous case when the elastic constants of the precpitate and matrix phases are equal. The only exception may be the origin expansion of the integral equation where the strain field is being approximated by a parabolic representation. Such an analytical situation is realized for the equivalency condition which is straightforward in that equation (14) in Part 1 reduces directly to $\beta_{ij}(\mathbf{x}) = e^{T*}_{ij}$. Such exact results are also ostensible when the integral equation is expanded about the point of interest. However as expected, when the results are calculated from an origin expansion, deviation from the exact result is observed. The extent of the deviation is negligible out to a distance of about $0.6a$ from the cube center especially in those directions approaching the cube face. As the cube edge or corner is approached this difference is magnified. The origin expansion of the integral equation does provide an overall feel for the strain field behavior and therefore has been retained. For a precipitate morphology with a more gently curving surface in which the strain field would not be expected to be such a strong function of position, the origin expansion approach may be highly desirable. Analytical expressions for the constrained strain are available [4, 5, 16] for the case when the stress-free transformation strain is a pure dilatation and $\mu^* = \mu$ but $\lambda^* \neq \lambda$. This provides another check on the different solution schemes and again the same behavior is observed as in the homogeneous system with all methods yielding exact results.

As a further indication of the reliability of the various modes of solution, the Lamé constant for the precipitate and matrix phases are equated ($\lambda^* = \lambda$) while the shear modulus of the precipitate is allowed to differ from that of the matrix by 5 percent. Since the elastic constants of the inhomogeneous system are almost identical, its strain field should behave similarly to that of the homogeneous system. This behavior is observed for all methods of solution with the difference between each method being negligible.

General Case. In order to measure the relative reliability of the various solution schemes, we have first allowed $\lambda^* = \lambda$ and $\mu^* = 1.3\mu$, where $\lambda = 12.14$ and $\mu = 7.54$ in units of 10^4MN/m^2 are those of Cu. Here, the 30 percent difference in shear moduli should have a profound influence on the convergence of the respective solutions. The following three solution schemes are considered in the present and subsequent examples:

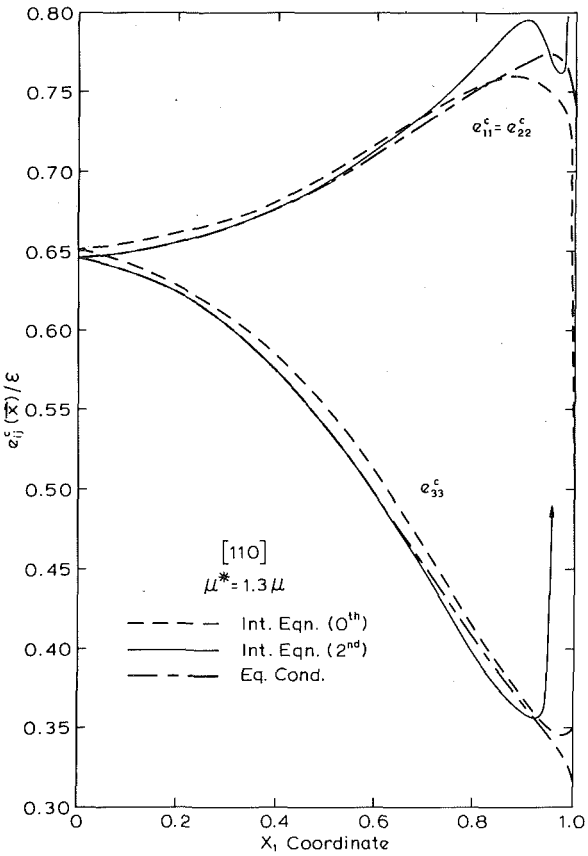


Fig. 2 Constrained strain as a function of the normalized x_1 coordinate for three different methods of solution in the [110] direction, $\mu^* = 1.3$

1 Modified equivalency method—parabolic expansion about origin.

2 Integral equation—zeroth-order expansion about the point of interest.

3 Integral equation—parabolic expansion about the point of interest.

Methods 1 and 3 yielded essentially the same results in that the strains agree to at least three or four significant figures in the [100] direction. Even in the nonsymmetrical direction of [210], the diagonal strain components differ by less than 1 percent between schemes 1 and 3 while differences in the shear term vary from a few percent as the interface is approached to almost 7 percent at the interface. Method 2 is within 1 percent of methods 1 and 3 in the [100] direction, within 2 percent of the diagonal terms, and 8 percent of the shear terms along the [210] direction. Method 2 tends to approach 1 more closely than 3 at close approach to the interface. No apparent preference is observed away from the interface.

Fig. 2 depicts the strain behavior of the diagonal components for the case of $\mu^* = 1.3\mu$ along the [110] direction. The constrained strain has been normalized in terms of the stress-free transformation strain and is plotted as a function of a reduced x_1-coordinate in units of the semiedge length, a, from the center of the cube. The dashed line represents the zeroth-order approximation to the integral equation (Method 2) and the solid line the second-order approximation to the integral equation, again expanded about the point of interest (Method 3). The parabolic approximation to the modified equivalency condition (Method 1) is depicted by the dashed line of alternating length. The [110] direction is chosen in that it depicts the largest differences between the respective solution schemes. Note the instability in Method 3 at close proximity to the cube edge. Method 2 appears to demonstrate some divergence at very small distances from the edge. The shear term as computed by Method 3 also experiences a large fluctuation near the edge discontinuity. The behavior of Method 3 is identical to Method 1 out to about $x_1 = 0.5a$ from which the difference in shear terms gradually increases to about 6 percent at $x_1 = 0.9a$ and about 12 percent at $x_1 = 0.95a$. The difference in the shear strain between solution Schemes 1 and 2 is within about 12 percent. The same relative behavior between Methods 1, 2, and 3 is observed in the [111] direction as in the [110] direction with roughly the same tolerances applying, though the instability in Method 3 is accentuated very close to the cube corner.

At this time we have examined the case in which the shear moduli of precipitate and matrix were almost equal, yielding practically identical results for all three methods of solution. We have also discussed the situation where $\mu^* = 1.3\mu$. Here there was an introduction of some degree of instability into the strain field of the integral equation method as the point of interest approached the discontinuities of the edge and corner. No such instability was indicated in the [100] or in nonsymmetric directions such as [210] in which the cube edges or corners are not approached. Since the agreement is found to be good between Methods 1 and 3, which theoretically should be more accurate, it is felt that either of these two methods can be used in determination of the strain field for a precipitate in which the difference in shear moduli is not too large. In the vicinity of the cube edge or corner the modified equivalency condition, Method 1, is felt to yield fairly reliable results for the strain field. Such a statement is based upon two observations: (a) good agreement between Methods 1 and 3 in directions which do not approach either the cube edge or corner and (b) the fact that the expansion of the equivalent stress-free transformation strain is taken about the origin and not the point of interest. Such an origin expansion is not much influenced by a fluctuant mode of strain near the precipitate-matrix interface.

When there is not a radical difference in the shear moduli between precipitate and matrix phases, the integral equation and modified equivalency approaches are seen to behave quite similarly. For small differences in the shear moduli, Method 2 may be used as a good approximation to the strain, particularly the diagonal components which are used in any calculation of the elastic self energy of a precipitate that had undergone a pure dilatation.

In an effort to gauge the reliability of the equivalency condition and integral equation methods for systems in which there is an extreme difference in elastic constants between precipitate and matrix, attention is focused on the case where $\lambda^* = \lambda$ and $\mu^* = 3\mu$. Here any inherent inconsistencies in the numerical solutions should be greatly magnified. Essentially the same respective behavior between the solution schemes is observed as was seen for smaller differences in the elastic constants. The fluctuation in the strain field of Method 3 is more pronounced and has spread to some of the nonsymmetrical directions. The difference between Methods 2 and 1 has also increased a little, about 3 percent in the [100] direction and slightly more in the other directions. If the precipitate is softer than the matrix, the discrepancy between Methods 1 and 3 is minimal, at least in the diagonal strain components. For the case of $\lambda^* = \lambda$ and $\mu^* = \mu/3$ in the nonsymmetric [210] direction, the difference between the two methods is less than 6 percent for the diagonal terms through the discrepancies in the shear terms are much more pronounced (15–20 percent). The [100] direction also exhibits quite close agreement between Schemes 1 and 3. As the cube edge or corner is approached, instability in Method 3 is again observed.

In light of these results, it becomes apparent that the present Taylor series solution technique, which accounts for only a few lower-order terms, as applied to the integral equation is incapable of adequately expressing the strain field of a cuboidal (and hence probably any parallelepiped) precipitate when the difference in shear moduli is very large. Such results are not a reflection on the integral equation for this equation is exact; rather it is the technique of solution that becomes incompatible, by virtue of the cutoff of higher-order terms, with the problem when the precipitate-matrix interface is approached and is magnified as the difference in elastic constants become more pronounced. For such drastic differences it may become necessary to adopt a more time-consuming convergence method of solution such as perhaps a Born approximation [13], for higher-order Taylor series

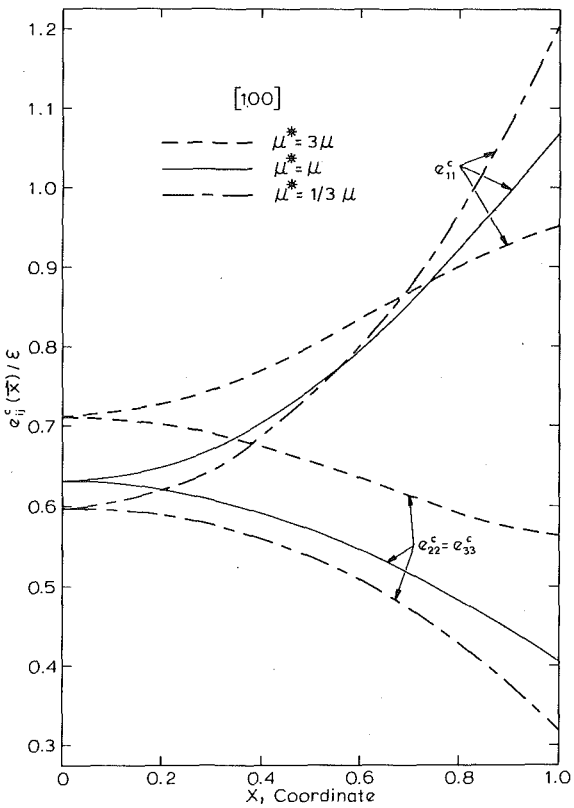


Fig. 3 Relative behavior of the constrained strain as a function of distance from the cube center for hard, soft, and homogeneous precipitates in the [100] direction

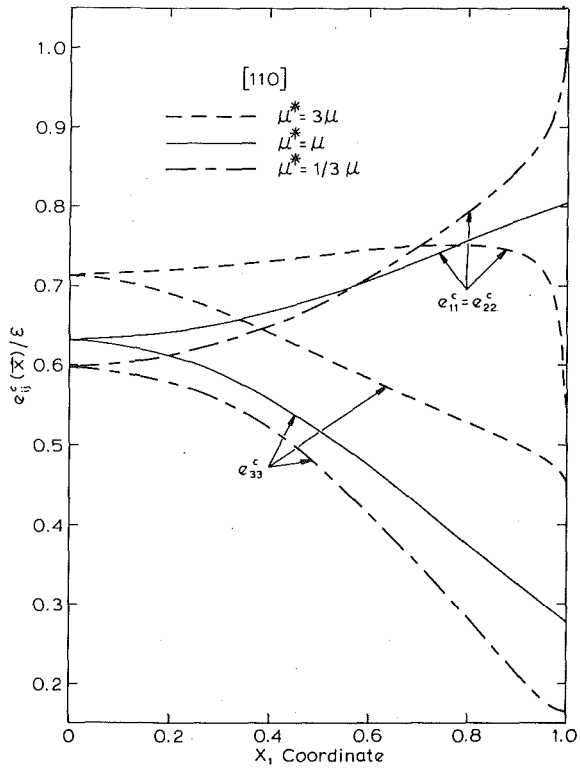


Fig. 4 Relative behavior of the constrained strain as a function of distance from the cube center for hard, soft, and homogeneous precipitates in the [110] direction

approximations may be too intrinsically complex. One may use a Taylor series expansion only for the central regions of the cube.

The modified equivalency condition, on the other hand, appears to give fairly reliable results near the cube interface, even for large differences in the shear moduli. Such results are predicated upon three distinct factors. First, is the very close agreement between Methods 1 and 3 in the interior of the cuboid for both very hard ($\mu^* = 3\mu$) and very soft ($\mu^* = \mu/3$) precipitates. This correlation is still retained for such systems, even as the interface is approached, along such directions as [421] and [100] though very near to the precipitate-matrix interface fluctuation is again observed. Second, is the ability of the zeroth-order approximation of the integral equation when expanded about the point of interest to detect all maxima and minima observed via the modified equivalency condition. The accuracy of the numerical results of Method 2 is questionable for such cases as $\mu^* = 3\mu$ but the fact that it always predicts the same trends as the modified equivalency condition indicates that the modified equivalency method is behaving properly at large distances from the origin. Actually, no fluctuation or instability would be expected from the modified equivalency method since the equivalent stress-free transformation strain is expressed as a polynomial expanded about the origin, and is far removed from a corner or edge. Its accuracy is maintained in that it still requires calculation of the potential functions at the point of interest. Another important factor is that the equivalent stress-free transformation strain does not appear to vary near as drastically as the constrained strain and is hence more susceptible to representation by a lower-order Taylor series expansion. Such a concept is reinforced in that at any point of calculation, the equivalent stress-free transformation strain is integrated over the entire precipitate and in the central regions of the cuboid the strains still agree quite closely with the parabolic expansion of the integral equation. Error bounds on the modified equivalency method would also increase, however, as the interface is approached for the equivalent stress-free transformation strain would presumably deviate more from the exact value where even higher-order terms may have an effect. Such differences are tempered to some extent because the integral encompasses large regions of accurate equivalent-strain representation.

The third point in favor of the equivalency condition is its satisfaction of the boundary conditions. The boundary conditions of the solution require that the constrained displacement as well as the surface traction must be continuous across the percipitate-matrix interface. From the displacement consideration it can be shown that in the [100] direction, $e^{cM}_{11} = e^{cP}_{11}$ where e^{cM}_{11} and e^{cP}_{11} are the constrained strains in the matrix and precipitate, respectively. Actual values computed for $\mu^* = 3\mu$ are $e^{cP}_{11} = -0.9502$ and $e^{cM}_{11} = -0.9429$, a difference of less than 1 percent. Similar analysis also indicates that the remaining diagonal terms must be continuous. Computed values give $e^{cP}_{22} = e^{cP}_{33} = 0.5634$ and $e^{cM}_{22} = e^{cM}_{33} = 0.5621$, a difference of less than 1 percent. The traction also requires computation of the strain immediately on both sides of the interface. Due to the discontinuous nature of the potential functions at this point, exact values are unobtainable in that the strains are calculated near but not on the interface. For the case of $\mu^* = 3\mu$ and the [100] direction, the stress component σ^P_{11} in the precipitate must equal σ^M_{11}, the stress component in the matrix. The values actually obtained are $\sigma^P_{11} = -1.79\mu\varepsilon$ and $\sigma^M_{11} = -1.59\mu\varepsilon$ or a difference of about 11 percent. Traction calculations employing the diagonal components demonstrate reasonable agreement while the tangential components show some divergence, an indication that the shear terms may not be as accurate as the diagonal terms.

In Figs. 3–5, the normalized constrained strain, $\varepsilon^c_{ij}(\mathbf{x})/\varepsilon$, is shown as a function of a reduced x_1-coordinate from the cube center for three different directions [100], [110], and [111]. The numerical values of $\varepsilon^c_{ij}(\mathbf{x})$ are obtained with Method 1, i.e., the modified equivalency method. For the purpose of comparison, three different types of precipitates are considered. The first, depicted by the solid line, is representative of a homogeneous inclusion. The dashed line depicts the case of a more rigid precipitate where $\mu^* = 3\mu$, and a softer precipitate where $\mu^* = \mu/3$ is represented by the dashed line of alter-

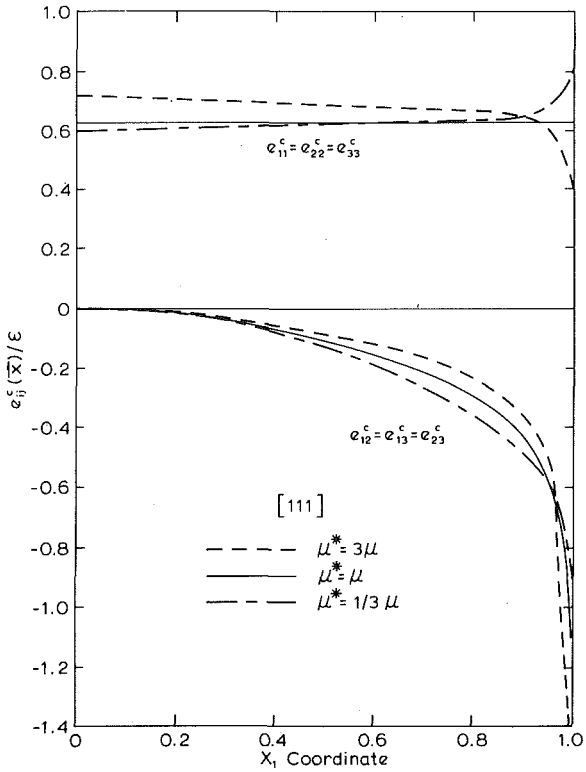


Fig. 5 Relative behavior of the constrained strain as a function of distance from the cube center for hard, soft, and homogeneous precipitates in the [111] direction

nating length. Along the [100] direction from the precipitate center, the e^c_{11} component increases smoothly up to the interface for all three precipitates. The rate of increase of e^c_{11}, however, depends upon the rigidity of the precipitate. The softer the precipitate the larger is the rate of increase in e^c_{11}. Hence, while the e^c_{11} value of a soft particle is smaller than that of hard particle in the central regions of the cube, the situation is reversed near the interface. The other two diagonal components, e^c_{22} and e^c_{33}, decrease up to the interface but their difference due to the rigidity of the precipitate is nearly constant along this direction. All shear components are zero in this case.

Similar behavior is observed along the [110] direction (Fig. 4), except that e^c_{11} and e^c_{22} are now increasing from the cube center while the e^c_{33} component decreases. As pointed out previously, the numerical values become unreliable near the cube edge or corner because of the cutoff of higher-order terms in the Taylor expansion. The sudden decrease of the e^c_{11} component near the interface in the case of $\mu^* = 3\mu$ seems to reflect some error in the solution technique. Fig. 5 shows both diagonal and shear components for the [111] direction. Again, a soft particle shows smaller e^c_{11} values than those of a hard particle in the central region, but yields higher values near the interface. The shear components approach a large negative value at the interface along this direction, which is characteristic of a faceted precipitate. When a precipitate is faceted, as is a cube, the sharp edges and corners are singularities where some strain (hence stress) components become unbounded [17, 18]. When the elastic constants of a precipitate are the same as those of the matrix phase, the dilatation, e^c_{kk}, becomes a constant regardless of the precipitate shape provided that the stress-free transformation is purely dilatational [4]. This is seen in Fig. 5 from the constancy of the e^c_{11}, e^c_{22}, and e^c_{33} components represented by the solid line. On the other hand when the elastic constants of the precipitate and matrix are different, the dilatation, e^c_{kk}, becomes a function of position as is also demonstrated in Fig. 5. This difference in behavior, which may also be manifested in the strain fields of the matrix phase, among other effects may have a profound influence on the rates of diffusion in the immediate neighborhood of the precipitate [19].

Summary

In this study we have applied the integral equation and modified equivalency methods, as derived in Part 1, to a cuboidal precipitate in an isotropic matrix. The establishment of reliability for the two methods and their respective techniques of solution is based solely upon their ability to reproduce degenerate cases in which analytical solutions are available and on the convergence of solutions for the two methods. Further verification comes from the reasonable satisfaction of the boundary conditions such as the continuity of displacement and traction at the precipitate-matrix interface.

For small changes in the elastic constants, the modified equivalency methods and the zeroth and parabolic approximations to the integral equation when expanded about the point of interest, all agree quite closely. Hence in this case, a zeroth-order approximation of the integral equation should be appropriate in determining the strain field, for it requires substantially less effort than the other two methods. For moderate changes in the shear modulus of the precipitate, either a second order approximation of the integral equation or the modified equivalency method yields fair results. The former does begin to show some fluctuation as the cube edges and corners are approached.

With large differences between the shear moduli of the precipitate and matrix ($\mu^* = \mu/3$ or $\mu^* = 3\mu$), large fluctuations are observed to spread from the corners and edges to other regions near the precipitate-matrix interface when the integral equation is employed. Such behavior reduces the confidence that can be placed in these results when strain values are computed near the interface. Such tendencies are probably mainfestations of the solution technique applied so near the discontinuities in the constrained strain field found at the interface and may be avoidable if a larger number of terms in the Taylor expansion are taken into account. The modified equivalency method is found to exhibit fair behavior even for large differences in shear moduli between the precipitate and matrix phases. Such a conclusion is based upon the reasonable satisfaction of the boundary conditions across the precipitate-matrix interface.

Using the modified equivalency method, the constrained strain $e^c_{ij}(\mathbf{x})$, is analyzed as a function of position and orientation for three different types of cuboidal precipitates with a dilatational stress-free transformation strain. In the central regions of the cube, the constrained strain increases as the shear modulus of the precipitate increases. Since the total elastic strain of the precipitate is the constrained strain minus the stress-free transformation strain, this result simply demonstrates the fact that the harder the particle phase the more the deformation is accommodated in the matrix phase. As the precipitate-matrix interface is approached, however, the behavior of the strain field becomes complicated because of facets, edges and corners associated with the precipitate. For example, along the [100] direction, the e^c_{11} component of a soft particle near the interface becomes much larger than that of the hard particle while the other strain components remain nearly independent of the precipitate rigidity. As a result, the dilatation of an inhomogeneous cuboidal precipitate is not constant as is found for a corresponding inclusion.

Acknowledgments

The authors would like to thank Prof. D. M. Barnett and T. Mura for their helpful discussions on this work. This work was supported by the Division of Materials Research of the National Science Foundation under Grants DMR76-06855 A01 and 79-01693 for which much appreciation is expressed.

References

1 Johnson, W. C., Earmme, Y. Y., and Lee, J. K., "Approximation of the Strain Field Associated With an Inhomogeneous Precipitate—Part 1: Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 775–780.

2 Eshelby, J. D., "Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proceedings of the Royal Society*, Vol. 241, Series A, 1957, pp. 376–396.

3 Eshelby, J. D., "The Elastic Field Outside an Ellipsoidal Inclusion," *Proceedings of the Royal Society,* Vol. 252, Series A, 1959, pp. 561–569.

4 Eshelby, J. D., "Elastic Inclusions and Inhomogeneities," *Progress in Solid Mechanics* Vol. 2, 1961, pp. 89–140.

5 Robinson, K., "Elastic Energy of an Ellipsoidal Inclusion in an Infinite Solid," *Journal of Applied Physics,* Vol. 22, 1951, pp. 1045–1054.

6 Shibata, M., and Ono, K., The Strain Energy of a Spheroidal Inclusion and Its Application to B.C.C.-H.C.P. Martensite Transformation," *Acta Met.,* Vol. 23, 1975, pp. 587–597.

7 Willis, J. R., "The Solution of Asymmetric Problems of Elasticity of Fourier Transforms," Adams Prize Essay, University of Cambridge, 1970, unpublished.

8 Kinoshita, N., and Mura, T., "Elastic Fields of Inclusions in Anisotropic Media," *Phys. Stat. Sol., Vol. A5,* 1971, pp. 759–768.

9 Asaro, R. J., and Barnett, D. M., "The Nonuniform Transformation Strain Problem for an Anisotropic Ellipsoidal Inclusion," *Journal of the Mechanics and Physics of Solids,* Vol. 23, 1975, pp. 77–83.

10 Lee, J. K., Barnett, D. M., and Aaronson, H. I., "The Elastic Strain Energy of Coherent Ellipsoidal Precipitates in Anisotropic Crystalline Solids," *Met. Trans.,* Vol. 8A, 1977, pp. 963–970.

11 Evans, J. H., Bullough, R., and Stoneham, A. M., "Radiation Induced Voids in Metals," eds., Corbett, J. W., and Ianniello, L. C., U. S. Atomic Energy Commission, Washington, 1972, pp. 522–532.

12 Ardell, A. J., and Nicholson, R. B., with Appendix by Eshelby, J. D., "On the Modulated Structure of Aged Ni-Al Alloys," *Acta Met.,* Vol. 14, 1966, pp. 1295–1309.

13 Chen, F. C., and Young, K., "Inclusions of Arbitrary Shape in an Elastic Medium," *Journal of Mathematics and Physics,* Vol. 12, 1977, pp. 1412–1416.

14 Moschovidis, Z. A., and Mura, T., "Two Ellipsoidal Inhomogeneities by the Equivalent Inclusion Method," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, 1975, pp. 847–851.

15 Faivre, G., "Deformations de Coherence d'un Precipitate Quadratique," *Phys. Stat. Sol.,* Vol. 35, 1969, pp. 249–259.

16 Goodier, J. N., "On the Integration of the Thermoelastic Equations," *Philosophical Magazine,* Vol. 23, 1937, pp. 1017–1032.

17 Sankaran, R., and Laird, C., "Deformation Field of a Misfitting Inclusion," *Journal of the Mechanics and Physics of Solids,* Vol. 24, 1976, pp. 251–261.

18 Lee, J. K., and Johnson, W. C., "Calculation of the Elastic Strain Field of a Cuboidal Precipitate in an Anisotropic Matrix," *Phys. Stat. Sol.,* Series(a), Vol. 46, 1978, pp. 267–272.

19 Li, J. C. M., "Physical Chemistry of Some Microstructural Phenomena," *Met. Trans.,* Vol. 9A, 1978, pp. 1353–1380.

APPENDIX *A*

$T^{\circ}_{im,mi} = -1/M$

$T^{\circ}_{im,miii} = 1.47021/M$

$T^{\circ}_{im,mi,ss} = 0.0$

$T^{\circ}_{is,ti,st} = 1.51520/M$

$T^{\circ}_{im,mi,mm} = 0.51520/M$

$T^{\circ}_{mi,immm,ss} = -5.71860/M$

$T^{\circ}_{mi,immm,ii} = -4.16225/M - 1.51329/\mu$

$T^{\circ}_{is,tiii,st} = -5.38025/M - 3.0265/\mu$

$\psi^{\circ}_{,iimmmm} = -36.95042$

$M = \lambda + 2\mu$

APPENDIX *B*

$D^{\circ}_{iikk} = 3K/M$

$D^{\circ}_{iikk,mm} = 0$

$D^{\circ}_{iikk,ii} = 4.41063K/M$

$D^{\circ}_{iikkmm} = 0$

$D^{\circ}_{iikkmm,pp} = 18K/M$

$D^{\circ}_{iikkmm,ii} = 17.15581K/M$

$D^{\circ}_{iiklkl} = -3.03041\mu/M$

$D^{\circ}_{iiklkl,pp} = 6\lambda/M + 24\mu/M$

$D^{\circ}_{iiklkl,ii} = 6.05316 + 5.71860\lambda/M + 10.76811\mu/M$

$D^{\circ}_{iikkkk} = -1.03041\mu/M$

$D^{\circ}_{iikkkk,pp} = 6$

$D^{\circ}_{iikkkk,ii} = 3.0265 + 5.71860\lambda/M + 8.32447\mu/M$

$M = \lambda + 2\mu$

G. G. Adams

Assistant Professor,
Department of Mechanical Engineering,
Northeastern University,
Boston, Mass. 02115
Assoc. Mem. ASME

A Semi-Infinite Elastic Strip Bonded to an Infinite Strip

The solution is obtained for the plane strain problem of a semi-infinite elastic strip whose end is bonded to and pressed against an infinite elastic strip. The infinite strip is supported by a pair of symmetrically located, concentrated forces. Using integral transform techniques, the solution is reduced to a set of singular integral equations of the second kind. The order of the singularity is determined and the equations are then solved numerically. The results show the normal and shear stress distributions as well as the stress-intensity factors for a range of support locations corresponding to various width ratios and material combinations.

Introduction

The elastostatic solution for a semi-infinite strip with arbitrary end conditions was obtained by Bogy [1]. In [2] this solution was used to solve the problem of joined semi-infinite strips of different material properties in tension. The corresponding bending problem was solved by Adams and Bogy [3]. They then obtained solutions for a semi-infinite strip in contact with a half plane [4] and for two contacting semi-infinite strips of different widths [5].

Here we consider the problem of a semi-infinite strip whose end is bonded to an infinite strip. Such a "T-joint" occurs in many physically relevant situations. Of particular interest here is the stress distribution along the bond and the effects of the width ratio, the composite material parameters, and the location of the supports.

First a solution is obtained for an infinite strip satisfying appropriate boundary conditions along the edges. This is combined with the semi-infinite strip solution [1] and continuity conditions are applied. This procedure finally yields a system of singular integral equations which is analyzed in order to find the order of the singularity at the corners. These equations are then solved numerically using the method of Erdogan and Gupta [6], yielding the normal and shear stress distributions along the interface. An interesting result is that for sufficiently narrow infinite strips, there exists a certain location of the supports for which shear stress is bounded. Another support location corresponds to bounded normal stress.

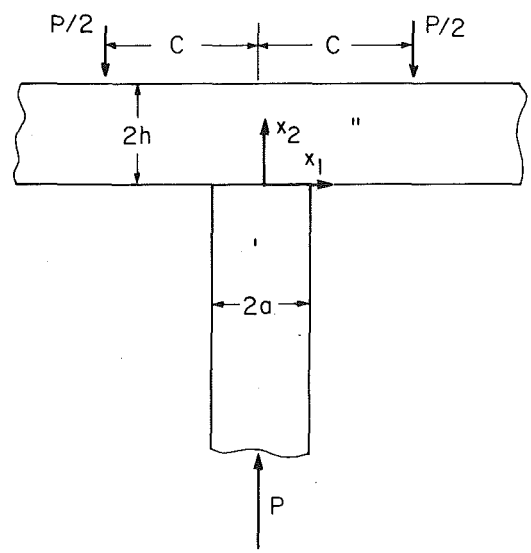


Fig. 1 A semi-infinite strip bonded to an infinite strip of different widths and material properties

Problem Formulation

The problem consists of an isotropic, homogeneous, semi-infinite elastic strip of constant width $2a$, with elastic constants μ', ν'. This is bonded to an infinite elastic strip of uniform width $2h$, having elastic constants μ'', ν''. The semi-infinite strip is pressed against the infinite strip by a resultant force P acting in the x_2-direction, whereas the infinite strip is supported by a pair of concentrated forces of magnitude $P/2$ (Fig. 1). We wish to find the two-dimensional (plane strain) stress and displacement fields which satisfy the appropriate elasticity equations for each region and which also satisfy the following boundary conditions:

$$\tau'_{11}(\pm a, x_2) = 0, \quad -\infty < x_2 < 0 \tag{1}$$

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, April, 1980. Paper No. 80-WA/APM-13.

$$\tau'_{12}(\pm a, x_2) = 0, \quad -\infty < x_2 < 0 \tag{2}$$

$$\tau'_{12}(x_1, x_2) \to 0, \quad \tau'_{22}(x_1, x_2) \to -P/2a, \quad |x_1| < a, x_2 \to -\infty \tag{3}$$

$$\tau''_{ij}(x_1, x_2) \to 0, \quad 0 < x_2 < 2h, \quad |x_1| \to \infty \tag{4}$$

$$\tau''_{12}(x_1, 2h) = 0, \quad |x_1| < \infty \tag{5}$$

$$\tau''_{22}(x_1, 2h) = -(P/2)\delta(|x_1| - c) \tag{6}$$

$$\tau''_{12}(x_1, 0) = 0, \quad |x_1| > a \tag{7}$$

$$\tau''_{22}(x_1, 0) = 0, \quad |x_1| > a \tag{8}$$

and continuity conditions

$$\tau'_{12}(x_1, 0) = \tau''_{12}(x_1, 0), \quad |x_1| < a \tag{9}$$

$$\tau'_{22}(x_1, 0) = \tau''_{22}(x_1, 0), \quad |x_1| < a \tag{10}$$

$$u'_1(x_1, 0) = u''_1(x_1, 0), \quad |x_1| < a \tag{11}$$

$$u'_2(x_1, 0) = u''_2(x_1, 0), \quad |x_1| < a \tag{12}$$

Infinite Strip Solution

Applying the exponential Fourier transform, with respect to x_1, to the elasticity equations for the infinite strip and taking advantage of symmetry, the stress and displacement fields become

$$\tau''_{11}(x_1, x_2) = -\sqrt{2/\pi}\int_0^\infty \omega\{[\omega A_1(\omega) + \omega x_2 B_1(\omega) + 2\nu'' B_2(\omega)]\sin\omega x_2 + [\omega A_2(\omega) + \omega x_2 B_2(\omega) + 2\nu'' B_1(\omega)]\cosh\omega x_2\}\cos\omega x_1 d\omega,$$

$$\tau''_{22}(x_1, x_2) = \sqrt{2/\pi}\int_0^\infty \omega\{[\omega A_1(\omega) + \omega x_2 B_1(\omega) - 2(1-\nu'')B_2(\omega)]\sinh\omega x_2 + [\omega A_2(\omega) + \omega x_2 B_2(\omega) - 2(1-\nu'')B_1(\omega)]\cosh\omega x_2\}\cos\omega x_1 d\omega,$$

$$\tau''_{12}(x_1, x_2) = -\sqrt{2/\pi}\int_0^\infty \omega\{[\omega A_1(\omega) + \omega x_2 B_1(\omega) - (1-2\nu'')B_2(\omega)]\cosh\omega x_2 + [\omega A_2(\omega) + \omega x_2 B_2(\omega) - (1-2\nu'')B_1(\omega)]\sinh\omega x_2\}\sin\omega x_1 d\omega,$$

$$2\mu'' u''_1(x_1, x_2) = -\sqrt{2/\pi}\int_0^\infty \omega\{[A_1(\omega) + x_2 B_1(\omega)]\sinh\omega x_2 + [A_2(\omega) + x_2 B_2(\omega)]\cosh\omega x_2\}\sin\omega x_1 d\omega,$$

$$2\mu'' u''_2(x_1, x_2) = \sqrt{2/\pi}\int_0^\infty \{[\omega A_1(\omega) + \omega x_2 B_1(\omega) - (3-4\nu'')B_2(\omega)]\cosh\omega x_2 + [\omega A_2(\omega) + \omega x_2 B_2(\omega) - (3-4\nu'')B_1(\omega)]\sinh\omega x_2\}\cos\omega x_1 d\omega. \tag{13}$$

Using the traction conditions (5) and (6) on the upper surface

$$\omega^2 A_1(\omega) = -(4\omega h + \sinh 4\omega h)\omega B_1(\omega)/2 + [(1-2\nu'') - \sinh^2 2\omega h]\omega B_2(\omega) + (P/\sqrt{2\pi})\sinh 2\omega h\cos\omega c,$$

$$\omega^2 A_2(\omega) = [(1-2\nu'') + \cosh^2 2\omega h]\omega B_1(\omega) + \tfrac{1}{2}(\sinh 4\omega h - 4\omega h)\omega B_2(\omega) - (P/\sqrt{2\pi})\cosh 2\omega h\cos\omega c, \tag{14}$$

is obtained. Now applying the traction conditions (7) and (8) to the lower surface

$$\Delta(\omega)\omega B_1(\omega) = \sinh^2 2\omega h\,\phi(\omega) + (4\omega h - \sinh 4\omega h)\psi(\omega)/2 + (P/\sqrt{2\pi})\,2\omega h\sinh 2\omega h\cos\omega c,$$

$$\Delta(\omega)\omega B_2(\omega) = -(4\omega h + \sinh 4\omega h)\phi(\omega)/2 + \sinh^2 2\omega h\,\psi(\omega) - (P/\sqrt{2\pi})(\sinh 2\omega h + 2\omega h\cosh\omega h)\cos\omega c,$$

$$\Delta(\omega) = (2\omega h)^2 - \sinh^2(2\omega h), \tag{15}$$

is determined where

$$\phi(\omega) = \sqrt{2/\pi}\int_0^a \phi_1(x_1)\cos\omega x_1 dx_1, \quad \phi_1(x_1) = \tau''_{22}(x_1, 0),$$

$$\psi(\omega) = \sqrt{2/\pi}\int_0^a \phi_2(x_1)\sin\omega x_1 dx_1, \quad \phi_2(x_1) = \tau''_{12}(x_1, 0)$$

$$\phi_3(x_1) = \mu'' u''_{2,1}(x_1, 0), \quad \phi_4(x_1) = \mu'' u''_{1,1}(x_1, 0). \tag{16}$$

Evaluating (13) at $x_2 = 0$, using (14)–(16), then interchanging the orders of integration, and finally using

$$\sqrt{2/\pi}\int_0^\infty \phi(\omega)\cos\omega x_1 d\omega = \phi_1(x_1),$$

$$\sqrt{2/\pi}\int_0^\infty \phi(\omega)\sin\omega x_1 d\omega = -\frac{1}{\pi}\int_{-a}^a \frac{\phi_1(t)}{t - x_1}dt,$$

$$\sqrt{2/\pi}\int_0^\infty \psi(\omega)\cos\omega x_1 d\omega = \frac{1}{\pi}\int_{-a}^a \frac{\phi_2(t)}{t - x_1}dt,$$

$$\sqrt{2/\pi}\int_0^\infty \psi(\omega)\sin\omega x_1 d\omega = \phi_2(x_1), \tag{17}$$

gives

$$\phi_4(x_1) = \gamma_2\phi_1(x_1) - \gamma_1\int_{-a}^a K''_{11}(x_1, t)\phi_1(t)dt - \gamma_1\int_{-a}^a\left[\frac{1}{t - x_1} - K''_{12}(x_1, t)\right]\phi_2(t)dt - P\gamma_1 L_1(x_1),$$

$$\phi_3(x_1) = -\gamma_1\int_{-a}^a\left[\frac{1}{t - x_1} + K''_{21}(x_1, t)\right]\phi_1(t)dt - \gamma_2\phi_2(x_1) + \gamma_1\int_{-a}^a K''_{22}(x_1, t)\phi_2(t)dt - P\gamma_1 L''_2(x_1), \tag{18}$$

where

$$K''_{ij}(x_1, t) = \int_0^\infty k_{ij}(2\omega h)R_i(\omega x_1)R_j(\omega t)d\omega,$$

$$L''_i(x_1) = \int_0^\infty l_i(2\omega h)R_i(\omega x_1)\cos\omega c d\omega,$$

$$k_{11}(\omega) = \omega^2/\alpha(\omega), \quad k_{12}(\omega) = (\omega^2 - \omega + e^{-\omega}\sinh\omega)/\alpha(\omega),$$
$$k_{21}(\omega) = (\omega^2 + \omega + e^{-\omega}\sinh\omega)/\alpha(\omega), \quad k_{22}(\omega) = k_{11}(\omega),$$
$$l_1(\omega) = \omega\sinh\omega/\alpha(\omega), \quad l_2(\omega) = (\omega\cosh\omega + \sinh\omega)/\alpha(\omega),$$
$$\alpha(\omega) = \omega^2 - \sinh^2\omega,$$
$$R_1(\omega) = \cos(\omega), \quad R_2(\omega) = \sin(\omega). \tag{19}$$

Due to the behavior of $k_{11}(\omega)$, $k_{21}(\omega)$, $l_1(\omega)$, $l_2(\omega)$ in the limit as $\omega \to 0$, the corresponding integrals $K''_{11}(x_1, t)$, $K''_{21}(x_1, t)$, $L''_1(x_1)$, $L''_2(x_1)$ do not exist. This situation can be rectified by writing

$$\int_{-a}^a K''_{11}(x_1, t)\phi_1(t)dt + PL''_1(x_1) = \int_{-a}^a \overline{K}''_{11}(x_1, t)\phi_1(t)dt,$$

$$\int_{-a}^a K''_{21}(x_1, t)\phi_1(t)dt + PL''_2(x_1) = \int_{-a}^a \overline{K}''_{21}(x_1, t)\phi_1(t)dt, \tag{20}$$

where

$$\overline{K}''_{11}(x_1, t) = \int_0^\infty [k_{11}(\omega h)\cos\omega t - l_2(\omega)\cos\omega c]\cos\omega x_1 d\omega,$$

$$\overline{K}''_{21}(x_1, t) = \int_0^\infty [k_{21}(\omega h)\cos\omega t - l_2(\omega)\cos\omega c]\sin\omega x_1 d\omega, \tag{21}$$

in which the resultant condition

$$\int_{-a}^a \phi_1(t)dt = -P \tag{22}$$

has been used in order to combine the two integrands on the left-hand side of (20) to obtain the properly defined integrands on the right side of (20).

Semi-Infinite Strip

The solution for a semi-infinite strip with arbitrary end conditions has been obtained by Bogy [1]. The resulting equations are

$$\mu' u'_{1,1}(x_1, 0) = (\gamma_3/2) \int_{-a}^{a} \left[\frac{a_{11}}{t - x_1} + K'_{11}(x_1, t)\right] \tau'_{12}(t, 0)dt - \mu'\gamma_3 \int_{-a}^{a} \left[\frac{a_{12}}{t - x_1} + K'_{12}(x_1, t)\right] u'_{2,1}(t, 0)dt + \frac{pv'}{4a},$$

$$\tau'_{22}(x_1, 0) + \gamma_3 \int_{-a}^{a} \left[\frac{a_{21}}{t - x_1} + K'_{21}(x_1, t)\right] \tau'_{12}(t, 0)dt - 2\gamma_3 \int_{-a}^{a} \left[\frac{a_{22}}{t - x_1} + K'_{22}(x_1, t)\right] \mu' u'_{2,1}(t, 0)dt = -\frac{P}{2a}, \quad (23)$$

where

$$\gamma_3 = 1/\pi(1 - \nu').$$

The constants a_{ij} and the kernels $K'_{ij}(x_1, t)$ are given in [1], and for brevity will not be repeated here.

Continuity at Interface

Applying the continuity conditions (9)–(12) across the interface, the following system of singular integral equations is obtained:

$$\sum_{j=1}^{3} \int_{-a}^{a} \left[\frac{b_{ij}}{t - x_1} + M_{ij}(x_1, t)\right] \phi_j(t)dt + \sum_{j=1}^{3} c_{ij}\phi_j(t)dt = f_i(x_1) \quad i = 1, 2, 3 \quad (24)$$

where

$$b_{12} = -(k\gamma_3 a_{11}/2 + \gamma_1), \quad b_{13} = \gamma_3 a_{12}, \quad b_{21} = \gamma_1,$$

$$b_{32} = a_{21}\gamma_3, \quad b_{33} = -2\gamma_3 a_{22}/k, \quad b_{11} = b_{22} = b_{23} = b_{31} = 0,$$

$$c_{11} = c_{22} = \gamma_2, \quad c_{23} = c_{31} = 1, \quad c_{12} = c_{13} = c_{21} = c_{32} = c_{33} = 0,$$

$$M_{11}(x_1, t) = -\gamma_1 \overline{K}''_{11}(x_1, t),$$

$$M_{12}(x_1, t) = \gamma_1 K''_{12}(x_1, t) - k\gamma_3 K'_{11}(x_1, t)/2,$$

$$M_{13}(x_1, t) = \gamma_3 K'_{12}(x_1, t), \quad M_{21}(x_1, t) = \gamma_1 \overline{K}''_{21}(x_1, t),$$

$$M_{22}(x_1, t) = -\gamma_1 K''_{22}(x_1, t), \quad M_{23}(x_1, t) = 0, \quad M_{31}(x_1, t) = 0,$$

$$M_{32}(x_1, t) = \gamma_3 K'_{21}(x_1, t), \quad M_{33}(x_1, t) = -(2\gamma_3/k)K'_{22}(x_1, t),$$

$$f_1(x_1) = Pkv'/4a, \quad f_2(x_1) = 0, \quad f_3(x_1) = -P/2a, \quad (25)$$

where $k = \mu''/\mu'$. The resultant conditions

$$\int_{-a}^{a} \phi_2(t)dt = 0, \quad (26)$$

$$\int_{-a}^{a} \phi_3(t)dt = 0, \quad (27)$$

along with (22) are also necessary.

Analysis of Singular Integral Equations and Numerical Solution

The system of singular integral equations (24) must now be analyzed in order to determine the order of the singularity at the corners. We note that the kernels $K''_{ij}(x_1, t)$ are bounded whereas $K'_{ij}(x_1, t)$ have regular as well as singular parts as discussed in [1]. (That analysis will not be repeated here.) Following the procedure of Muskhelishvili [7, chapter 4], we define $H_j(t)$ through

$$\phi_j(x_1) = H_j(x_1)/(a^2 - x_1^2)^\gamma = \frac{H_j(a)}{(2a)^\gamma (a - x_1)^\gamma} + \frac{H_j(-a)}{(2a)^\gamma (a + x_1)^\gamma} + O(1) \quad \text{as} \quad x_1 \to \pm a \quad (28)$$

and it follows that

$$\frac{1}{\pi} \int_{-a}^{a} \frac{\phi_j(t)}{t - x_1} dt = \frac{H_j(-a) \cot \pi\gamma}{(2a)^\gamma (a + x_1)^\gamma} - \frac{H_j(a) \cot \pi\gamma}{(2a)^\gamma (a - x_1)^\gamma} + O(1) \quad \text{as} \quad x_1 \to \pm a \quad (29)$$

where γ is the order of the singularity. Then we rewrite (24) so that only the singular terms are on the left-hand side, use (28) and (29), and take the limit as $x_1 \to a$. Since the sum of the singular terms on the left-hand side of each equation must be bounded, a 3×3 set of homogeneous equations results. Setting the determinant equal to zero yields the order of the singularity (γ). The result is

$$2 \sin^2(\pi\gamma) [1 + \cos(\pi\gamma) - 2(1 - \gamma)^2]\beta^2 + 4(1 - \gamma)^2 \sin^2(\pi\gamma)\alpha\beta + [1 + \cos(\pi\gamma) - 2(1 - \gamma)^2]\alpha^2/2 - 4(1 - \gamma)^2 \sin^2(\pi\gamma)\beta + [4(1 - \gamma)^2 \sin^2(\pi\gamma) + 1 + \cos(\pi\gamma) - 2(1 - \gamma)^2 - \sin^2(\pi\gamma)]\alpha + 1/2 - (3/2) \cos(\pi\gamma) + 2 \cos^3(\pi\gamma) - (1 - \gamma)^2 = 0, \quad (30)$$

which is written in terms of the composite material parameters α, β introduced by Dundurs [8]

$$\alpha = \frac{(1 - \nu'') - (1 - \nu')k}{(1 - \nu'') + (1 - \nu')k}, \quad \beta = \frac{(1 - 2\nu'') - (1 - 2\nu')k}{(1 - \nu'') + 2(1 - \nu')k}. \quad (31)$$

The details involved in obtaining (30) have been omitted since the singular terms are similar to those encountered in [4], in which a semi-infinite strip with a half plane is treated. The results also agree with the bonded wedge solutions of Bogy [9].

Having determined γ, we now proceed to solve (24) numerically using the method of Erdogan and Gupta [6]. Applying that method to (24) along with (22), (26), (27) yields a system of $3N$ linear algebraic equations

$$\sum_{j=1}^{3} \sum_{J=1}^{N} A_J \left[\frac{b_{ij}}{\tau_J - y_K} + aM_{ij}(ay_K, a\tau_J)\right] \overline{\phi}(\tau_J) + \sum_{j=1}^{3} C_{ij}\overline{\phi}_j(y_K)/(1 - y_K^2)^\gamma = f_i(hy_K),$$

$$i = 1, 2, 3, \quad K = 1, 2, \ldots N - 1 \quad (32)$$

$$\sum_{J=1}^{N} A_J \overline{\phi}_j(\tau_J) = d_j, \quad j = 1, 2, 3, \quad (33)$$

where

$$d_1 = -P/a, \quad d_2 = d_3 = 0,$$

$$\overline{\phi}_j(\tau) = \phi_j(a\tau)(1 - \tau^2)^\gamma.$$

The quadrature points τ_J, $J = 1, 2, \ldots N$ are the N roots of the Jacobi polynomial $P_N^{(-\gamma,-\gamma)}(\tau)$, the collocation points y_K, $K = 1, 2, \ldots N - 1$ are the $N - 1$ roots of $P_{N-1}^{(1-\gamma,1-\gamma)}(y)$ and the weights of the Gaussian quadrature are

$$A_J = -\frac{2(N - \gamma + 1)}{(N + 1)!(N - 2\gamma + 1)} \frac{\Gamma^2(N - \gamma + 1)}{\Gamma(N - 2\gamma + 1)} \times \frac{2^{-2\gamma}}{P'_N{}^{(-\gamma,-\gamma)}(\tau_J) P_{N+1}^{(-\gamma,-\gamma)}(\tau_J)} \quad (34)$$

Because the equations are of the second kind there are $3N$ equations but $6N$-3 unknowns. We expand $\overline{\phi}_j(\tau)$ in terms of Jacobi polynomials

$$\overline{\phi}_j(\tau) = \sum_{I=0}^{N-1} G_{jI} P_I^{(-\gamma,-\gamma)}(\tau),$$

in order to reduce the number of unknowns to $3N$ [4].

Results and Discussion

Using the method just described, numerical results are obtained for the three material composites shown in Fig. 2. These correspond to some of the material pairs used in the strip half-plane solution [4]. The results are shown in Figs. 3–8 in terms of the stress-intensity factors K_{22}, K_{21}, defined as

$$K_{22} = -\lim_{x_1 \to a} (a^2 - x_1^2)^\gamma \tau_{22}(x_1, 0),$$

$$K_{21} = -\lim_{x_1 \to a} (a^2 - x_1^2)^\gamma \tau_{21}(x_1, 0). \quad (35)$$

	k	ν'	ν''	α	β	γ
①	1	0.5	0.5	0.0	0	0.456
②	1/9	0.5	0.5	0.8	0	0.491
⑤	9	0.5	0.5	-0.8	0	0.416

Fig. 2 Material composites used

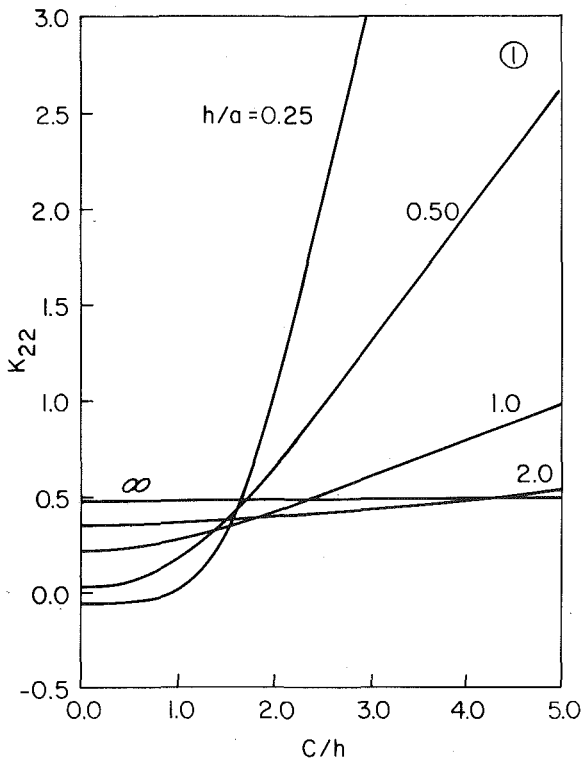


Fig. 3 Stress-intensity factor K_{22} versus c/h for various width ratios (composite No. 1)

Evaluating these quantities numerically gives

$$K_{22} = -a^{\gamma} \sum_{I=0}^{N-1} G_{1I} P_I^{(-\gamma,-\gamma)}(1),$$

$$K_{21} = -a^{\gamma} \sum_{I=0}^{N-1} G_{2I} P_I^{(-\gamma,-\gamma)}(1). \tag{36}$$

In Fig. 3 is shown the variation of stress-intensity factor K_{22}, with the support location c/h for different width ratios h/a and with identical materials. Note the variation of stress-intensity factor for equal width strips. As h/a increases the variation of K_{22} with support location decreases, and K_{22} becomes a constant for very large h/a. This corresponds to the strip half plane results [4]. Conversely, for smaller width ratios K_{22} depends very strongly on c/h. This is because for large c/h the correspondingly larger bending moment causes a greater bowing effect, resulting in the load being transmitted closer to the corners of the bond. For sufficiently small values of c/h, the thin infinite strip allows the bending effect to be reversed, actually resulting in small tensile regions at the corners. This transition from tension to compression will be discussed later. Fig. 4 shows similar results for the stress-intensity factor K_{21}. They are qualitatively similar to Fig. 3 and will not be discussed in detail.

Fig. 5 shows the variation of K_{22} with c/h for composite No. 2, which corresponds to a relatively stiff semistrip. Large values of c/h correspond to even larger K_{22} than in Fig. 3 due to the combination of a stiffer indenter along with the larger bending effect. Also large values of h/a yield greater K_{22} than in Fig. 3 due to a less uniform distribution of normal stress across the bond. Qualitatively similar results are shown in Fig. 6 for K_{21}.

In Fig. 7 is shown the variation of stress-intensity factor K_{22} with c/h for composite No. 5, which corresponds to a relatively stiff infinite strip. The variation of K_{22} with c/h is not as pronounced as in the

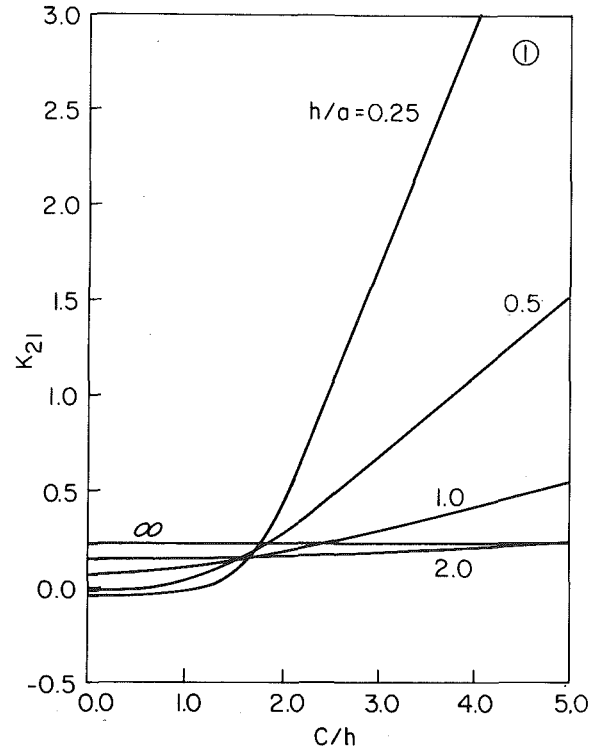


Fig. 4 Stress-intensity factor K_{21} versus c/h for various width ratios (composite No. 1)

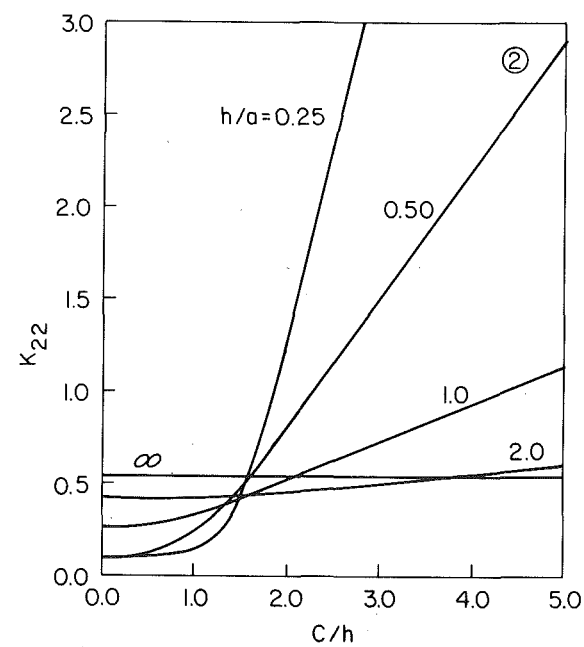


Fig. 5 Stress-intensity factor K_{22} versus c/h for various width ratios (composite No. 2)

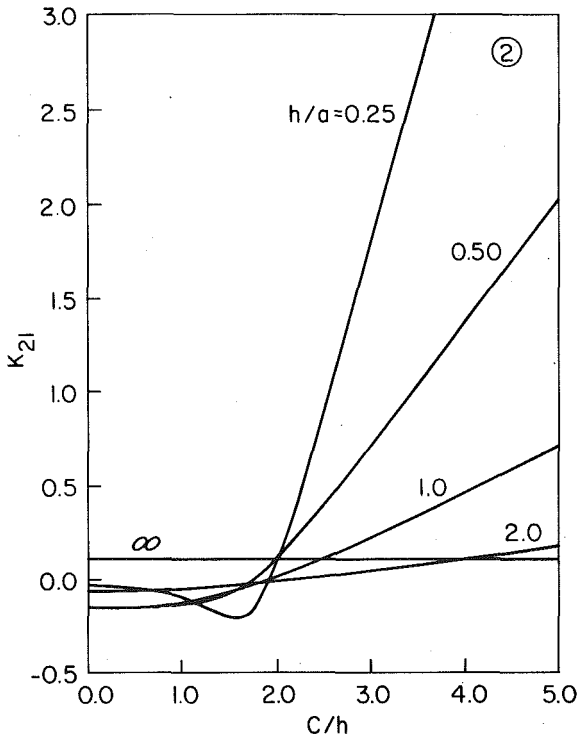


Fig. 6 Stress-intensity factor K_{21} versus c/h for various width ratios (composite No. 2)

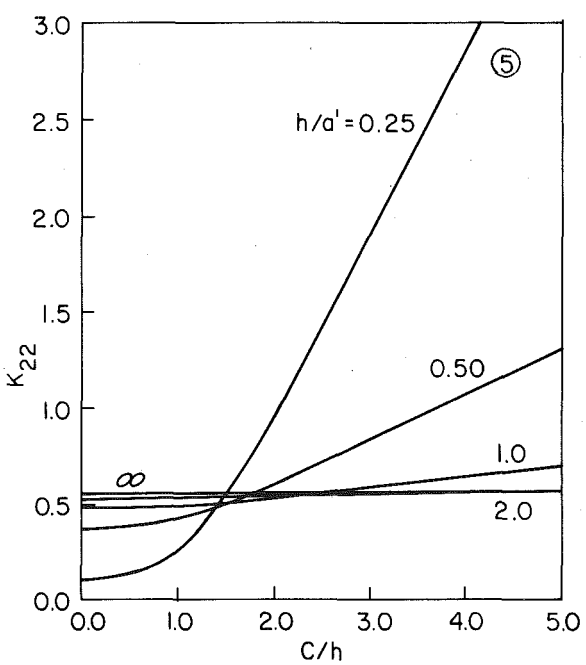


Fig. 7 Stress-intensity factor K_{22} versus c/h for various width ratios (composite No. 5)

other cases since the stiffness of the infinite strip tends to shelter the semistrip from the direct effects of the support locations. Also note that the results for $h/a = 2.0$ are are barely distinguishable from those for a semistrip and half plane [4]. Related results are shown in Fig. 8 for K_{21}. In both cases the stiffness of the infinite strip is sufficient to prevent either tensile regions or a change in the sign of the shear stress near the corners.

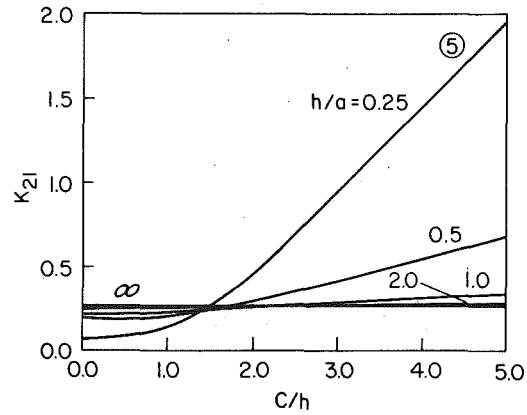


Fig. 8 Stress-intensity factor K_{21} versus c/h for various width ratios (composite No. 5)

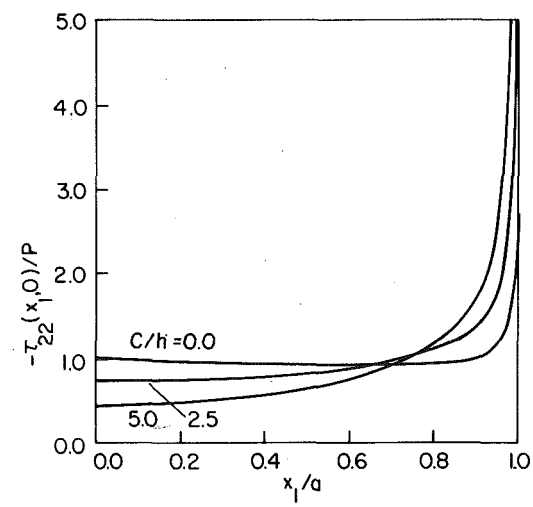


Fig. 9 Normal stress variation along the interface for $h/a = 1$ and various values of c/h (composite No. 1)

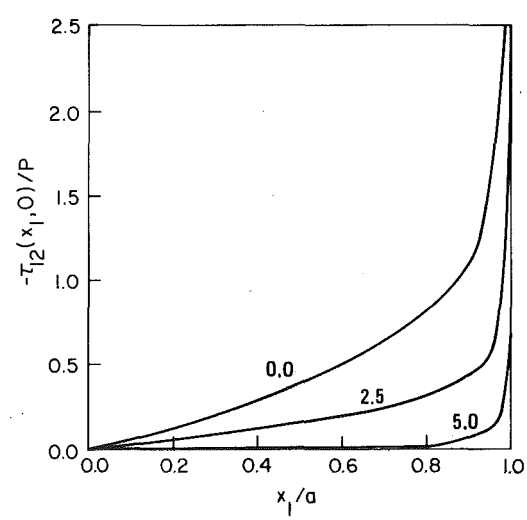


Fig. 10 Shear stress variation along the interface for $h/a = 1$ and various values of c/h (composite No. 1)

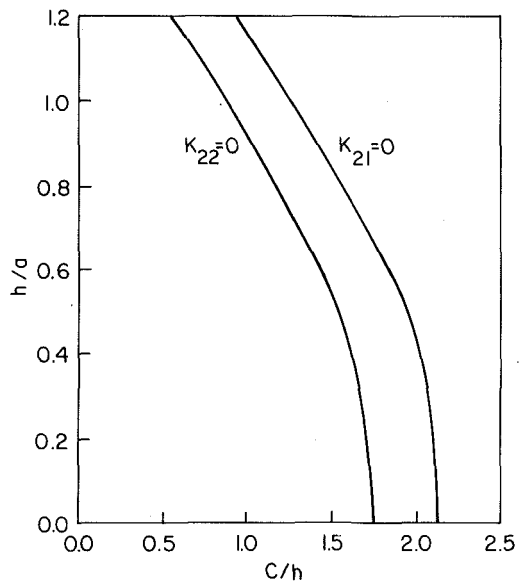


Fig. 11 Values of c/h and h/a such that $K_{21} = 0$ or $K_{22} = 0$

The distribution of normal and shear stress along the bond is determined by

$$\phi_j(t) = (1 - \tau^2)^{-\gamma} \sum_{I=0}^{N-1} G_{jI} P_I^{(-\gamma,-\gamma)}(\tau), \quad j = 1, 2 \tag{37}$$

The results for the normal stress are shown in Fig. 9 for various support locations with $h/a = 1.0$ and identical materials (composite No. 1). Notice that as c/h increases, the distribution of normal stress becomes less uniform. This is due to the bending effect discussed earlier which causes the normal stress to be greater nearer the corners and less in the central region. Fig. 10 shows the corresponding results for the shear stress distribution.

As mentioned earlier, for each sufficiently small value of h/a there exists a value of c/h for which the stress-intensity factor vanishes. This is because the flexibility of the infinite strip allows a small enough value of c/h to cause the strip to bend upward rather than downward. Fig. 11 shows the values of h/a and c/h for which either K_{22} or K_{21} vanishes (composite No. 1). For values of c/h to the left of these curves, the sign of the stress-intensity factor is negative.

Acknowledgments

The author is grateful to the National Science Foundation for their support during part of this research (ENG 77-22181).

References

1 Bogy, D. B., "Solution of the End Problem for a Semi-Infinite Elastic Strip," *ZAMP*, Vol. 26, 1975, pp. 749–769.

2 Bogy, D. B., "The Plane Solution for Joined Dissimilar Elastic Semistrips Under Tension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, 1975, pp. 93–98.

3 Adams, G. G., and Bogy, D. B., "The Plane Solution for Bending of Joined Dissimilar Elastic Semi-Infinite Strips," *International Journal of Solids and Structures*, Vol. 12, 1976, pp. 239–249.

4 Adams, G. G., and Bogy, D. B., "The Plane Solution for the Elastic Contact Problem of a Semi-Infinite Strip and Half Plane," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, 1976, pp. 603–607.

5 Adams, G. G., and Bogy, D. B., "The Plane Symmetric Contact Problem for Dissimilar Elastic Semi-Infinite Strips of Different Widths," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 604–610.

6 Erdogan, F., and Gupta, G. D., "On the Numerical Solution of Singular Integral Equations," *Quarterly of Applied Mathematics*, Vol. 30, 1972, pp. 525–534.

7 Muskhelishvili, N. I., *Singular Integral Equations*, P. Noordhoff, 1958.

8 Dundurs, J., Discussion, ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, 1969, p. 650.

9 Bogy, D. B., "Two Edge-Bonded Elastic Wedges of Different Materials and Wedge Angles Under Surface Tractions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 38, 1971, pp. 377–384.

H. P. Rossmanith

II. Institut für Mechanik,
Technical University Vienna,
Karlsplatz 13,
A—1040,
Vienna, Austria

The Dynamic Three-Parameter Method for Determination of Stress-Intensity Factors From Dynamic Isochromatic Crack-Tip Stress Patterns

Correction methods for the determination of dynamic stress-intensity factors from isochromatic crack-tip stress patterns are developed within the framework of a Westergaard-type stress-function analysis where higher-order terms of the series expansions of the stress functions are retained. The addition σ_{ox} to the extensional stress σ_x, is regarded as a first correction term, and the far-field correction term which is proportional to $r^{1/2}$ is referred to as the β-correction. The β-term represents effects that are due to particular loading systems and situations including finite specimen boundaries. The associated method to determine K can be termed a three-parameter method since it contains K, α, and β as parameters. The correction methods, i.e., velocity correction and higher-order term corrections, permit modification of the "static" crack velocity versus stress-intensity factor (c-K) relationship by correcting the static K for the influence of crack speed and higher-order terms. The results show that both corrections assist the interpretation of current photoelastic c-K-data even though the crack speeds do not exceed one third of the shear wave speed.

Introduction

In the past substantial effort has been devoted to the determination of crack-tip stress fields and associated stress-intensity factors by using photoelastic techniques.

Post [1] and Post and Wells [2] in the early 50's were the first investigators to apply photoelastic methods to fracture mechanics. Irwin [3] in a discussion to reference [2] showed that the stress-intensity factor K could be determined from a single isochromatic fringe loop at the tip of a crack. According to Irwin's two-parameter method, the stress-intensity factor K and an additional uniform stress field σ_{ox} are functions of the radius r_{max} of the fringe loop and its angle of tilt θ_{max} as defined in Fig. 1. The two-parameter method and its modifications [4, 5] for determining the stress-intensity factors from photoelastic isochromatic fringe data have been critically reviewed recently by Etheridge and Dally [6].

Kobayashi, et al. [7], have shown that isochromatic fringe loops obtained analytically within the framework of the static two-parameter method match to a high degree the shape of the dynamic isochromatic fringe loops if small loops (of higher order) close to the crack tip are employed. They recommend that small fringe loops whose apogee distance, r_{max}, is about 2–5 mm should be taken for measurement purposes.

There are, however, practical objections to use of small fringe loops, such as the transition of the state of stress, nonlinear effects close to the crack tip, and uncertain localization of the exact crack-tip position. Regarding all the uncertainties included in experimental close-to-crack-tip measurements it is desirable to abandon the crack-tip vicinity and take measurements from larger loops. These fringe loops, however, exceed the range of applicability of the singular solution and a two-parameter representation of the stress field may not be adequate.

It is thus natural to consider higher-order terms in the expansion of the Westergaard functions which enable one to analyze fringe loops further away from the crack tip. The addition σ_{ox} to the extensional stress, σ_x, which is proportional to r to the zero power and thus has

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, September, 1979; final revision, February, 1980.

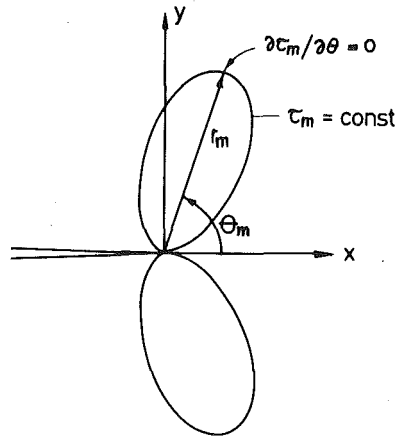


Fig. 1 Definition of geometrical parameters associated with an isochromatic fringe loop

the appearance of a superimposed uniform stress is regarded as a first correction term and will be referred to here as the α-correction. A natural choice for a second far-field correction term would be one proportional to $r^{1/2}$ and will be referred to here as the β-correction. The associated method to determine K can be termed a three-parameter method since it contains K, α, and β as parameters. Physically, α and β account for significant influences on the stress pattern near the crack tip of the loading and geometry of the specimen investigated. Irwin, et al. [8], and Etheridge and Dally [9] compared the two and three-parameter dynamic solutions, respectively, introduced several approximate relations pertaining to the dynamic analysis and proved the convergence of the dynamic solution to the static solution. The comparison between static and dynamic calculations then showed that the static solution would overestimate the K-value only slightly for the high velocity range where $c/c_2 > 0.2$. This was supported by the strong similarity between the computer-generated static and dynamic fringe loops which showed practically no difference.

The use of the exact formulas, however, reveals appreciable differences between the calculated static and dynamic K-values for moderate and high crack propagation velocities. The results exhibit smaller K-values for the dynamic case and the difference between static and dynamic K-value exceeds 5 percent in almost all cases.

Analysis

The components σ_x, σ_y, and τ_{xy} of the stress field associated with a constant velocity semi-infinite tensile crack opened by a traveling line force may be represented in the form [10]

$$\sigma_x = A\mu\,\{(2r_1^2 - r_2^2 - 1)\,\mathrm{Re}\,Z_1 - \Omega\,\mathrm{Re}\,Z_2\} + \sigma_{ox}$$

$$\sigma_y = A\mu\,\{-(1 + r_2^2)\,\mathrm{Re}\,Z_1 + \Omega\,\mathrm{Re}\,Z_2\}$$

$$\tau_{xy} = A\mu\,2r_1\,\{\mathrm{Im}\,Z_2 - \mathrm{Im}\,Z_1\} \tag{1}$$

The coordinates x and y refer to a Cartesian coordinate system whose origin remains fixed at the crack tip. The expressions $r_j^2 = 1 - (c/c_j)^2$ $(j = 1, 2)$ and $\Omega = 4r_1r_2/(1 + r_2^2)$ have been employed, where c_1 and c_2 are the longitudinal (or P) and transversal (or S) wave velocities for plane waves in an infinite medium.

The stress functions Z_j are chosen (following Irwin's procedure [11] of factoring out $K/\sqrt{2\pi z_j}$) as

$$Z_j = \frac{K}{\sqrt{2\pi z_j}}\left[1 + \beta\frac{z_j}{r_s} + O(z^2)\right] \tag{2}$$

The $K/\sqrt{2\pi z_j}$ term describes the crack-tip singularity of any opening mode crack where the nonlinear zone containing the fracture process is relatively small, the plate exhibits uniform thickness and has no finite boundaries. The factor β models the effect of the presence of near boundaries, and r_s is a reference length-factor constant which can be selected to equal the size, r_m, of a particular isochromatic fringe loop, or the crack size, a, etc.

The near crack-tip stress field may be represented by

$$\sigma_x = \frac{A\mu K}{\sqrt{2\pi r_s}}\sqrt{\frac{r_s}{r}}\left\{E_{11} + \beta^* E_{12} + \frac{\alpha^*}{A\mu}\right\}$$

$$\sigma_y = \frac{A\mu K}{\sqrt{2\pi r_s}}\sqrt{\frac{r_s}{r}}\,\{E_{21} + \beta^* E_{22}\}$$

$$\tau_{xy} = \frac{A\mu K}{\sqrt{2\pi r_s}}\sqrt{\frac{r_s}{r}}\,\{E_{31} + \beta^* E_{32}\} \tag{3}$$

where the coefficients E_{ij} depend on the velocity ratios c/c_j and the angular coordinate θ but not on the radial coordinate r/r_s.

The expressions (9) contain three elementary parts: the leading term E_{i1} $(i = 1, 2, 3)$ correspond to the singular near crack-tip stress field, the β^*-term indicates the higher-order term influence when expanding the range of analysis, and the α^*-term which appears in the expression for stress σ_x only is a measure for the degree of biaxiality of the crack-tip stress field.

The introduction of α^* and β^*

$$\alpha^* = \alpha\sqrt{\frac{r}{r_s}} = \frac{\sigma_{ox}\sqrt{2\pi r_s}}{K}\sqrt{\frac{r}{r_s}};\quad \beta^* = \beta\frac{r}{r_s} \tag{4}$$

permits joint parametric studies for varying α and for β for fixed r/r_s as well as for varying r/r_s for α and for β fixed.

Equations (3) reduce in the limit $c \to 0$ to the well-known static crack-tip stress expressions [9] and for $\beta = 0$ to the classical two-parameter crack-tip stress equations [3].

The maximum shear stress τ_m is expressed in terms of the Cartesian stress components as

$$\tau_m^2 = \left[\frac{\sigma_y - \sigma_x}{2}\right]^2 + \tau_{xy}^2. \tag{5}$$

Substitution of equations (3) into equation (5) yields

$$\left[\frac{2\tau_m\sqrt{2\pi r_s}}{2A\mu K}\right]^2\frac{r}{r_s} = \left[F(\beta^*) - \frac{\alpha^*}{2A\mu}\right]^2 + [G(\beta^*)]^2 \tag{6}$$

where F and G now are functions of the angle θ, the velocity ratios c/c_j $(j = 1, 2)$, *and* the parameter β^*:

$$F(\beta^*) = {}^1/_2\,\{E_{21} - E_{11} + \beta^*(E_{22} - E_{12})\}$$

$$G(\beta^*) = E_{31} + \beta^* E_{32}. \tag{7}$$

Irwin observed the geometry of the fringe loops and noted that (Fig. 1)

$$\frac{\partial\tau_m}{\partial\theta}\Big|_{\substack{r = r_{mk}\\ \theta = \theta_{mk}}} = 0\ (k = 1, 2) \tag{8}$$

holds at the apogee $(r = r_{mk},\ \theta = \theta_{mk})$ of each fringe loop.

Differentiating equation (6) with respect to θ and using equation (8) gives

$$\alpha^* = 2A\mu[G(\beta^*)g(\beta^*) + F(\beta^*)]^{r = r_{mk}}_{\theta = \theta_{mk}}\quad (k = 1, 2, \ldots) \tag{9}$$

where

$$(\ )^{\cdot} = \partial(\ )/\partial\theta \quad\text{and}\quad g = \dot{G}/\dot{F}.$$

Inspection shows that equation (9) holds for any fringe loop around the crack tip. The functions F, G, and their derivatives depend essentially on the velocity of the running crack and on β^*.

Combining equations (6) and (9) yields the fundamental equation for the normalized stress-intensity factor K_n [12]

$$K_n \equiv \frac{K}{2\tau_m\sqrt{2\pi r_m}} = \frac{1}{2A\mu}\frac{1}{|G|\sqrt{1 + g^2}} \equiv H(\theta_m, r_m; c, \beta^*) \tag{10}$$

Equation (10) reduces in the static case $(c = 0)$ to

$$K_n = \left\{\sin^2\theta_m(1 - 2\beta^*\cos\theta_m + \beta^{*2}) - 2\alpha^*\sin\theta\left(\sin\frac{3\theta}{2} - \beta^*\sin\frac{\theta}{2}\right) + \alpha^{*2}\right\}^{-1/2} \tag{11}$$

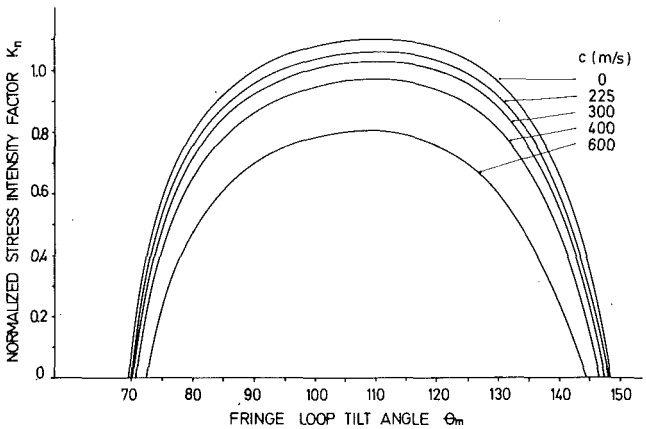


Fig. 2 Normalized stress-intensity factor K_n versus fringe loop tilt angle θ_m relationship for varying crack velocity

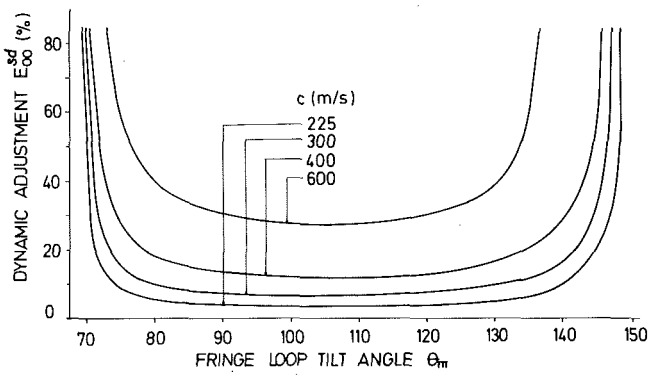


Fig. 3 Dynamic adjustment (or relative error) $E_{00}{}^{sd}$ in percent versus θ_m relationship for varying crack velocities

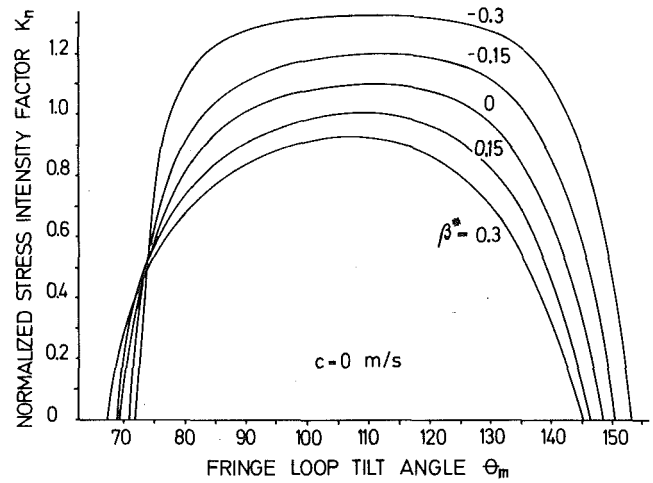


Fig. 4(*a*) Normalized stress-intensity factor K_n versus fringe loop tilt angle θ_m relationship for varying β^*-term for a static situation ($c/c_2 = 0$)

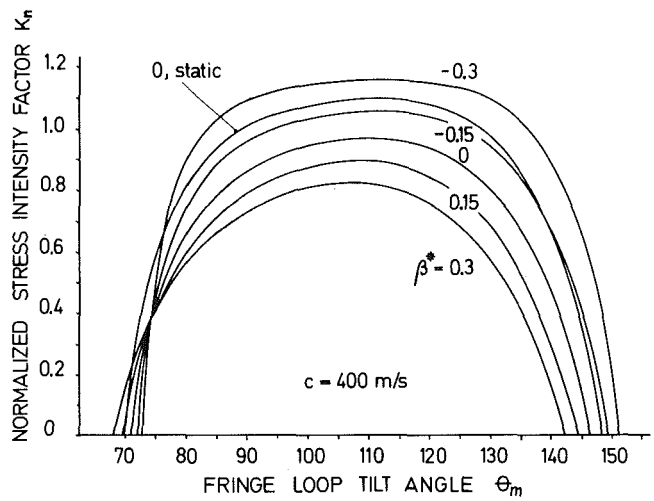


Fig. 4(*b*) Normalized stress-intensity factor K_n versus fringe loop tilt angle θ relationship for varying β^*-term for a medium velocity propagating crack ($c/c_2 \sim 0.2$)

c-K Characterization for Homalite 100

Dynamic crack propagation in birefringent polymers has been the subject of extensive study. Homalite 100 is a commercially available transparent birefringent polymer of appreciable photoelastic sensitivity to be used in crack propagation studies employing high-speed photography in conjunction with dynamic photoelasticity. Thus the preceeding analysis is applied to Homalite 100. The models were fabricated from uniform sheets (12.5 mm). The assumption of plane-stress conditions is appropriate outside of a small region around the crack tip where the problem is a three-dimensional one. The following input data for Homalite 100 have been used: state of plane stress, Poisson' ratio $\nu = 0.31$, $c_1 = 2662$ m/s, $c_2 = 1234$ m/s, and plate wave velocity $c_L = 2148$ m/s.

Dynamic Two-Parameter Method. Results for Homalite 100 for the normalized stress-intensity factor K_n obtained from equation (10) are shown in Fig. 2, as a function of θ_m over the range $69.4° < \theta_m < 148.5°$. This range for the tilt angle θ_m is associated with the main loop of the isochromatic fringe pattern around a running crack tip. The cutoff points of the K_n θ_m-curves, i.e., the intersection of the curves with the abscissa are velocity-dependent and define the range of applicability of the two-parameter method. The difference between static and dynamic stress-intensity factor for different velocities depicted in Fig. 3 reveals that the static analysis provides acceptable results for low and medium velocity crack propagation. For high-speed running cracks the differences in K between static and dynamic analysis are by far too large and the static analysis gives too high values for K.

Dynamic Three-Parameter Method. When higher-order terms are retained in the analysis both the velocity effect and the β-effect influence the distribution of the stress-intensity factor, and their interplay eventually generates complex situations. Figs. 4(*a*) and (*b*) show results pertaining to Homalite 100 for the normalized stress-intensity factor K_n as a function of the tilt angle over the range $68° < \theta_m < 153°$. This range is related to the main loop. Different curves correspond to different β^*-values. For the sake of comparison and completeness the curves for $\beta^* = 0$, $c = 0$ have been added. The K_n versus θ_m-curves for the approximately static case are given in Fig. 4(*a*). An important result obtained in the two-parameter method was that the K-curves for higher velocities lie below the K-curves for low crack velocities. This fact holds also for the three-parameter method and implies that the sets of K-curves for a variety of β^*-values give lower K-values for higher velocities. From this one concludes that the dynamic effect and the β^*-term effect work in opposite directions for $\beta^* < 0$. The magnitude of (negative) β^* required to annihilate the velocity-induced correction depends on the tilt angle. The dynamic effect and the β^*-term effect, however, superimpose for $\beta^* > 0$ and, hence, may yield to appreciable corrections of the K-value.

Corrections Due to High Velocity and Higher-Order Terms. Several dynamic correction factors may be defined by relating the results of different methods used to one another. Figs. 5(*a*) and (*b*) show the dynamic correction factor $K_{00}{}^{sd}$ as a function of crack velocity c for various tilt angles θ_m.

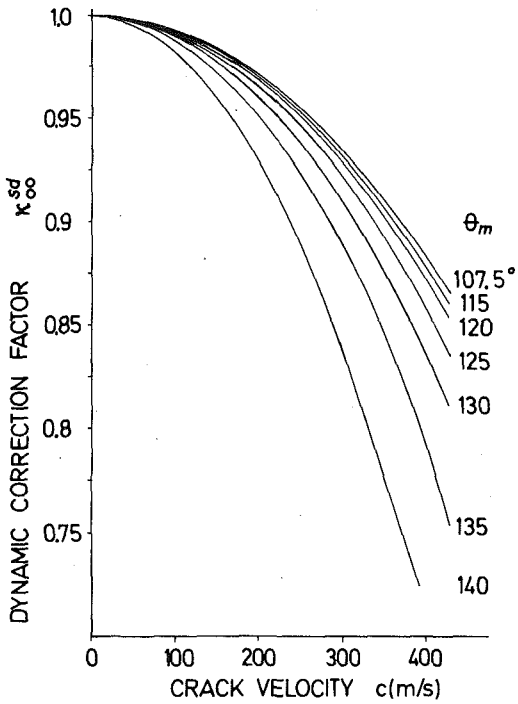


Fig. 5(*a*) Dynamic correction factor $\kappa_{00}{}^{sd}$ versus crack velocity *c* for different fringe loop tilt angles ($75° < \theta_m < 107.5°$)

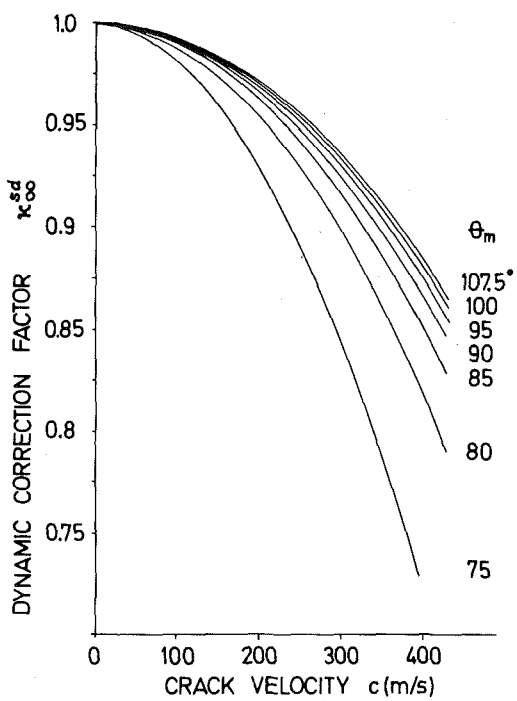
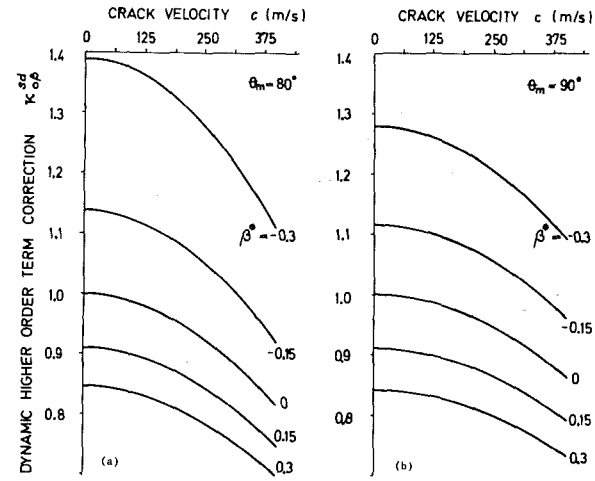
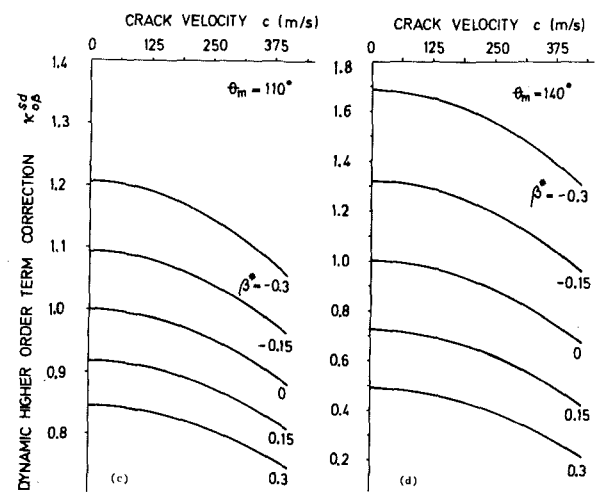


Fig. 5(*b*) Dynamic correction factor $\kappa_{00}{}^{sd}$ versus crack velocity *c* for different fringe loop tilt angles ($107.5° < \theta_m < 140°$)



6(*a*) $\theta_m = 80°$ (forward leaning loops: SEN); (*b*) $\theta_m = 90°$ (straight-up loop: M-CT)



6(*c*) $\theta_m = 110°$ (slightly backward leaning loops: M-CT); (*d*) $\theta_m = 140°$ (strongly backward leaning loops: DCB)

Fig. 6 Combined *c*-β correction factor $\kappa_{0\beta}{}^{sd} = K_\beta{}^d/K_0{}^s$ versus velocity *c* for varying β^* and fixed fringe loop tilt angle θ_m

The combined β^*-c correction factor $K_{o\beta}{}^{sd}$ shown in Figs. 6 (*a–d*) represents the correction which is introduced when the dynamic higher-order stress field around the tip of a propagating crack as determined by the dynamic three-parameter method is approximated by the static stress field associated with the staic two-parameter method. Inspection of Figs. 6(*a–d*) shows that the correction for K becomes inevitable in the high velocity region whenever $\beta^* > 0$. No correction is necessary for situations where the effects of negative β^* and high velocity compensate. Situations of low crack velocity and $\beta^* < 0$, however, require correction in the reverse direction.

A critical examination of the β-terms reveals that the value of β depends on the crack type (Griffith-crack, semi-infinite crack, . .) and loading conditions (e.g., crack-line loading, internal pressure, remote biaxial loading . . .). Unlike in the static case, where it is possible to assign a restricted range of β-values to each specimen configuration, the extraction of an explicit expression for β on the basis of a Taylor-series expansion is not apparent for running cracks. The reason for this may be found in the continuous change of β during the process of crack propagation. Nevertheless, it is believed that some qualitative estimates derived from a staic situation may hold in the dynamic case too [12].

Comments on the Dynamic *c*-*K* Curve

The analysis of the constant speed running crack under Mode-I stress conditions is also valid for any nonuniformly propagating crack because the singular part of the stress field, i.e., the $-1/2$ term in the series expansion of the Westergaard stress function is invariant with

respect to nonuniform crack propagation. Thus a dynamic 2-parameter method gives reasonable K-values also in a situation where the crack velocity changes provided the fracture process zone and the range of influence of the higher-order terms do not overlap or interfere.

When material points close to the crack tip experience an increasing effect of higher-order terms (rapid crack acceleration or velocity changes in small specimens) and/or the fracture process zone is large (rough fracture) higher-order term effects have to be taken into account. Inertia effects of material particles in the extended fracture process zone influence the dynamic fracture behavior when sudden crack velocity changes occur. Moreover, in-plane wave propagation effects cause continuous redistribution of the crack loading which in turn changes the β-term. During a general dynamic fracture process c and β are functions of the crack length or time and the resulting β-behavior is extremely complicated. Nethertheless it is believed that the qualitative features of a quasistatic β-study can be transferred to the dynamic situation of nonuniform crack extension. The K_n versus θ_m tables generated involving corrections due to dynamic effects and higher-order terms, may be utilized helpfully also in the most general dynamic situation provided the crack advances under pure opening mode conditions.

Assume that the fracture process zone of a running crack is relatively small and that changes in crack speed have limited gradients relative to time and crack extension. One can then assume that observed values of c and K are restricted to a natural c-K curve which may be defined as one which is associated with dynamic fracture of a larger specimen where the fracture process zone is confined to the immediate vicinity of the crack tip and higher-order terms have no effect. Any c-K curve obtained by analytical reduction of experimentally obtained data possibly involving velocity effects and higher-order terms is considered a "test" c-K curve. These test c-K curves must not be confused with the natural c-K curve.

Fracture experiments utilizing specimens of any arbitrary shape and size reproduced similar c-K curves within error bounds set by the experiment when a static 2-parameter analysis of data reduction was employed. This fact led to the idea that the c-K curve may be a material property regardless of the type of specimen used to perform the fracture test [13, 14].

The independence of a c-K relationship from specimen geometry was questioned when experimental results obtained by different research teams employing specimens of the same shape, but different in size, did not agree. Large and small size members of a specimen-type family produced different test c-K curves which approximate a natural c-K curve with different degree of accuracy.

A possible explanation for this size effects is offered by the β-term study. Let us consider two equal sets of crack-tip fringe loops, one set in a large specimen and the other set in a small specimen. The specimens are of the same type and subject to similar loading conditions. The set of isochromatics in the smaller specimen is more influenced by the presence of near boundaries than the corresponding set in the larger specimens. This causes different α and β-values for the two sets and therefore different K-values.

A (dynamic) three-parameter based c-K curve is constructed by applying (velocity corrections) and higher-order term corrections to a c-K curve which was obtained by means of the static two-parameter method of K-determination.

The influence of the higher-order terms onto the c-K relationship is more complex than the velocity effect because the β-correction affects the entire c-K curve.

Positive β-values require a correction toward smaller K-values, i.e., velocity correction and β-term correction superimpose and amplify the backward correction of c-K curves. Note, that a positive β-correction shifts the entire β-curve toward lower K-values. The amount of shift depends on the β-value at crack initiation. This is illustrated in Fig. 7(a). SEN-CPL-type specimens are associated with positive initial β-values and therefore require positive β-corrections.

Negative β-values introduce a correction toward larger K-values. Velocity effect and negative β-effect subtract and the combined effect is to increase K in the lower velocity region (where the β-effect dominates) but it may cause K to increase in the high velocity region (where the c-effect may dominate). The root of the c-K curve independent of velocity corrections shifts to higher K-values for negative β-values at crack initiation. SEN-CLL, M-CT, and DCB-specimens are associated with slightly negative initial β-values, and thus require negative β-corrections. This type of correction is shown in Fig. 7(b). Note, that *any* negative β-correction causes a negative slope at the root and the lower part of the stem of the dynamic 3-parameter c-K curve provided the corresponding static 2-parameter c-K curve has a vertical stem.

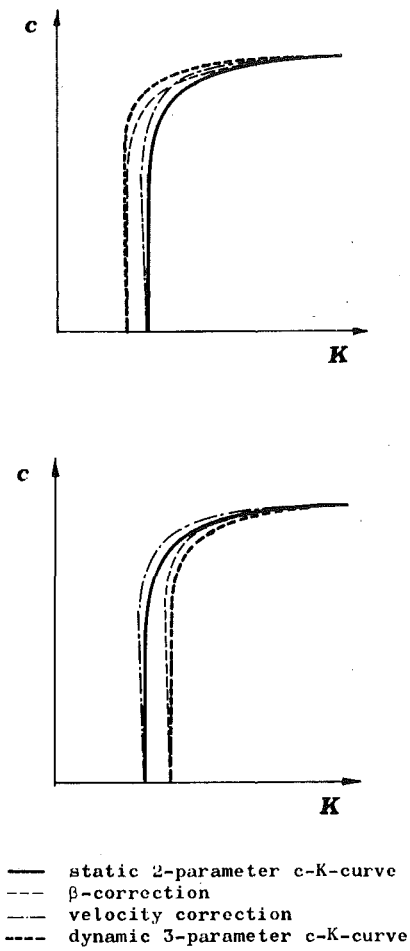


Fig. 7 Influence of the combined c-β correction on the c-K relationship; (a) positive β correction (β^+); (b) negative β correction (β^-)

The upper part of the c-K curve, the so-called "second plateau" where even large changes in K produce small c-changes alters appreciably. The dynamic correction does not change the altitude of the plateau, however, it shifts points on the plateau to the left, i.e., to considerably lower K-values. This implies that unsuccessful branching attempts and successful crack branching occur at lower K-values in SEN specimens than assumed in the past.

Conclusions

The preceding dynamic 3-parameter analysis reveals several important implications for the practical determination of the stress-intensity factor from isochromatic fringe loops.

1 Dynamic K-values are smaller than K-values obtained from a static analysis. Their difference exceeds the 5 percent error found for moderate crack velocities and exceeds the 12 percent error bound for high crack velocities encountered in experiments with Homalite 100.

2 SEN-specimens (forward leaning fringe loops) and DCB-

specimens (strongly backward leaning fringe loops) are very sensitive to velocity effects and a dynamic correction for K becomes inevitable.

3 Dynamic fringe loops are slightly larger in size than the corresponding static counterparts.

4 The incorporation of the β-term into the analysis and the results obtained provide an explanation of the specimen geometry influence. Corrections due to different β-values results in different c-K curves. Dynamic and higher-order term correction amplify (subtract) for positive (negative) values of β.

5 The results imply that test c-K relationships obtained by employing specimens of different type and size are in fact dependent of the specimen geometry. The size effect is predominant.

6 The influence of higher-order terms disappears at the crack tip and hence the use of small isochromatic fringes close to the crack tip has been recommended in the past. Practical objections to use of small fringe loops are the transition of the state of stress in the vicinity of the crack tip and some materials show considerable nonlinear effects close to the crack tip. Moreover, the determination of the exact crack-tip position is often a difficult task for several reasons.

7 Smaller fringe loops give slightly higher values for the stress-intensity factor with β-correction absent. The dynamic K-values associated with unsuccessful branching attempts and successful crack branching in SEN-samples are considerably lower than the K-values obtained from a static analysis.

8 The dynamic analysis provides helpful information for future fracture test specimen design philosophy. It is found that the CT-specimen is least sensitive with respect to dynamic fracture behavior.

Acknowledgments

The author greatly acknowledges the cooperation, guidance, and critical discussions of Prof. Dr. G. R. Irwin during the research project. This work was supported in part by a postdoctoral fellowship from the Max-Kade Foundation, Inc., New York and Grant No. NRC-04-76-112 from the U.S. Nuclear Regulatory Commission to the Photomechanics Laboratory of the Department of Mechanical Engineering, University of Maryland.

References

1 Post, D., "Photoelastic Stress Analysis for an Edge Crack in a Tensile Field," *Proceedings of SESA,* Vol. 12, 1954, pp. 99–116.

2 Wells, A., and Post, D., "The Dynamic Stress Distribution Surrounding a Running Crack—A Photoelastic Analysis," *Proceedings SESA,* Vol. 16, 1958, pp. 69–92.

3 Irwin G. R., Discussion to reference [2], *Proceedings SESA,* Vol. 16, 1958, pp. 93–96.

4 Bradley, W. B., and Kobayashi, A., "An Investigation of Propagating Cracks by Dynamic Photoelasticity," *Experimental Mechanics,* Vol. 10, 1970, pp. 106–113.

5 Schroedl, N. A., and Smith, C. S., "Local Stress Near Deep Surface Flaws Under Cylindrical Loading Fields," *Progress in Flaw Growth and Fracture Toughness Testing,* ASTM STP 536, ATM 45–63, 1973.

6 Etheridge, J. M., and Dally, J. W., "A Critical Review of Methods for Determining Stress-Intensity Factors From Isochromatic Fringes," *Experimental Mechanics,* Vol. 17, 1977, pp. 248–254.

7 Kobayashi, A. S., Wade, R. G., and Bradley, W. B., "Fracture Dynamics of Homalite 100," *Deformation and Fracture of High Polymers,* eds., Kansch, H. H., Haseele, T. A., and Jaffee, R. I., Plenum Press, New York, 1973, pp. 487–500.

8 Irwin, G. R., et al., "A Photoelastic Characterization of Dynamic Fracture," U.S. NRC-Report NUREG-0072, University of Maryland, 1976.

9 Etheridge, J. M., and Dally, J. W., "A Three Parameter Method for Determining Stress-Intensity Factors From Isochromatic Fringe Loops," *Journal of Strain Analysis,* Vol. 13, No. 2, 1978, pp. 91–94.

10 Irwin, G. R., "Relatively Unexplored Aspects of Fracture Mechanics," University of Illinois, T and AM Report 240, Feb. 1963, and "Constant Speed Semi-Infinite Crack," *Lehigh University Lecture Notes,* 1968.

11 Irwin, G. R., "Analysis of Stresses and Strains Near the End of a Crack Traversing a Plate," ASME JOURNAL OF APPLIED MECHANICS, Vol. 24, Sept. 1957, pp. 361–364.

12. Rossmanith, H. P., and Irwin, G. R., "Analysis of Dynamic Isochromatic Crack-Tip Stress Patterns," University of Maryland Report, July 1979, 430 pages.

13 Kobayashi, T., Dally, J. W., and Fourney, W. L., "Influence on Specimen Geometry on Crack Propagation and Arrest Behavior," SESA Fall Meeting, Munich, 1978.

14 Kalthoff, J. F., Beinert, J., and Winkler, S., "Effect of Specimen Geometry on Crack Arrest Toughness, ASME/CSME Pressure Vessels and Piping Conference, Montreal, Canada, June 1978.

S. Itou

Associate Professor,
Department of Mechanical Engineering,
Hachinohe Institute of Technology,
Hachinohe 031, Japan

Transient Response of a Finite Crack in a Strip With Stress-Free Edges

We have considered the problem of determining the dynamic stress distribution in an infinitely long isotropic homogeneous elastic strip containing a Griffith crack which is perpendicular to the edges of the strip. The crack is opened by internal pressure with the Heaviside function time-dependence. By using the Fourier and Laplace transforms, we can solve the problem with a set of dual integral equations in the Laplace transform domain. These equations are solved using the Schmidt method. The Laplace inversion of the stress-intensity factor is carried out numerically.

1 Introduction

From an engineering point of view, the dynamic crack problems are of particular interest, because the dynamic stress-intensity factors are 1.2 or 1.6 times larger than the corresponding static values. From this reason, many studies have been carried out to determine the dynamic stress fields around a finite crack in an infinite elastic solid. Among these, the work carried out by Sih and his coworkers [1–8], Mal [9, 10] and Thau and Lu [11, 12] are immediately useful in designing the various parts of a machine or structure which are composed of brittle materials.

In comparison with the studies that are concerned with a single crack in an infinite elastic solid, relatively little work has been performed concerning problems which involve other boundaries. Recently, the dynamic stress-intensity factors for a centrally cracked rectangular bar under impact loads were computed by the finite-element method [13], the finite-difference methods [14] and the finite-element technique with the calibrated element [15]. Later, an analytical approach to the transient crack problem including the effects of the boundaries was investigated by Chen [16]. He considered the impact response of a finite crack in an elastic strip under antiplane shear, to the edges of which the crack is placed perpendicularly. He solved the mixed boundary-value problem with a Fredholm integral equation of the second kind in the Laplace transform domain and inversed the Laplace transforms using the numerical technique [17]. The dynamic stress-intensity factor for an in-plane problem was also obtained by Chen [18]. In these investigations [16, 18], when the a/h ratio approaches zero, the solutions reduce to those of an infinite medium with $2a$ being width of the crack and $2h$ being the width of the strip.

On the other hand, the dynamic solution obtained by Thau and Lu for a finite crack on which a plane transient compressional wave impinges [12] contains Chen's limiting one. In the symmetric problem, the incident P wave produces a zero shear stress along the surfaces of the crack, and then this solution corresponds with that for $a/h \rightarrow 0.0$ in reference [18]. Quite recently, Kim studied the transient problem for a finite crack whose tips propagate nonuniformly with time due to the action of an arbitrary time-dependent normal load on the face of the crack [19]. The special case for a stationary crack also agreed with that of the limiting solution for $a/h \rightarrow 0.0$ in reference [18]. However, there is disagreement between Chen's results and those in references [12, 19].

In this paper, the same problem which was treated by Chen [18] is reworked using a somewhat different approach. Laplace and Fourier transforms are applied and a mixed boundary-value problem is reduced to dual integral equations by a method similar to that employed by Sneddon and Srivastav [20] for the corresponding static problem. In solving the equations, the crack surface displacement is expanded in a series using Jacobi's polynomials and Schmidt's method is used. This process is quite different from that adopted in references [1–10]. The obtained Laplace transform solution is inverted numerically by the method developed by Miller and Guy [17]. Numerical calculations are carried out for the transient stress-intensity factor.

2 Fundamental Equation

We consider a strip bounded in the x, y-plane by the lines $x = \pm h$ and a finite crack located along the x-axis from $-a$ to $+a$ as shown in Fig. 1.

For the plane elastodynamic problems, the displacement potential functions are usually introduced in the following way:

$$u = \phi_{,x} - \psi_{,y}, \qquad (1)$$
$$v = \phi_{,y} + \psi_{,x},$$

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, May, 1980.

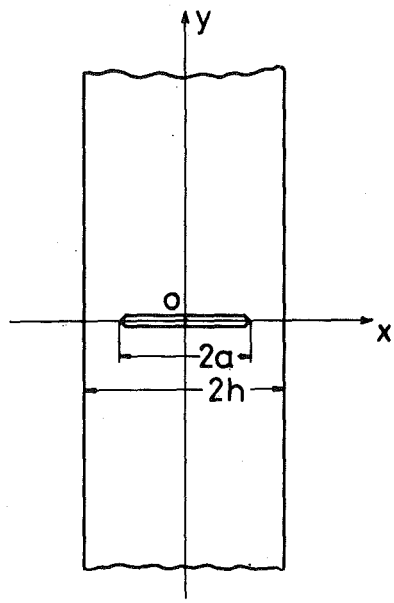


Fig. 1 Geometry and coordinate system

where u and v are defined as the x and y components of the displacement, respectively, and the indices following the comma indicate the partial differentiation with respect to the variable, e.g., $\phi_{,x} = \partial\phi/\partial x$.

Substituting equation (1) into the motion equation reduces it to

$$\phi_{,xx} + \phi_{,yy} = \frac{1}{c_L^2}\phi_{,tt}, \tag{2}$$

$$\psi_{,xx} + \psi_{,yy} = \frac{1}{c_T^2}\psi_{,tt},$$

where the medium is assumed to be homogeneous and isotropic and $c_L = \{(\lambda + 2\mu)/\rho\}^{1/2}$, $c_T = (\mu/\rho)^{1/2}$ are the dilatational and shear wave velocities with λ and μ being the Lamé constants, ρ being the density of the material. The stresses are written in terms of ϕ and ψ as

$$\begin{aligned} \tau_{yy}/(2\mu) &= -\phi_{,xx} + \tfrac{1}{2}\kappa^2(\phi_{,xx} + \phi_{,yy}) + \psi_{,xy}, \\ \tau_{xx}/(2\mu) &= -\phi_{,yy} + \tfrac{1}{2}\kappa^2(\phi_{,xx} + \phi_{,yy}) - \psi_{,xy}, \\ \tau_{yx}/(2\mu) &= \phi_{,xy} + \psi_{,xx} - \tfrac{1}{2}(\psi_{,xx} + \psi_{,yy}), \end{aligned} \tag{3}$$

with

$$\kappa^2 = (c_L/c_T)^2, \tag{4}$$

where elastic constant κ^2 takes the value $2(1-\nu)/(1-2\nu)$ for the plane strain and $2/(1-\nu)$ for the generalized plane stress with ν denoting Poisson's ratio.

The boundary conditions for the problem to be studied are as follows:

$$\begin{aligned} \tau_{yy}^0 &= -P\,\mathrm{H}(t), && \text{for} \quad y = 0, |x| < a, \\ v^0 &= 0, && \text{for} \quad y = 0, a < |x| \leqq h, \\ \tau_{yx}^0 &= 0, && \text{for} \quad y = 0, 0 \leqq |x| \leqq h, \end{aligned} \tag{5}$$

$$\begin{aligned} \tau_{xx} &= 0, \\ \tau_{yx} &= 0, \end{aligned} \quad \text{for} \quad x = \pm h, |y| < \infty, \tag{6}$$

where P is the constant, $\mathrm{H}(t)$ is the Heaviside unit step function and the superscript means that the values with it are those at $y = 0$. Because of the symmetry conditions in equations (5) and (6), it is possible to consider only the problem for the half plane, $y \geqq 0$.

3 Analysis

A Laplace transform pair is defined by equations

$$f^*(s) = \int_0^\infty \exp(-st) f(t)dt, \tag{7}$$

$$f(t) = \frac{1}{2\pi i}\int_{\mathrm{Br}} \exp(st) f^*(s)ds, \tag{8}$$

where the second integral is over the Bromwich path. Applying equation (7) to equation (2) results in

$$\begin{aligned} \phi_{,xx}^* + \phi_{,yy}^* &= \frac{s^2}{c_L^2}\phi^*, \\ \psi_{,xx}^* + \psi_{,yy}^* &= \frac{s^2}{c_T^2}\psi^*. \end{aligned} \tag{9}$$

The solution for equation (9) will be in the following forms in terms of its unknown coefficients $A_1(s, \xi)$, $A_2(s, \xi)$, $B_1(s, \zeta)$, and $B_2(s, \zeta)$

$$\begin{aligned} \phi^* &= 2\int_0^\infty A_1(s, \xi)\exp(-\gamma_1 y)\cos(\xi x)d\xi \\ &\quad + 2\int_0^\infty B_1(s, \zeta)\cosh(\beta_1 x)\cos(\zeta y)d\zeta, \\ \psi^* &= 2\int_0^\infty A_2(s, \xi)\exp(-\gamma_2 y)\sin(\xi x)d\xi \\ &\quad + 2\int_0^\infty B_2(s, \zeta)\sinh(\beta_2 x)\sin(\zeta y)d\zeta, \end{aligned} \tag{10}$$

with

$$\begin{aligned} \gamma_1 &= (\xi^2 + s^2/c_L^2)^{1/2}, \\ \gamma_2 &= (\xi^2 + \kappa^2 s^2/c_L^2)^{1/2}, \end{aligned} \tag{11}$$

$$\begin{aligned} \beta_1 &= (\zeta^2 + s^2/c_L^2)^{1/2}, \\ \beta_2 &= (\zeta^2 + \kappa^2 s^2/c_L^2)^{1/2}. \end{aligned} \tag{12}$$

The Laplace transforms of the boundary conditions (5) and (6) are

$$\tau_{yy}^{0*} = -P/s, \quad \text{for} \quad y = 0, \quad |x| < a, \tag{13a}$$

$$v^{0*} = 0, \quad \text{for} \quad y = 0, \quad a < |x| \leqq h, \tag{13b}$$

$$\tau_{yx}^{0*} = 0, \quad \text{for} \quad y = 0, \quad 0 \leqq |x| \leqq h, \tag{13c}$$

$$\begin{aligned} \tau_{xx}^* &= 0, \\ \tau_{yx}^* &= 0, \end{aligned} \quad \text{for} \quad x = \pm h, \quad |y| < \infty. \tag{14}$$

From equation (13c), we obtain

$$A_2(s, \xi) = \xi\gamma_1 A_1(s, \xi)/\{\xi^2 + \kappa^2 s^2/(2c_L^2)\}. \tag{15}$$

For convenience, we represent $A_1(s, \xi)$ by Fourier transformed displacement $\bar{v}^{0*}$ as follows:

$$A_1(s, \xi) = -\{\xi^2 + \kappa^2 s^2/(2c_L^2)\}\,\bar{v}^{0*}/(\gamma_1\kappa^2 s^2/c_L^2), \tag{16}$$

where $\bar{v}^{0*}$ is defined by

$$\bar{v}^{0*} = 2\int_0^\infty v^{0*}\cos(\xi x)dx. \tag{17}$$

Then, equation (14) can be satisfied if we choose $B_1(s, \zeta)$ and $B_2(s, \zeta)$ to be such that

$$\begin{aligned} B_1(s, \zeta) &= \frac{a_4}{D}\int_0^\infty \bar{v}^{0*}K_1(s)\cos(\xi h)d\xi + \frac{a_2}{D} \\ &\quad \times \int_0^\infty \bar{v}^{0*}K_2(s)\sin(\xi h)d\xi, \\ B_2(s, \zeta) &= -\frac{a_1}{D}\int_0^\infty \bar{v}^{0*}K_1(s)\sin(\xi h)d\xi \\ &\quad - \frac{a_3}{D}\int_0^\infty \bar{v}^{0*}K_2(s)\cos(\xi h)d\xi, \end{aligned} \tag{18}$$

with

$$\begin{aligned} a_1 &= \{\zeta^2 + \kappa^2 s^2/(2c_L^2)\}\cosh(\beta_1 h), \\ a_2 &= -\beta_2\zeta\cosh(\beta_2 h), \end{aligned} \tag{19}$$

$$a_3 = -\beta_1\zeta \sinh(\beta_1 h),$$
$$a_4 = \{\beta_2^2 - \kappa^2 s^2/(2c_L^2)\}, \tag{19}$$
(Cont.)

$$D = a_1 a_4 - a_2 a_3, \tag{20}$$

$$K_1(s) = 2\{\xi^2 + \kappa^2 s^2/(2c_L^2)\}\, k_1(s)/(\pi^2\kappa^2 s^2/c_L^2),$$
$$K_2(s) = -2\{\xi^2 + \kappa^2 s^2/(2c_L^2)\}\, k_2(s)/(\pi^2\kappa^2 s^2/c_L^2),$$
$$k_1(s) = \gamma[\{-\gamma_1^2 + \kappa^2 s^2/(2c_L^2)\}\{\xi^2 + \kappa^2 s^2/(2c_L^2)\}(\gamma_2^2 + \zeta^2) + \xi^2\gamma_2^2(\gamma_1^2 + \zeta^2)]/[(\gamma_1^2 + \zeta^2)(\gamma_2^2 + \zeta^2)\{\xi^2 + \kappa^2 s^2/(2c_L^2)\}],$$
$$k_2(s) = \gamma\xi\zeta(\gamma_2^2 - \gamma_1^2)/\{(\gamma_1^2 + \zeta^2)(\gamma_2^2 + \zeta^2)\}. \tag{21}$$

Here, boundary condition equations (13c) and (14) have been satisfied and remaining equations (13a) and (13b) give dual integral equations,

$$\tau_{yy}^{0*}/(2\mu) = 2\int_0^\infty \bar{v}^{0*}\, [\![[\xi^2\gamma_1\gamma_2 - \{\xi^2 + \kappa^2 s^2/(2c_L^2)\}]/(\pi\gamma_1\kappa^2 s^2/c_L^2) \times \cos(\xi x) + \cos(\xi h)\int_0^\infty f_1(\zeta)K_1(s)d\zeta + \sin(\xi h)\int_0^\infty f_2(\zeta)K_2(s)d\zeta]\!]\, d\xi = -P/(2\mu s),$$
$$\text{for} \quad y = 0, |x| < a, \tag{22a}$$

$$v^{0*} = \frac{1}{\pi}\int_0^\infty \bar{v}^{0*}\cos(\xi x)d\xi = 0, \quad \text{for} \quad y = 0, a < |x| \leqq h, \tag{22b}$$

with

$$f_1(\zeta) = \frac{1}{D}[a_4\{-\beta_1^2 + \kappa^2 s^2/(2c_L^2)\}\cosh(\beta_1 x) - a_3\beta_2\zeta\cosh(\beta_2 x)],$$
$$f_2(\zeta) = \frac{1}{D}[a_2\{-\beta_1^2 + \kappa^2 s^2/(2c_L^2)\}\cosh(\beta_1 x) - a_1\beta_2\zeta\cosh(\beta_2 x)]. \tag{23}$$

To solve integral equation (22), we represent displacement v^{0*} by the following series:

$$2\mu v^{0*} = \sum_{n=1}^\infty c_n(s)\, P_{2n-2}^{(1/2,1/2)}(x/a)(1 - x^2/a^2)^{1/2}, \text{ for } \quad y = 0, |x| < a,$$
$$= 0, \quad \text{for} \quad y = 0, a < |x| \leqq h, \tag{24}$$

where $c_n(s)$ are unknown coefficients to be determined and $P_n^{(1/2,1/2)}(x)$ is a Jacobi polynomial [21]. The Fourier transformation for equation (24) is [21]

$$2\mu\bar{v}^{0*} = \sum_{n=1}^\infty c_n(s)\, 2\sqrt{\pi}\,(-1)^{n-1}\frac{\Gamma(2n - \frac{1}{2})}{(2n-2)!\xi}J_{2n-1}(\xi a), \tag{25}$$

where $\Gamma(x)$ and $J_n(x)$ are the Gamma and Bessel functions, respectively.

Finally, substituting equation (25) into equation (22a), we obtain for $|x| < a$ after integrating with respect to x

$$\sum_{n=1}^\infty c_n(s)\Bigg[\!\Bigg[4\sqrt{\pi}(-1)^n\frac{\Gamma(2n-\frac{1}{2})}{(2n-2)!}\Bigg[\!\Bigg[\frac{c_L^2}{\pi\kappa^2 s^2} \times \int_0^\infty g(\xi)\frac{1}{\xi}J_{2n-1}(\xi a) \times \sin(\xi x)d\xi - \int_0^\infty \frac{1}{\xi}J_{2n-1}(\xi a)\cos(\xi h)\Bigg[\int_0^\infty f_3(\zeta) \times K_1(s)d\zeta\Bigg]d\xi - \int_0^\infty \frac{1}{\xi}J_{2n-1}(\xi a)\sin(\xi h)\Bigg[\int_0^\infty f_4(\zeta) \times K_2(s)d\zeta\Bigg]d\xi\Bigg]\!\Bigg] = -Px/s, \quad \text{for} \quad |x| < a, \tag{26}$$

with

$$g(\xi) = [\{\xi^2 + \kappa^2 s^2/(2c_L^2)\}^2 - \xi^2\gamma_1\gamma_2]/(\xi\gamma_1), \tag{27}$$

$$f_3(\zeta) = a_4\{\kappa^2 s^2/(2c_L^2) - \beta_1^2\}\sinh(\beta_1 x)/(\beta_1 D) - a_3\zeta\sinh(\beta_2 x)/D,$$
$$f_4(\zeta) = a_2\{\kappa^2 s^2/(2c_L^2) - \beta_1^2\}\sinh(\beta_1 x)/(\beta_1 D) - a_1\zeta\sinh(\beta_2 x)/D. \tag{28}$$

The semi-infinite integrals in equation (26) with respect to variable ζ can be easily evaluated numerically because the integrands almost all decrease exponentially. For a large value of ξ, $K_1(s)$ and $K_2(s)$ behave as

$$K_1(s) \to 0(\xi^{-2}),$$
$$K_2(s) \to 0(\xi^{-2}), \tag{29}$$

so that the last two semi-infinite integrals concerned with variable ξ can also be calculated numerically by Filon's method [22]. The first semi-infinite integral in equation (26) is modified as

$$\int_0^\infty g(\xi)\frac{1}{\xi}J_{2n-1}(\xi a)\sin(\xi x)d\xi = \int_0^\infty \{g(\xi) - g(\delta)\}\frac{1}{\xi}J_{2n-1}(\xi a)\sin(\xi x)d\xi + g(\delta)\frac{1}{(2n-1)}\sin\{(2n-1)\sin^{-1}(x/a)\}, \tag{30}$$

where δ represents the infinite limit of ξ, or

$$g(\delta) = \lim_{\xi\to\infty} g(\xi) = (\kappa^2 - 1)s^2/(2c_L^2). \tag{31}$$

The function $\{g(\xi) - g(\delta)\}$ behaves as ξ^{-2} for a large ξ, so that the integral in equation (30) can be evaluated numerically.

Thus equation (26) can be solved for coefficients $c_n(s)$ by the Schmidt method [23]. For brevity, we have rewritten equation (26) as

$$\sum_{n=1}^\infty c_n(s)E_n(s,x) = -u(s,x), \quad \text{for} \quad 0 \leqq x < a, \tag{32}$$

where $E_n(s,x)$ and $u(s,x)$ are known functions and coefficients $c_n(s)$ are unknown and to be determined. A set of functions $P_n(s,x)$ which satisfy the orthogonality condition

$$\int_0^a P_m(s,x)P_n(s,x)dx = N_n\delta_{mn},\ N_n = \int_0^a P_n^2(s,x)dx, \tag{33}$$

can be constructed from the function, $E_n(s,x)$, such that

$$P_n(s,x) = \sum_{i=1}^n \frac{M_{in}}{M_{nn}}E_i(s,x), \tag{34}$$

where M_{in} is the cofactor of the element d_{in} of D_n, which is defined as

$$D_n = \begin{vmatrix} d_{11}\, d_{12} \ldots d_{in} \\ d_{21} \qquad\qquad \vdots \\ \vdots \qquad\qquad \vdots \\ d_{n1} \ldots\ldots d_{nn} \end{vmatrix}, \quad d_{in} = \int_0^a E_i(s,x)E_n(s,x)dx. \tag{35}$$

Using equations (32) and (34), we obtain

$$c_n(s) = \sum_{j=n}^\infty q_j\frac{M_{nj}}{M_{jj}}, \tag{36}$$

with

$$q_j = \frac{-1}{N_j}\int_0^a u(s,x)P_j(s,x)dx. \tag{37}$$

4 Stress-Intensity Factor

Coefficients $c_n(s)$ are known, so that the entire stress field is obtainable. However, in fracture mechanics, it is of importance to determine stress τ_{yy} in the vicinity of the crack's tip. τ_{yy}^* at $y = 0$ is given by

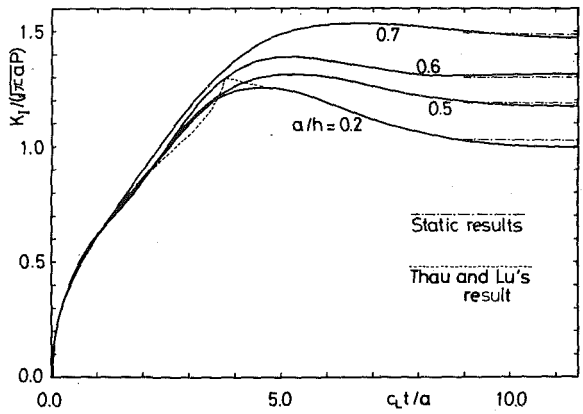


Fig. 2 Dynamic stress-intensity factor K_I

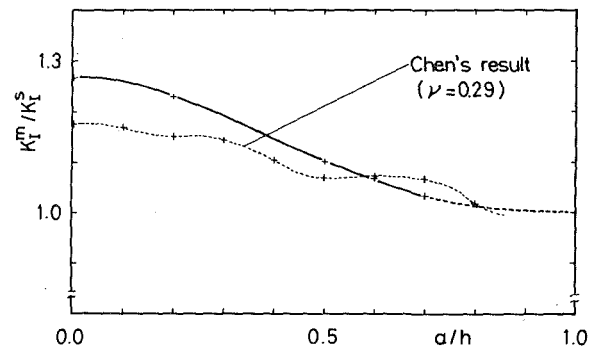


Fig. 3 Ratio K_I^m/K_I^s versus a/h

$$\tau_{yy}^{0*} = \sum_{n=1}^{\infty} c_n(s) \Bigg[\Bigg[4\sqrt{\pi}\,(-1)^n \frac{\Gamma(2n-\frac{1}{2})}{(2n-2)!} \times \int_0^{\infty} \Big[g(\xi)c_L^2/(\pi\kappa^2 s^2)\, J_{2n-1}(\xi a) \times \cos(\xi x) - 1/\xi\, J_{2n-1}(\xi a)\cos(\xi h) \int_0^{\infty} f_1(\zeta)K_1(s)d\zeta - 1/\xi \times J_{2n-1}(\xi a)\sin(\xi h) \int_0^{\infty} f_2(\zeta)K_2(s)d\zeta\Big]d\xi\Bigg]\Bigg]. \quad (38)$$

The singular portion in the stress field results from the relationship,

$$\int_0^{\infty} J_{2n-1}(\xi a)\cos(\xi x)d\xi = (-1)^n a^{2n-1}/\{(x^2-a^2)^{1/2} \times (x+\sqrt{x^2-a^2})^{2n-1}\}, \quad \text{for} \quad x > a > 0. \quad (39)$$

Then, we obtain stress intensity factor K_I^* in the Laplace transform domain as

$$K_I^* = \sqrt{2\pi(x-a)}\,\tau_{yy}^{0*}|_{x\to a+} = \frac{2(\kappa^2-1)}{\sqrt{a}\,\kappa^2}\sum_{n=1}^{\infty} c_n(s)\frac{\Gamma(2n-\frac{1}{2})}{(2n-2)!}. \quad (40)$$

The Laplace inverse transformation in equation (40) is carried out by the numerical method given by Miller and Guy [17]. When the Laplace transform $f^*(s)$ can be evaluated at discrete points given by

$$s = (\beta+1+k)\delta', \quad k = 0, 1, 2, \ldots\ldots\ldots\ldots, \quad (41)$$

we can determine coefficients C_m from the following set of equations:

$$\delta' f^*\{(\beta+1+k)\delta'\} = \sum_{m=0}^{k} k!/\{(k+\beta+1)(k+\beta+2)\ldots\ldots (k+\beta+1+m)(k-m)!\}C_m, \quad (42)$$

where $\delta' > 0$ and $\beta > -1.0$. If coefficients are calculated up to C_{N-1}, an approximate value of $f(t)$ can be found as

$$f(t) = \sum_{m=0}^{N-1} C_m P_m^{(0,\beta)}\{2\exp(-\delta' t) - 1\}, \quad (43)$$

where $P_m^{(\alpha,\beta)}(z)$ is a Jacobi polynomial. Parameters δ', β, and N are selected such that $f(t)$ can best be described within a particular range of time t.

5 Numerical Example and Results

In the plane strain state, the dynamic stress-intensity factor K_I is computed numerically for Poisson's ratio $\nu = 0.25$. The semi-infinite numerical integrations, which occur, are evaluated easily by Filon and Simpson's methods because of the rapid diminution of the integrands. From references [24, 25], it can be seen that the Schmidt method is performed satisfactorily if the first five terms of the infinite series to equation (32) are retained.

To invert the Laplace transforms numerically, we must select the values of β, δ', and N, the number of terms in equation (43). However, there is no best way of selecting these values. From the results given by Sih and his coworkers [6–8], it can be considered that stress-intensity factor K_I may have a form which is similar to

$$f(T) = q_1[1 - \exp(-q_2 T) + q_3\{1 - \cos(q_4 T)\}/T], \quad (44)$$

where T is a dimensionless time variable. The Laplace transform of equation (44) is

$$f^*(s) = q_1[1/s - 1/(s+q_2) + \tfrac{1}{2}q_3 \log\{(s^2+q_4^2)/s^2\}]. \quad (45)$$

Numerical inversions to equation (40) are first carried out with some combinations of β, δ', and N. From these results, the three best curves are selected and they are approximated with equation (44) separately, adjusting the values for q_1, q_2, q_3, q_4. Next, the Laplace inversions of $f^*(s)$ are carried out numerically for the three cases of the constants q_1, q_2, q_3, q_4. Comparing $f(T)$ with the numerical inversion of $f^*(s)$, we can see which combination awards the best approximation. As a result, we know that the numerical Laplace inversions of K_I^* can be carried out satisfactorily for $a/h = 0.6$ by using $\beta = 0$, $\delta' = 0.3$, $N = 7$ and for $a/h = 0.7$, by $\beta = 0$, $\delta' = 0.4$, $N = 7$. It can also be seen that $\beta = 0$, $\delta' = 0.2$, $N = 7$ covers for $a/h = 0.0$–0.5.

In Fig. 2, transient stress-intensity factor K_I is plotted against $c_L t/a$, in which the broken line is Thau and Lu's exact results for $a/h = 0.0$ [12], and the corresponding static values given by Ishida [26] are also shown in it. The curve for $a/h = 0.0$ is omitted because, for the scale shown, the results for $a/h = 0.0$ and $a/h = 0.2$ are indistinguishable. The present calculations cannot detect discontinuity which appears in the exact solution owing to its numerical analysis. However, as a whole, both curves for $a/h = 0.0$ almost agree and further the exact peak value exceeds the other one by only 2.5 percent. Therefore, it is considered that the numerical results of K_I are satisfactory from an engineering point of view. In Fig. 3, the ratio of the peak value of K_I to the corresponding static value, namely, K_I^m/K_I^s is shown graphically, where the slender broken line shows the results given by Chen for $\nu = 0.29$ [18]. Chen insisted that the wavelike curve is attributed to the interaction between the dynamic and finite boundary effects. However, the author considers that this is caused by the crude numerical Laplace inversion.

In general, stress-intensity factor K_I shows a similar tendency which was observed in the investigations by Chen [16, 18]. The peak value of K_I is increased by 23, 19, 7, and 3 percent over its static value for $a/h = 0.2, 0.5, 0.6$, and 0.7, respectively. The dynamic stress-intensity factor approaches the corresponding static value after some time. This means that the K_I^m/K_I^s ratio is not less than unity. Therefore, it may be very possible for the ratio to approach unity for the limit of $a/h \to 1.0$.

Acknowledgments

The author wishes to express thanks to Prof. A. Atsumi, Tohoku University, for his suggestions. A part of this research was carried with a Scientific Research Grant from the Ministry of Education, to which the author is deeply grateful.

References

1 Loeber, J. F., and Sih, G. C., "Diffraction of Antiplane Shear Waves by a Finite Crack," *Journal of the Acoustical Society of America,* Vol. 44, 1968, pp. 90–98.

2 Sih, G. C., and Loeber, J. F., "Wave Propagation in an Elastic Solid With a Line of Discontinuity or Finite Crack," *Quarterly of Applied Mathematics,* Vol. 27, 1969. pp. 193–213.

3 Sih, G. C., and Loeber, J. F., "Torsional Vibration of an Elastic Solids Containing a Penny-Shaped Crack," *Journal of the Acoustical Society of America,* Vol. 44, 1968, pp. 1237–1245.

4 Sih, G. C., and Loeber, J. F., "Normal Compression and Radial Shear Waves Scattering at a Penny-Shaped Crack in an Elastic Solid," *Journal of the Acoustical Society of America,* Vol. 46, 1969, pp. 711–721.

5 Ravera, R. J., and Sih, G. C., "Transient Analysis of Stress Waves Around Cracks Under Antiplane Strain," *Journal of the Acoustical Society of America,* Vol. 47, 1970, pp. 875–881.

6 Sih, G. C., Ravera, R. S., and Embley, G. T., "Impact Response of a Finite Crack in Plane Extension," *International Journal of Solids and Structures,* Vol. 8, 1972, pp. 977–993.

7 Sih, G. C., and Embley, G. T., "Sudden Twisting of a Penny-Shaped Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, 1971, pp. 395–400.

8 Embley, G. T., and Sih, G. C., "Response of a Penny-Shaped crack to impact waves," *Proceedings of the 12th Midwestern Mechanics Conference,* Vol. 6, 1972, pp. 473–487.

9 Mal, A. K., "Interaction of Elastic Waves With a Penny-Shaped Crack," *International Journal of Engineering Science,* Vol. 8, 1970, pp. 381–388.

10 Mal, A. K., "Interaction of Elastic Waves With a Griffith Crack," *International Journal of Engineering Science,* Vol. 8, 1970, pp. 763–776.

11 Thau, S. A., and Lu, T. H, "Diffraction of Transient Horizontal Shear Waves by a Finite Crack and a Finite Rigid Ribbon," *International Journal of Engineering Science,* Vol. 8, 1970, pp. 857–874.

12 Thau, S. A., and Lu, T. H., "Transient Stress-Intensity Factors for a Finite Crack in an Elastic Solid Caused by a Dilatational Wave," *International Journal of Solids and Structures,* Vol. 7, 1971, pp. 731–750.

13 Anderson, J. M., Aberson, J. A., and King. W. W., "Finite Element Analysis of Cracked Structures Subjected to Shock Loads," *Computational Fracture Mechanics,* Rybicki, E. F., ed., ASME, New York, 1975.

14 Chen, Y. M., "Numerical Computation of Dynamic Stress-Intensity Factors by a Lagrangian Finite-Difference Method (The Hemp Code)," *Engineering Fracture Mechanics,* Vol. 7, 1975, pp. 653–660.

15 Glazik, J. L., Jr., "Numerical Analysis of Elastodynamic Near-Tip Stress Fields for Stationary and Propagating Cracks," PhD Dissertation, Northwestern University, Evanston, Ill., 1976.

16 Chen, E. P., "Impact Response of a Finite Crack in a Finite Strip Under Antiplane Shear," *Engineering Fracture Mechanics,* Vol. 9, 1977, pp. 719–724.

17 Miller, M. K., and Guy, W. T., "Numerical Inversion of the Laplace Transform by Use of Jacobi Polynomials," *SIAM Journal of Numerical Analysis,* Vol. 3, 1966, pp. 624–635.

18 Chen, E. P., "Sudden Appearance of a Crack in a Stretched Finite Strip," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 277–280.

19 Kim, K. S., "Dynamic Propagation of a Finite Crack," *International Journal of Solids and Structures,* Vol. 15, 1979, pp. 685–699.

20 Sneddon, I. N., and Srivastav, R. P., "The Stress Field in the Vicinity of a Griffith Crack in a Strip of Finite Width," *International Journal of Engineering Science,* Vol. 9, 1971, pp. 479–488.

21 Erdelyi, A., ed., *Tables of Integral Transforms,* Vol. 1, McGraw-Hill, New York, 1954.

22 Amemiya, A., and Taguchi, T., *Numerical Analysis and Fortran,* Maruzen, Tokyo, 1969.

23 Morse, P. M., and Feshbach, H., *Methods of Theoretical Physics,* Vol. 1, McGraw-Hill, New York, 1958.

24 Itou, S., "Three-Dimensional Wave Propagation in a Cracked Elastic Solid," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 807–811.

25 Itou, S., "Three-Dimensional Problem of a Running Crack," *International Journal of Engineering Science,* Vol. 17, 1979, pp. 59–71.

26 Ishida, M., "Method of Laurent Series Expansion for Internal Crack Problems," *Mechanics of Fracture 1, Methods of Analysis and Solutions of Crack Problems,* Sih, G. C., ed., Noordhoff International Publishing, Leyden, 1973.

H. Hirai
Graduate Student.

M. Satake
Professor.

Department of Civil Engineering,
Faculty of Engineering,
Tohoku University,
Sendai 980, Japan

Stress Analysis of a Penny-Shaped Crack Located Between Two Spherical Cavities in an Infinite Solid

The problem of a penny-shaped crack located between two spherical cavities in an infinite solid subjected to uniaxial loads is considered. Using transformations between harmonic functions in cylindrical coordinates and those in spherical ones, the problem is reduced to nonhomogeneous linear equations. The obtained equations are solved numerically and the influence of the two spherical cavities upon the stress-intensity factors at the penny-shaped crack tip is shown graphically.

Introduction

Recently interest in crack problems has been focused on interaction of cracks and cavities for various technical applications. Many recent papers on the subject are reviewed by Sih [1–2]. However, most of them are results of two-dimensional problems and few studies on three-dimensional problems have been reported. Srivastava and Mahajan [3] investigated the problem of finding stresses in an infinite solid containing a spherical cavity and an external crack. Sternberg, Eubanks, and Sadowsky [4] studied the axisymmetric problem for a solid bounded by two concentric spheres. The problem of stress concentration around a spherical cavity in a semi-infinite body was analyzed by Tsuchida and Nakahara [5], who used the method of the transformations between harmonic functions in cylindrical coordinates and those in spherical ones. The problem of a solid which contains discontinuities such as a penny-shaped crack and some other flaws has not been analyzed as yet.

The objective of this paper is to investigate the stress-intensity factor for a penny-shaped crack located between two spherical cavities in an infinite solid uniformly stretched. Using Hankel transform technique and the transformation of a harmonic function from spherical coordinates into cylindrical ones, the boundary conditions of the penny-shaped crack lead to the set of dual integral equations. Next, the transformation of a harmonic function from cylindrical coordinates into spherical ones is used. Thus the boundary conditions of the two spherical cavities are reduced to the nonhomogeneous linear equations of arbitrary constants.

Formulation of Problem

Consider a penny-shaped crack located between two spherical cavities in an infinite solid subjected to uniaxial loads p in the z-direction (Fig. 1). As the problem is symmetrical for the plane of the penny-shaped crack, let cylindrical coordinates (r, ϕ, z) and spherical ones (R, θ, ϕ) be set as shown in Fig. 2. The boundary conditions for the problem can be specified as follows:

On $z = -c$

$$\tau_{rz} = 0, \quad 0 \leqq r \tag{1}$$

$$\sigma_z = 0, \quad 0 \leqq r \leqq a \tag{2}$$

$$u_z = 0, \quad a < r \tag{3}$$

On $R = b$

$$\sigma_R = 0 \tag{4}$$

$$\tau_{R\theta} = 0 \tag{5}$$

Assuming that there is symmetry about z-axis and the displacement u_ϕ vanishes, we can use the solution of Navier's equation in the form

$$2Gu_r = \frac{\partial \Phi}{\partial r} + z\frac{\partial \Psi}{\partial r}$$

$$2Gu_z = \frac{\partial \Phi}{\partial z} + (4\nu - 3)\Psi + z\frac{\partial \Psi}{\partial z} \tag{6}$$

where the unknown harmonic functions Φ and Ψ are expressed in the following forms, so that uniaxial loads p appear at infinity, i.e.,

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, December, 1979; final revision, April, 1980. Paper No. 80-WA/APM-12.

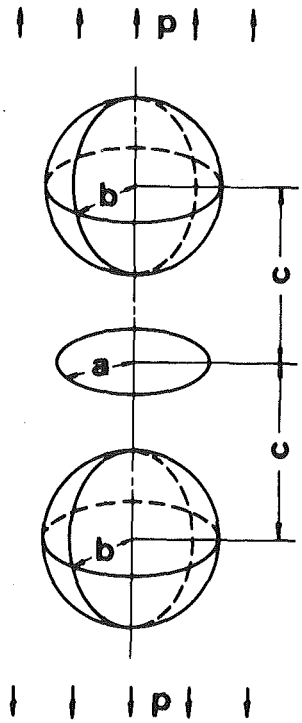


Fig. 1 A penny-shaped crack located between two spherical cavities in an infinite solid subjected to uniaxial loads

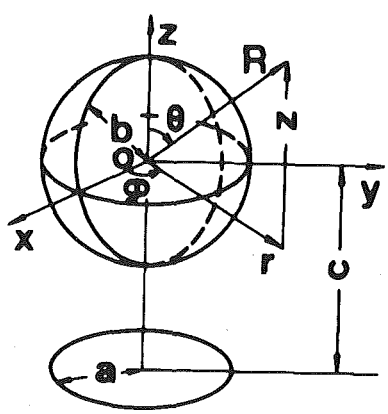


Fig. 2 Coordinates systems

$$\Phi = p\left[\int_0^\infty \phi(\lambda)J_0(\lambda r)e^{-\lambda z}d\lambda + \sum_{m=0}^{\infty} A_m \frac{P_m(\mu)}{R^{m+1}} + \frac{1}{1+\nu}\left\{\nu\left(z^2 - \frac{r^2}{2}\right) + cz\right\}\right] \quad (7)$$

$$\Psi = p\left\{\int_0^\infty \lambda\psi(\lambda)J_0(\lambda r)e^{-\lambda z}d\lambda + \sum_{m=0}^{\infty} B_m \frac{P_m(\mu)}{R^{m+1}} - \frac{z}{2(1+\nu)}\right\} \quad (8)$$

and where $\phi(\lambda)$ and $\psi(\lambda)$ are unknown functions and A_m and B_m are arbitrary constants, which are to be determined from the boundary conditions. In order to transform the harmonic functions in spherical coordinates into those in cylindrical ones, the following relation [6] is useful:

$$\frac{P_n(\mu)}{R^{n+1}} = \frac{(-1)^n}{n!}\int_0^\infty \lambda^n J_0(\lambda r)e^{\lambda z}d\lambda, \quad (z < 0) \quad (9)$$

Rewriting the spherical harmonic functions in equations (7) and (8) by using equation (9), we have the displacement u_z and the stresses σ_z and τ_{rz} as follows:

$$\frac{2G}{p}u_z = \int_0^\infty \left[-\lambda\{\phi(\lambda) + (3 - 4\nu + \lambda z)\psi(\lambda)\}e^{-\lambda z} + \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\lambda^m \times \{\lambda A_m + (4\nu - 3 + \lambda z)B_m\}e^{\lambda z}\right]J_0(\lambda r)d\lambda + \frac{z+c}{1+\nu}$$

$$\frac{\sigma_z}{p} = \int_0^\infty \lambda\left[\lambda\{\phi(\lambda) + (2 - 2\nu + \lambda z)\psi(\lambda)\}e^{-\lambda z} + \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\lambda^m \times \{\lambda A_m - (2 - 2\nu - \lambda z)B_m\}e^{\lambda z}\right]J_0(\lambda r)d\lambda + 1$$

$$\frac{\tau_{rz}}{p} = \int_0^\infty \lambda\left[\lambda\{\phi(\lambda) + (1 - 2\nu + \lambda z)\psi(\lambda)\}e^{-\lambda z} - \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\lambda^m \times \{\lambda A_m - (1 - 2\nu - \lambda z)B_m\}e^{\lambda z}\right]J_1(\lambda r)d\lambda \quad (10)$$

Applying equations (1)–(3) to equation (10), we get the following set of dual integral equations:

$$\int_0^\infty \lambda\left[\lambda\psi(\lambda)e^{\lambda c} + \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\lambda^m\{2\lambda A_m + (4\nu - 3 - 2\lambda c)B_m\}e^{-\lambda c}\right]J_0(\lambda r)d\lambda = -1, \quad 0 \leqq r \leqq a \quad (11)$$

$$\int_0^\infty \left\{\lambda\psi(\lambda)e^{\lambda c} + \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\lambda^m B_m e^{-\lambda c}\right\}J_0(\lambda r)d\lambda = 0, \quad a < r \quad (12)$$

where the unknown function $\phi(\lambda)$ is related to the unknown function $\psi(\lambda)$ and the arbitrary constants A_m and B_m as follows:

$$\phi(\lambda) = (2\nu - 1 + \lambda c)\psi(\lambda) + \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\lambda^{m-1} \times \{\lambda A_m - (1 - 2\nu + \lambda c)B_m\}e^{-2\lambda c} \quad (13)$$

Equation (12) can be satisfied by letting

$$\psi(\lambda) = \frac{e^{-\lambda c}}{\lambda}\int_0^a g(t)\sin(\lambda t)dt - \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\lambda^{m-1}B_m e^{-2\lambda c} \quad (14)$$

where $g(t)$ is an auxiliary function. Substituting equation (14) into equation (11), we obtain

$$g(t) = -\frac{4}{\pi}\sum_{m=0}^{\infty}\frac{(-1)^m}{m!}[\gamma_{m+2}(t)A_m + \{2(\nu - 1)\gamma_{m+1}(t) - c\gamma_{m+2}(t)\}B_m] - \frac{2}{\pi}t \quad (15)$$

where

$$\gamma_m(t) = (m-1)!\sin\left\{m\cdot\arctan\left(\frac{t}{c}\right)\right\}\Big/(c^2 + t^2)^{m/2} \quad (16)$$

Nomenclature

G = shear modulus
ν = Poisson's ratio
$\mu = \cos\theta$
$P_n(\mu)$ = Legendre polynomial
$J_n(\alpha r)$ = Bessel function of the first kind
a = radius of a penny-shaped crack
b = radius of a spherical cavity
c = distance from the plane of a penny-shaped crack to the center of a spherical cavity
$m, n = 0, 1, 2, \ldots$

From equations (13)–(15), we have

$$\phi(\lambda) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!}\Bigg[\Bigg\{\frac{4e^{-\lambda c}}{\pi\lambda}(1-2\nu-\lambda c)\omega_{m+2}(\lambda) + \lambda^m e^{-2\lambda c}\Bigg\}A_m + \Bigg[\frac{4e^{-\lambda c}}{\pi\lambda}(1-2\nu-\lambda c)\{2(\nu-1)\omega_{m+1}(\lambda) - c\omega_{m+2}(\lambda)\} - 2c\lambda^m e^{-2\lambda c}\Bigg]B_m\Bigg] + (1-2\nu-\lambda c)\frac{2e^{-\lambda c}}{\pi\lambda}\int_0^a t\sin(\lambda t)dt \quad (17)$$

$$\psi(\lambda) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!}\Bigg[-\frac{4e^{-\lambda c}}{\pi\lambda}\omega_{m+2}(\lambda)A_m - \Bigg[\frac{4e^{-\lambda c}}{\pi\lambda}\{2(\nu-1)\omega_{m+1}(\lambda) - c\omega_{m+2}(\lambda)\} + \lambda^{m-1}e^{-2\lambda c}\Bigg]B_m\Bigg] - \frac{2e^{-\lambda c}}{\pi\lambda}\int_0^a t\sin(\lambda t)dt \quad (18)$$

where

$$\omega_m(\lambda) = \int_0^a \gamma_m(t)\sin(\lambda t)dt \quad (19)$$

In order to transform the harmonic functions in cylindrical coordinates into those in spherical ones, the following identity [7] is useful:

$$J_0(\lambda r)\,e^{-\lambda z} = \sum_{n=0}^{\infty}\frac{(-\lambda R)^n}{n!}P_n(\mu) \quad (20)$$

Using equation (20), we can write the cylindrical harmonic functions in equations (7) and (8) as

$$\int_0^\infty \phi(\lambda)\,J_0(\lambda r)e^{-\lambda z}d\lambda = \sum_{n=0}^{\infty}\alpha_n R^n P_n(\mu)$$

$$\int_0^\infty \lambda\psi(\lambda)\,J_0(\lambda r)\,e^{-\lambda z}d\lambda = \sum_{n=0}^{\infty}\beta_n R^n P_n(\mu) \quad (21)$$

where

$$\alpha_n = \frac{(-1)^n}{n!}\int_0^\infty \phi(\lambda)\,\lambda^n d\lambda$$

$$\beta_n = \frac{(-1)^n}{n!}\int_0^\infty \psi(\lambda)\,\lambda^{n+1}d\lambda \quad (22)$$

Substituting equations (17) and (18) into equation (22), we obtain

$$\alpha_n = \sum_{m=0}^{\infty}(C_n{}^{m+1}A_m + D_n{}^{m+1}B_m) + E_n$$

$$\beta_n = \sum_{m=0}^{\infty}(F_n{}^{m+1}A_m + G_n{}^{m+1}B_m) + H_n \quad (23)$$

where

$$C_n{}^{m+1} = \frac{(-1)^{m+n}}{m!\,n!}\left[\frac{4}{\pi}\{(1-2\nu)\xi_n{}^{m+2} - c\xi_{n+1}{}^{m+2}\} + \frac{(m+n)!}{(2c)^{m+n+1}}\right]$$

$$D_n{}^{m+1} = \frac{(-1)^{m+n}}{m!\,n!}\left[\frac{4}{\pi}(1-2\nu)\{2(\nu-1)\xi_n{}^{m+1} - c\xi_n{}^{m+2}\} - \frac{4}{\pi}c\{2(\nu-1)\xi_{n+1}{}^{m+1} - c\xi_{n+1}{}^{m+2}\} - \frac{(m+n)!}{(2c)^{m+n}}\right]$$

$$E_n = \frac{2(-1)^n}{\pi\,n!}\{(1-2\nu)\zeta_n - c\zeta_{n+1}\}$$

$$F_n{}^{m+1} = \frac{4(-1)^{m+n+1}}{\pi\,m!\,n!}\xi_{n+1}{}^{m+2}$$

$$G_n{}^{m+1} = \frac{(-1)^{m+n+1}}{m!\,n!}\left[\frac{4}{\pi}\{2(\nu-1)\xi_{n+1}{}^{m+1} - c\xi_{n+1}{}^{m+2}\} + \frac{(m+n)!}{(2c)^{m+n+1}}\right]$$

$$H_n = \frac{2(-1)^{n+1}}{\pi\,n!}\zeta_{n+1} \quad (24)$$

and where

$$\xi_n{}^m = \int_0^a \gamma_m(t)\gamma_n(t)dt$$

$$\zeta_n = \int_0^a t\,\gamma_n(t)dt \quad (25)$$

Making use of equations (6)–(8) and (21), the stress components σ_R and $\tau_{R\theta}$ in spherical coordinates can be expressed as

$$\frac{\sigma_R}{p} = \sum_{n=0}^{\infty}\Bigg\{\frac{(n+1)(n+2)}{R^{n+3}}A_n + \frac{n(n^2+3n-2\nu)}{(2n-1)R^{n+1}}B_{n-1} + \frac{(n+1)(n+2)(n+5-4\nu)}{(2n+3)R^{n+3}}B_{n+1} + (n-1)nR^{n-2}\alpha_n + \frac{(n-1)n(n-4+4\nu)}{2n-1}R^{n-2}\beta_{n-1} + \frac{(n+1)(n^2-n-2-2\nu)}{2n+3}R^n\beta_{n+1}\Bigg\}P_n(\mu) + \frac{1}{3}\{P_0(\mu) + 2P_2(\mu)\} \quad (26)$$

$$\frac{\tau_{R\theta}}{p\sin\theta} = \sum_{n=1}^{\infty}\Bigg\{\frac{n+2}{R^{n+3}}A_n + \frac{n^2-2+2\nu}{(2n-1)R^{n+1}}B_{n-1} + \frac{(n+2)(n+5-4\nu)}{(2n+3)R^{n+3}}B_{n+1} + (1-n)R^{n-2}\alpha_n + \frac{(n-1)(4-n-4\nu)}{2n-1}R^{n-2}\beta_{n-1} + \frac{(1-2n-n^2-2\nu)}{2n+3}R^n\beta_{n+1}\Bigg\}\frac{\partial}{\partial\mu}P_n(\mu) - \frac{1}{3}\frac{\partial}{\partial\mu}P_2(\mu) \quad (27)$$

In deriving equations (26) and (27), the recurrence formulas of Legendre polynomials are used. Applying equations (4) and (5) to equations (26) and (27), respectively, we get the following equations:

For $n = 0$

$$\frac{2}{b^3}A_0 + \frac{2(5-4\nu)}{3b^3}B_1 + \frac{2}{3}(\nu-1)\beta_1 = -\frac{1}{3}$$

For $n = 1$

$$\frac{6}{b^4}A_1 + \frac{2(2-\nu)}{b^2}B_0 + \frac{12(3-2\nu)}{5b^4}B_2 - \frac{4}{5}(1+\nu)b\beta_2 = 0$$

For $n = 2$

$$\frac{12}{b^5}A_2 + \frac{4}{3b^3}(5-\nu)B_1 + \frac{12}{7b^5}(7-4\nu)B_3 + 2\alpha_2 + \frac{4}{3}(2\nu-1)\beta_1 - \frac{6}{7}\nu b^2\beta_3 = -\frac{2}{3}$$

For $n \geqq 3$

$$\frac{(n+1)(n+2)}{b^{n+3}}A_n + \frac{n(n^2+3n-2\nu)}{(2n-1)b^{n+1}}B_{n-1} + \frac{(n+1)(n+2)(n+5-4\nu)}{(2n+3)b^{n+3}}B_{n+1} + (n-1)nb^{n-2}\alpha_n + \frac{n(n-1)(n-4+4\nu)}{2n-1}b^{n-2}\beta_{n-1} + \frac{(n+1)(n^2-n-2-2\nu)}{2n+3}b^n\beta_{n+1} = 0 \quad (28)$$

For $n = 1$

$$\frac{3}{b^4}A_1 + \frac{2\nu-1}{b^2}B_0 + \frac{6(3-2\nu)}{5b^4}B_2 - \frac{2(1+\nu)}{5}b\beta_2 = 0$$

For $n = 2$

$$\frac{4}{b^5}A_2 + \frac{2(1+\nu)}{3b^3}B_1 + \frac{4(7-4\nu)}{7b^5}B_3 - \alpha_2 \quad (29)$$

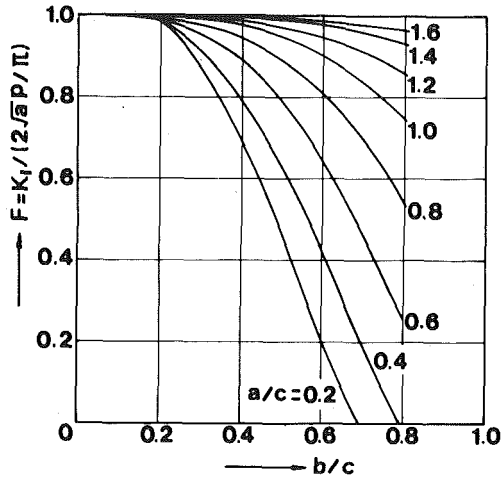


Fig. 3 Normalized stress-intensity factors for a penny-shaped crack located between two spherical cavities in an infinite solid subjected to uniaxial loads p

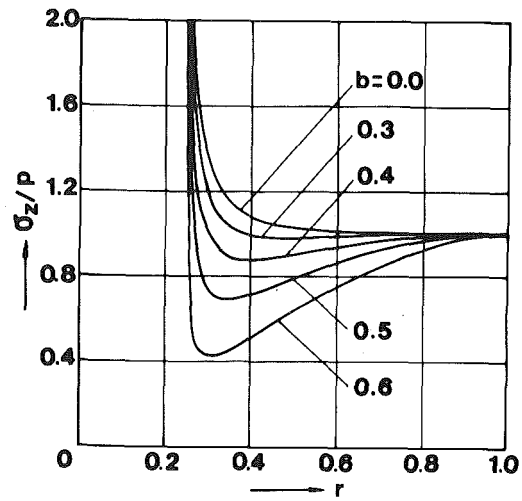


Fig. 4 Normal stress in the neighborhood of a penny-shaped crack (a = 0.25, c = 1.0)

$$+\frac{2(1-2\nu)}{3}\beta_1-\frac{(7+2\nu)}{7}b^2\beta_3=\frac{1}{3}$$

For $n \geqq 3$

$$\frac{n+2}{b^{n+3}}A_n+\frac{n^2-2+2\nu}{(2n-1)b^{n+1}}B_{n-1}+\frac{(n+2)(n+5-4\nu)}{(2n+3)b^{n+3}}B_{n+1}$$
$$+(1-n)b^{n-2}\alpha_n+\frac{(n-1)(4-n-4\nu)}{2n-1}b^{n-2}\beta_{n-1}$$
$$+\frac{1-2n-n^2-2\nu}{2n+3}b^n\beta_{n+1}=0 \quad (29)$$

(Cont.)

Substituting equation (23) into equations (28) and (29), we get the nonhomogeneous linear equations of the arbitrary constants A_n and B_n. The unknown functions $\phi(\lambda)$ and $\psi(\lambda)$ are given by equations (17) and (18) in terms of A_n and B_n just obtained.

The stress-intensity factor, K_I, is defined in terms of the stress on the plane $z = -c$ as

$$K_I=\lim_{r\to a^+}\{2(r-a)\}^{1/2}\sigma_z(r,\phi,-c)$$
$$=2p\frac{a^{1/2}}{\pi}+\frac{4p}{\pi a^{1/2}}\sum_{m=0}^{\infty}\frac{(-1)^m}{m!}[\gamma_{m+2}(a)A_m+\{2(\nu-1)\gamma_{m+1}(a)$$
$$-c\gamma_{m+2}(a)\}B_m] \quad (30)$$

The first term in equation (30) represents the stress-intensity factor in the case where the two spherical cavities are absent. The second term represents the influence of the two spherical cavities on the penny-shaped crack.

Numerical Results and Discussion

Computations were carried out for Poisson's ratio of 0.3. Fig. 3 shows the normalized stress-intensity factors plotted with b/c for a/c. As b/c increases, all the curves decrease in magnitude very sharply at the beginning and this tendency is more noticeable for the lower value of a/c. The two-dimensional problem of a crack located between two circular holes in an infinite plate subjected to uniaxial loads was discussed by Newman [8]. In the two-dimensional problem, we assume that a, b, and c stand for the half-crack length, the radius of a hole, and the distance from the plane of a crack to the center of a hole, respectively. Comparing the results shown in Fig. 3 in this paper with those contained in the paper of Newman, we note that, as b/c increases, the normalized stress-intensity factors in two-dimensional case decrease more rapidly than those in three-dimensional case and this tendency is more conspicuous for the lower value of a/c. Figs. 4 and 5 show the stress component σ_z and the displacement u_z on the plane $z = -c$, respectively. As b increases, the normal stress σ_z near the crack tip and the displacement u_z on the crack surface decrease. The foregoing curves are drawn on the results which have 3 digits of efficient numbers by taking 10 terms of the arbitrary constants A_n and B_n, respectively.

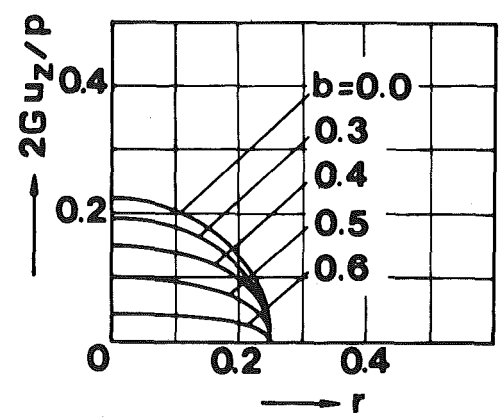


Fig. 5 Displacement in the z-direction on the surface of a penny-shaped crack (a = 0.25, c = 1.0)

Concluding Remarks

In the present paper, the interaction of a penny-shaped crack and two spherical cavities in an infinite body uniformly stretched was investigated. It was found that the magnitude of the stress-intensity factor K_I decreases rapidly as the radius of the penny-shaped crack becomes smaller than those of the spherical cavities and in addition as the spherical surfaces approach the crack surface.

As for the problem of a penny-shaped crack and two spherical cavities in an infinite body, the technique employed here will also be applicable to determine the stress intensity factors K_{II} for shear loads and K_{III} for torsional loads.

Acknowledgment

The authors wish to thank Dr. Y. Shindo, Tohoku University, for his valuable suggestions and careful review of the manuscript.

References

1 Sih, G. C., ed., "Method of Analysis and Solution of Crack Problems," *Mechanics of Fracture*, Vol. 1, Nordhoff, 1973.

2 Sih, G. C., ed., "Three-Dimensional Crack Problems," *Mechanics of Fracture*, Vol. 2, Nordhoff, 1975.

3 Srivastava, K. N., and Mahajan, R. L., "Distribution of Stress in an Infinite Solid Containing a Spherical Cavity and an External Crack," *Indian Journal of Technology,* Vol. 8, 1970, p. 398.

4 Sternberg, E., Eubanks, R. A., and Sadowsky, M. A., "On the Axisymmetric Problem of Elasticity Theory for a Region Bounded by Two Concentric Spheres," *Proceedings of the First U.S. National Congress of Applied Mechanics,* 1951, pp. 209–215.

5 Tsuchida, E., and Nakahara, I., "Three-Dimensional Stress Concentration Around a Spherical Cavity in a Semi-Infinite Elastic Body," *Bulletin of the Japan Society of Mechanical Engineers,* Vol. 13, No. 58, 1970, pp. 499–508.

6 Sneddon, I. N., *Fourier Transform,* McGraw-Hill, New York, 1951, p. 514.

7 Morse, P. M., and Feshbach, H., *Method of Theoretical Physics,* Part II, McGraw-Hill, New York, 1953, p. 1318.

8 Newman, J. C., "An Improved Method of Collocation for the Stress Analysis of Cracked Plates With Various Shaped Boundaries," NASA Technical Note TN D-6376, 1971, pp. 1–45.

K. K. Lo

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

Fatigue Crack Closure Following a Step-Increase Load

A Dugdale-Barenblatt model is used to examine the effects of crack closure following a step increase in the applied cyclic loading. Complex function formulation is employed to calculate opening and contact loads. It is shown that the effect of previous history of loading on crack growth is significant only when the extent of crack growth is within about one plastic zone size.

Introduction

Experimentally, it has been observed that a sudden increase in the level of loading for a crack under constant amplitude cyclic stress has a large influence on the crack-growth rate (see, for example [1–2]). This behavior of crack growth acceleration and retardation has been explained on the basis of fatigue crack closure first proposed by Elber [3] and subsequently made use of by various authors [4–10].

According to Elber's concept, the residual plastic stretch left in the wake of a steadily advancing crack interacts with the plastic zone ahead of the crack tip and causes closure above zero load.

Earlier, finite-element calculations have been carried out by Newman [4], Ohji, et al. [5, 6], on crack closure. More recently, there have been combined numerical-analytic studies [7–10] and analytical studies [11, 12] on this subject. McCartney [11], using the Dugdale model, studied crack closure for a finite-length crack in an infinite medium. In [12], Budiansky and Hutchinson examined the effects of crack closure for a steadily growing semi-infinite crack under *constant* amplitude cyclic loading. Their study [12] is different from some of the previous studies [4, 8] in that no explicit reference is given to the actual crack growth process

In this paper, we study the problem of crack closure following a *step increase* in loading in the same spirit as [12]. A modified Dugdale-Barenblatt model under the assumptions of small-scale yielding is used. Numerical results are given for the opening and closing loads and the influence of the prior loading history is assessed.

Formulation and Analysis

We base our theoretical study on the two-dimensional Dugdale-Barenblatt model. According to this model, the plastic yielding region is confined to a narrow slitlike region directly ahead of the crack tip where the material is yielding at some maximum constant tensile stress σ_Y. We assume *small-scale yielding* conditions under which the *far-field behavior* is governed by the asymptotic crack-tip *elastic* stress field,

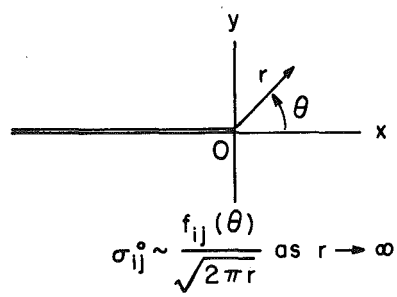


Fig. 1 Crack-tip geometry

$$\sigma_{ij} \sim K f_{ij}(\theta)/\sqrt{2\pi r} \quad \text{as} \quad r \to \infty \tag{1}$$

where K, the elastic intensity factor, is considered a *prescribed load parameter.*

For a stationary crack, the results of the Dugdale model are summarized as follows [12]:

$$w = \frac{\pi}{8}\left(\frac{K}{\sigma_Y}\right)^2, \; \delta_o = \frac{K^2}{E\sigma_Y} = \frac{8\sigma_Y w}{\pi E},$$

$$\delta/\delta_o = g(x/w), \; g(\xi) \equiv \sqrt{1-\xi} - \frac{\xi}{2}\ln\left|\frac{1+\sqrt{1-\xi}}{1-\sqrt{1-\xi}}\right| \tag{2}$$

where w is the plastic zone size, K is the Mode-I elastic stress-intensity factor, δ_o is the crack opening displacement and δ is the plastic stretch in the interval $(0, w)$, and E is Young's modulus.

This model was modified in [12] to account for the residual plastic strain that is left behind as the crack advances through the material. This introduces an incompatibility which gives rise to the effects of crack closure. Under steady cyclic loading between $K_{min} = 0$ and $K = K_{max}{}^L$, the crack approaches a steady-state situation as shown in Fig. 2. This is the picture analyzed in [12] and from the associated

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, May, 1980.

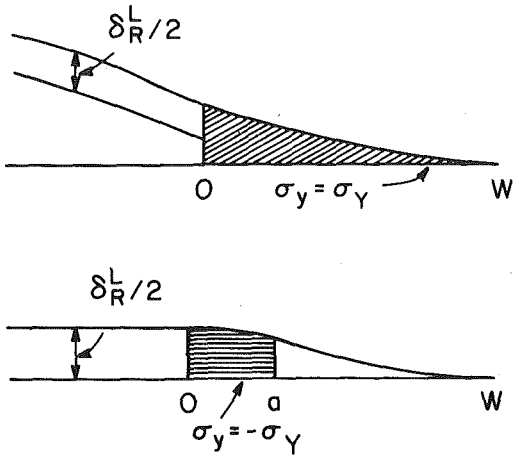


Fig. 2 Growing crack with residual plastic deformation and plastic zone at (*a*) maximum and (*b*) minimum load

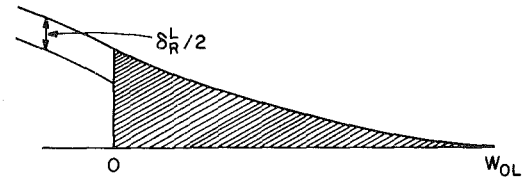


Fig. 3(*a*) Growing crack at overload $K = K_{\max}{}^U$

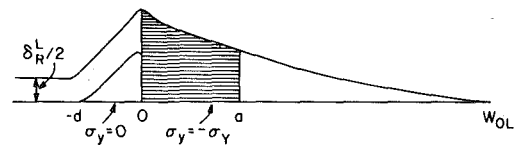


Fig. 3(*b*) Growing crack at minimum load $K_{\min} = 0$ immediately after $K_{\max}{}^U$

boundary-value problem at $K_{\min} = 0$, the residual stretch was found to be $\delta_R{}^L = 0.8562\ \delta_o{}^L$ (see equation (26) of [12]), where $\delta_o{}^L$ is the crack-tip opening at load $K_{\max}{}^L$. At this point we consider the situation where the load is raised to $K_{\max}{}^U$ ($> K_{\max}{}^L$). The plastic stretch just created ahead of the crack tip by the overload $K_{\max}{}^U$ is greater than the existing stretch $\delta_R{}^L$ caused by $K_{\max}{}^L$. Consequently, due to the incompatibility of the plastic stretches caused by two different load levels, there is a region (on the crack face) in which the crack is "propped" open. The picture now at $K_{\min} = 0$ is given by Fig. 3(*b*). Further crack growth will reduce the strain incompatibility until the plastic stretch becomes completely compatible with that due to a constant cyclic load between $K = 0$ and $K = K_{\max}{}^U$ (see Fig. 2(*b*).

Since the load $K_{\max}{}^U$ is greater than $K_{\max}{}^L$, the plastic zone size is increased to

$$w_{OL} = \frac{\pi}{8}(K_{\max}{}^U/\sigma_Y)^2$$

(see equation (1)). Reversed plastic yielding occurs directly ahead of the crack tip as soon as unloading takes place. The crack opening displacement at load K is now given by

$$\delta = \delta_o{}^U\,[g(x/w_{OL}) - 2(a_k/w_{OL})\,g(x/a_k)]$$

where

$$a_k = \frac{\pi}{8}[(K_{\max}{}^U - K)^2/2\sigma_Y]^2 = \frac{w_{OL}}{4}(1 - K/K_{\max}{}^U)^2. \qquad (3)$$

On lowering the load to $K = 0$ corresponding to the picture in Fig. 3(*b*), contact first occurs when the crack opening displacement δ in (3) is equal to the residual stretch $\delta_R{}^L$ created previously during the cyclic load between $K = 0$ and $K_{\max}{}^L$, i.e.,

$$\left.\begin{aligned} &g(x/w_{OL}) - 2(a_k/w_{OL})\,g(x/a_k) = \delta_R{}^L/\delta_o{}^U \\ &\text{and} \\ &g'(x/w_{OL}) - 2g'(x/a_k) = 0 \end{aligned}\right\} \qquad (4)$$

It turns out that with the function $g(x)$ given by (2), equations (4) can be solved exactly to give the two roots $K = K_{\text{cont}}$, $x = x_c$,

$$\begin{aligned} K_{\text{cont}}/K_{\max}{}^U &= 1 - \sqrt{1 - (\delta_R{}^L/\delta_o{}^U)^2} \\ &= 1 - \sqrt{1 - (\delta_R{}^L/\delta_o{}^L)^2(\delta_o{}^L/\delta_o{}^U)^2} \end{aligned} \qquad (5a)$$

and the point of contact at x_c,

$$\xi_c \equiv x_c/w_{OL} = -[1 - (\delta_R{}^L/\delta_o{}^U)^2]^2/4(\delta_R{}^L/\delta_o{}^U)^2 \qquad (5b)$$

As mentioned earlier, $\delta_R{}^L/\delta_o{}^L$ was calculated in [12] to be 0.8562 and hence, with the use of (2), this gives

$$K_{\text{cont}}/K_{\max}{}^U = 1 - \sqrt{1 - 0.733\,(K_{\max}{}^L/K_{\max}{}^U)^4} \qquad (6)$$

It is noted that equation (6) in general predicts a very low contact load.

The problem corresponding to Fig. 3(*b*) is illustrated in Fig. 4. An elastic field is sought with the following boundary conditions given on the real axis: $\delta = \delta_R{}^L$ in $(-\infty, -d)$, $\sigma_y = 0$ in $(-d, 0)$, and $\sigma_y = -\sigma_Y$ in $(0, a)$, the yield stress in compression, and $\delta = \delta_0{}^U g(x/w_{OL})$ in (a, w_{OL})·($\delta_o{}^U = 8\sigma_Y w_{OL}/\pi E$ from (1).) By the Barenblatt hypothesis, we require the stresses to be bounded. Because $K_{\min} = 0$, the condition at infinity must correspond to that of a dislocation $\delta_R{}^L$ along the negative x-axis. Finally, it has to be checked *a posteriori* that the solution satsifies $|\sigma_y| \leqslant \sigma_Y$ on the x-axis and $\sigma_y \leqslant 0$ in the contact region.

The method of analysis involves using Muskhelishvili complex potentials ϕ, ψ [13] in two-dimensional linear elasticity. Since similar analysis has been carried out in considerable detail in [12], the theoretical formulation in this study will only be sketched. For completeness, the Appendix summarizes all the results needed for the present analysis.

The problem for $K = 0$ immediately after the change in loading ($K_{\max}{}^L$ to $K_{\max}{}^U$), illustrated in Fig. 3(*b*), can be reduced to the following problem for the potential $\Phi \equiv \phi'$ which, except along a branch cut on the x-axis, is analytic everywhere. The boundary conditions given on the x-axis are

$$\left.\begin{aligned} \Phi_+ - \Phi_- &= 0 && \text{for } x < -d \\ \Phi_+ + \Phi_- &= 0 && -d < x < 0 \\ \Phi_+ + \Phi_- &= -\sigma_Y && 0 < x < a \\ \Phi_+ - \Phi_- &= iEd\delta_M/dx && a < x < w_{OL} \end{aligned}\right\} \qquad (7)$$

where $\delta_M(x) = \delta_o{}^U g(x/w_{OL})$. In order that the solution at infinity corresponds to that of a dislocation $\delta_R{}^L$ along the negative x-axis, the function $\Phi \equiv \phi'(z)$ must behave like $E\delta_R{}^L/8\pi z$ as $|z| \to \infty$. The solution of the aforementioned boundary-value problem, with the use of Plemelj integrals [14] and the auxillary function $\lambda(z) \equiv \sqrt{(z+d)(z-a)}$, is then found to be given by

$$\begin{aligned} \lambda(z)\Phi(z) = \frac{E\delta_R{}^L}{8\pi} - \frac{\sigma_Y}{2\pi}\int_o^a \frac{\sqrt{(x+d)(a-x)}\,dx}{x-z} \\ + \frac{E}{8\pi}\int_a^{w_{OL}} \frac{\sqrt{(x+d)(x-a)}}{x-z}\,\frac{d\delta_M}{dx}\,dx \end{aligned} \qquad (8)$$

where the branch cut is along $(-d, a)$ ($\lambda(z) \sim z$ as $|z| \to \infty$).

Introducing the following dimensionless variables:

$\zeta = z/w_{OL}$, $\xi = x/w_{OL}$, $\alpha = a/w_{OL}$, $\gamma = d/w_{OL}$, $F = (\pi^2/\sigma_Y)\,\Phi$ we can express the result (8) as

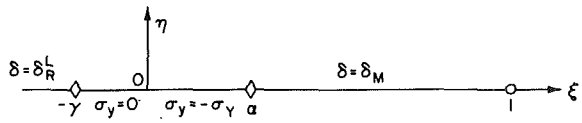


Fig. 4 Boundary conditions along crack line corresponding to deformations shown in Fig. 3

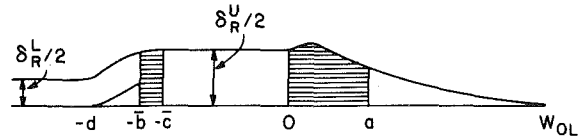


Fig. 5 Growing crack at ($K = 0$) minimum load ($b > 0$)

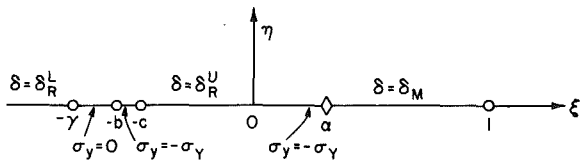


Fig. 6 Boundary conditions along crack line corresponding to deformations shown in Fig. 5

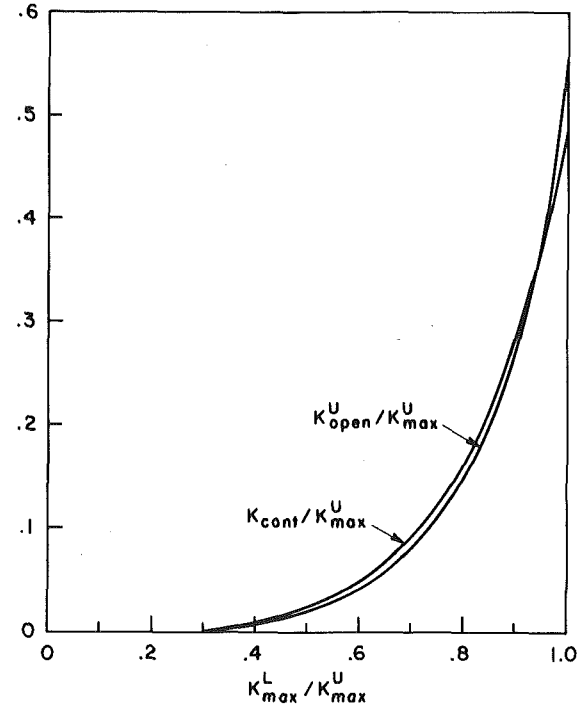


Fig. 7 Opening and contact loads immediately after $K_{max}{}^U$ ($b = 0$)

$$\chi(\zeta)F(\zeta) = s - \frac{\pi}{2}\int_0^\alpha \frac{\sqrt{(\xi+\gamma)(\alpha-\xi)}}{\xi-\zeta}\,d\xi - \int_\alpha^1 \frac{\sqrt{(\xi+\gamma)(\xi-\alpha)}}{\xi-\zeta} f_1(\xi)d\xi \quad (9)$$

where

$$s = \delta_R{}^L/\delta_o{}^U = (\delta_R{}^L/\delta_o{}^L)(\delta_o{}^L/\delta_o{}^U) = 0.8562\,(K_{max}{}^L/K_{max}{}^U)^2,$$

$$f_1(t) \equiv -\frac{d}{dt}(\delta_M/\delta_o{}^U), \quad \chi(\zeta) \equiv \sqrt{(\zeta+\gamma)(\zeta-\alpha)}\,.$$

Requiring that the stresses be bounded at $\zeta = \alpha, -\gamma$ gives two conditions for the two unknowns α, γ: $F(\alpha) = F(-\gamma) = 0$. This gives two equations for the determination of α and γ,

$$s + \sqrt{\alpha\gamma} + (\alpha+\gamma)\sin^{-1}\sqrt{\frac{\alpha}{\alpha+\gamma}} - \int_\alpha^1 f_1(t)\sqrt{\frac{t+\gamma}{t-\alpha}}\,dt = 0$$

$$s + \sqrt{\alpha\gamma} - (\alpha+\gamma)\sin^{-1}\sqrt{\frac{\alpha}{\alpha+\gamma}} - \int_\alpha^1 f_1(t)\sqrt{\frac{t-\alpha}{t+\gamma}}\,dt = 0 \quad (10)$$

These equations were solved numerically and the results are given in Fig. 8 for different values of $K_{max}{}^L/K_{max}{}^U$. Having determined α, γ for a given $K_{max}{}^L/K_{max}{}^U$, we can calculate the displacement gradient δ' in $(0, \alpha)$ from (18), (9), and (10). It is given by

$$\frac{d}{d\xi}(\delta/\delta_o{}^U) \equiv -f_2(\xi) = -\frac{1}{\pi}\sqrt{(\xi+\gamma)(\alpha-\xi)}\left\{\int_\alpha^1 \frac{f_1(\tau)d\tau}{(\tau-\xi)\sqrt{(\tau+\gamma)(\tau-\alpha)}} - \frac{\pi}{2}\mathcal{J}_0^\alpha \frac{d\tau}{(\tau-\xi)\sqrt{(\tau+\gamma)(\alpha-\tau)}}\right\} \quad (11)$$

where the second integral is to be interpreted in the Cauchy principal sense. We note that this integral gives a logarithmic singularity at the crack tip $\xi = 0$. The stresses were calculated numerically according to the formulas in the Appendix and they were found not to exceed yield in the intervals $(-\infty, -\gamma)$ and $(\alpha, 1)$.

During the reloading process from $K = 0$ to $K = K_{max}{}^U$, the picture for $K > K_{open}$ is identical to that analyzed in [12]. The opening load K_{open} ((equation (42) of [12]) is given by (see (20) in the Appendix)

$$K_{open}/K_{max}{}^U = \frac{1}{\pi}\int_0^\alpha \frac{f_2(\xi)d\xi}{\sqrt{\xi}} + \frac{1}{\pi}\int_\alpha^1 \frac{f_1(\xi)d\xi}{\sqrt{\xi}} \quad (12)$$

where $f_2(\xi)$ is calculated numerically from (11). The results are summarized in Fig. 9 and, consistent with the results for K_{cont}, (equation (6)) K_{open} is drastically reduced. For example, $K_{open}/K_{max}{}^U = 0.078$ for $K_{max}{}^L/K_{max}{}^U = 0.7$ compared to $K_{open}/K_{max} = 0.557$ in [12].

Consider now the situation in which the load is allowed to remain between $K = 0$ and $K_{max}{}^U$. The parameter $\bar{b}$ in in Fig. 5 denotes the extent of the crack growth once the load has been raised to $K_{max}{}^U$. Relative to the position of the current crack tip, which is fixed at 0, the point $-\bar{b}$ moves from right to left as the crack advances in the position x-direction. As in [12], the actual kinetics of crack growth is not considered here; rather, our main objective is to analyze crack closure as it affects K_{open}, K_{cont} and the crack-tip deformation as the crack advances. As $\bar{b} \to \infty$, the steady-state situation in Fig. 2 is recovered with the maximum load now given by $K_{max}{}^U$. The residual plastic stretch in this limit is given by $\delta_R{}^U = 0.8562\,\delta_o{}^U$ where $\delta_o{}^U$ is the crack opening displacement created by $K_{max}{}^U$ and equal to $(\pi/8)\,(K_{max}{}^U/\sigma_Y)^2$.

As the crack advances ($\bar{b}$ increases), the crack-tip field is given in Fig. 5 and the associated boundary-value problem is illustrated in Fig. 6. In terms of the dimensionless variables, we want F to satisfy on the x-axis:

$$\left.\begin{array}{lll} F_+ - F_- = 0 & \text{for} & \xi < -\gamma \\ F_+ + F_- = 0 & & -\gamma < \xi < -c \\ F_+ + F_- = -\pi^2 & & -c < \xi < -b \\ F_+ - F_- = 0 & & -b < \xi < 0 \\ F_+ + F_- = -\pi^2 & & 0 < \xi < \alpha \\ F_+ - F_- = 2\pi i \dfrac{d}{d\xi}(\delta_M/\delta_o{}^U) & & \alpha < \xi < 1 \end{array}\right\} \quad (13)$$

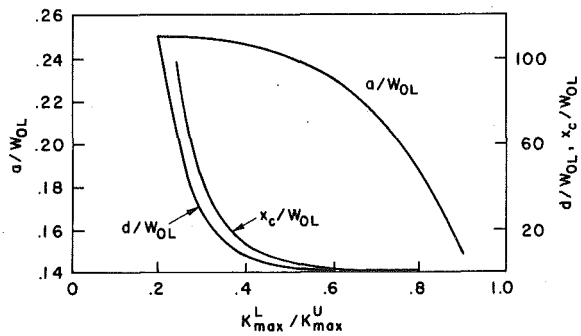


Fig. 8 Contact points at $K = 0$, K_{cont}, immediately following overload ($b = 0$)

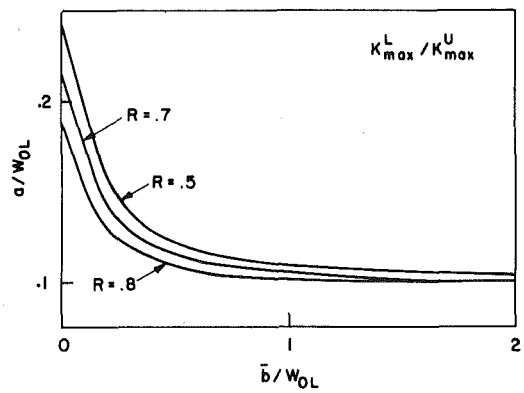


Fig. 9 Reversed yielding zone versus crack-tip position

where $c = \bar{c}/w_{OL}$, $b = \bar{b}/w_{OL}$. We now have to introduce the auxillary function $\chi(\zeta) \equiv \sqrt{\zeta(\zeta-\alpha)(\zeta+c)(\zeta+\gamma)}$ and the branch cuts are along $(-\gamma, -c)$ and $(0, \alpha)$. ($\chi(\zeta) \sim \zeta^2$ as $|\zeta| \to \infty$). The far-field behavior is, as before, given by $F(\zeta) \sim (\delta_R{}^L/\delta_o{}^U)\zeta^{-1}$ as $|\zeta| \to \infty$. The solution of the problem is given by

$$\chi(\zeta)F(\zeta) = A + s\zeta + \frac{\pi}{2}\int_{-b}^{-c}\frac{dt}{\bar{\chi}(t)(t-\zeta)} - \frac{\pi}{2}\int_{o}^{\alpha}\frac{dt}{\bar{\chi}(t)(t-\zeta)} - \int_{\alpha}^{1}\frac{f_1(t)dt}{\bar{\chi}(t)(t-\zeta)} \quad (14)$$

where $\bar{\chi}(t) = |\chi(\zeta)|^{-1}$ and s, as before, is equal to 0.8562 $(K_{max}{}^L/K_{max}{}^U)^2$.

Imposing the boundedness conditions $F(-\gamma) = F(-c) = F(0) = F(\alpha) = 0$ gives four equations for the unknowns A, α, c, γ. They can in turn be combined to yield three equations for α, c, γ. Omitting details, we merely mention that these equations were solved numerically and the results are summarized in Fig. 9, in which the reversed yield size is plotted against the extent of crack growth b, ($b \equiv \bar{b}/w_{OL}$). Having determined α, c, γ as functions of b for a given value of $K_{max}{}^L/K_{max}{}^U$, we are in a position to calculate the opening loads. The displacement gradient in $(0, \alpha)$ which is needed to compute the opening loads may be calculated numerically from (21) (see Appendix) after α, c, γ are determined.

Discussion of Results

The results for $K_{cont}/K_{max}{}^U$ are plotted in Fig. 7 for various values of the load ratios $K_{max}{}^L/K_{max}{}^U$. We note that over a range of values of $K_{max}{}^L/K_{max}{}^U$, the initial contact load K_{cont} when little or no crack growth has taken place is typically quite small. For example, for an overload of a 100 percent ($K_{max}{}^L/K_{max}{}^U = 0.5$), the contact load immediately after $K_{max}{}^U$ is $K_{cont}/K_{max}{}^U = 0.023$. This means that initially the effective stress range $\Delta K_{eff} = K_{max}{}^U - K_{cont}$ is increased. If we were to assume that crack growth rate is primarily affected by

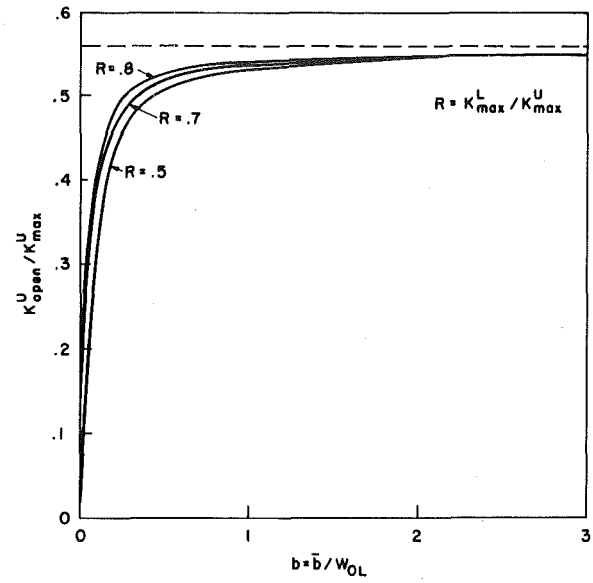


Fig. 10 Opening loads $K_{open}{}^U$ versus crack-tip position

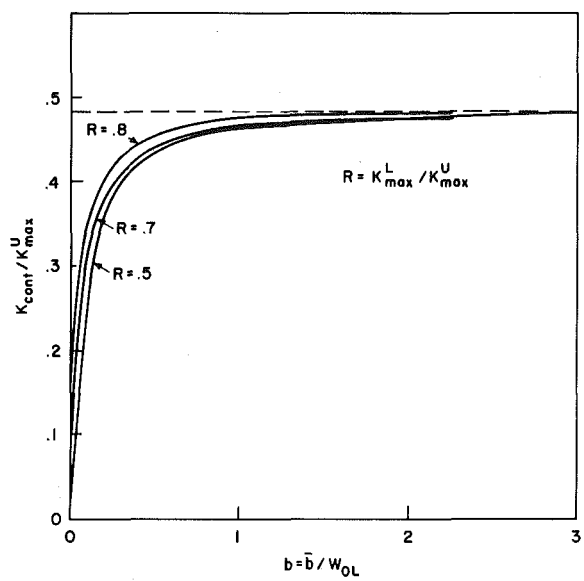


Fig. 11 Contact loads K_{cont} versus crack-tip position

$\Delta K_{eff} = K_{max}{}^U - K_{open}{}^U$ (or $K_{max}{}^U - K_{cont}{}^U$), our theoretical result is at least consistent with the often observed experimental fact [1] that there is an initial increase of crack growth rate for a very short period. We note also that the point of contact (see Fig. 8) as indicated by x_c (equation 5(b)) in general extends a large distance behind the crack tip; this is in contrast to the result in [12] that the contact point is very close to the crack tip ($x_c/w = -0.0243$) for a crack under constant amplitude cyclic load.

It is interesting to note that, according to the picture given in Fig. 5, two different values of opening loads can be defined. The lower $K_{open}{}^L$ is the load when the crack faces are opened up to the point $x = -\bar{c}$. Of greater interest, however, is the actual (higher) opening load $K_{open}{}^U$ and it is still given by equation (12) (with $f_2(\xi)$ calculated according to (21)). The numerical values of $K_{open}{}^U$ are summarized in Fig. 10. The closing loads K_{cont} are calculated according to (5a) with $\delta_R{}^U$ now computed from $f_2(\xi)$ in (21) and $f_1(\xi)$ in (2) for a range of values $\bar{b}$. We note that, from Figs. 10 and 11,

$K_{\text{open}}{}^U$ and $K_{\text{cont}}{}^U$ approach the steady-state situation ($\bar{b} \to \infty$) quite rapidly over a range of $K_{\max}{}^L/K_{\max}{}^U$. In fact, even for a 100 percent increase in the step load, $K_{\text{open}}{}^U$ reaches 95 percent of its limiting value (0.577 $K_{\max}{}^U$) once the crack has grown about one plastic zone size w_{OL}, assuming that small-scale yielding conditions are maintained. This means that the effect of prior history of loading is significant only in a range in which crack growth is less than w_{OL}. In the limit as $\bar{b} \to \infty$, Fig. 2(b) is recovered except that the plastic stretch is now given by $\delta_R{}^U$ associated with $K_{\max}{}^U$.

Acknowledgment

This work is supported by the Material Research Laboratory at Brown University funded by the NSF. The research was initiated at Harvard University under Grant AFOSR 77-3330 and part of the work was performed at the California Institute of Technology. Their support is gratefully acknowledged. The author would also like to thank Profs. J. W. Hutchinson and B. Budiansky for helpful discussions in connection with this work.

References

1 Matthews, W. T., Baratta, F. I., and Driscoll, G. W., "Experimental Observations of a Stress-Intensity History Upon Fatigue Crack Growth Rate," *International Journal of Fracture Mechanics*, Vol. 7, 1971, pp. 224–228.

2 Brown, R. D., and Weertmann, J., "Effect of Crack Closure and Crack Propagation Rates in 7050 Al.," *Engineering Fracture Mechanics*, Vol. 10, No. 4, 1978, pp. 867–878.

3 Elber, W., "Fatigue Crack Closure Under Cyclic Tension," *Engineering Fracture Mechanics*, Vol. 2, No. 1, July 1970, pp. 37–45.

4 Newman, J. C., Jr., "A Finite-Element Analysis of Fatigue Crack Closure," *Mechanics of Crack Growth*, ASTM STP 590, 1976, pp. 281–301.

5 Ohji, K., Ogua, K., and Ohkubo, K., "Cyclic Analysis of a Propagating Crack and Its Correlation With Fatigue Crack Growth," *Engineering Fracture Mechanics*, Vol. 7, 1975, pp. 457–464.

6 Ogura, K., and Ohji, K., "FEM Analysis of Crack Closure and Delay Effect in Fatigue Crack Growth Under Variable Amplitude Loading," *Engineering Fracture Mechanics*, Vol. 9, 1977, pp. 471–480.

7 Dill, H. D., and Saff, C. R., "Analysis of Crack Growth Following Compressive High Loads Based on Crack Surface Displacements and Contact Analysis," MCAIR 76-006, McDonnell Aircraft Co., 1976.

8 Führing, H., and Seeger, T., "Acceleration and Retardation Effects With Fatigue Crack Growth and Their Calculations Based on Fatigue Fracture Mechanics," *Proceedings of 2nd International Conference of Mechanical Behavior of Materials*, Boston, 1976, pp. 721–725.

9 Führing, H., and Seeger, T., "Dugdale Crack Analysis of Fatigue Cracks Under Constant Amplitude Loading," *Engineering Fracture Mechanics*, Vol. 11, No. 1, 1979, pp. 99–122.

10 Führing, H., and Seeger, T., "Remarks on Load-Interaction Effects Based on Fatigue Fracture Mechanics Calculations," *Proceedings of 9th ICAF Symposium on Aero. Fatigue*, 1977, Darmstadt.

11 McCartney, L. N., "A Theoretical Explanation of Delaying Effects of Overloads on Fatigue Crack Growth," *International Journal of Fracture*, Vol. 14, No. 2, Apr. 1978, pp. 213–232.

12 Budiansky, B., and Hutchinson, J. W., "Analysis of Closure in Fatigue Crack Growth," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, No. 2, June 1978, pp. 267–275.

13 Muskhelishvili, N. L., *Some Basic Problems of Mathematical Theory of Elasticity*, P. Noordhoff, 1958.

14 Rice, J. R., "Mathematical Analysis of Mechanics of Fracture," *Fracture*, ed., Liebowitz, H., Academic Press, 1968, Chapter 3.

APPENDIX

The general formulation of two-dimensional linear elasticity in terms of the complex potentials ϕ, ψ can be found in [13, 14]. The potential ψ can be expressed in terms of ϕ from stress continuity across the x-axis and the stresses are then given by

$$\left.\begin{aligned} \sigma_x + \sigma_y &= 4\,\mathrm{Re}\,[\phi'(z)] \\ \sigma_y - i\,\tau_{xy} &= \phi'(z) + \phi'(\bar{z}) + (z - \bar{z})\,\bar{\phi}''(\bar{z}) \end{aligned}\right\} \tag{15}$$

Since the solution we seek is symmetric and hence $\tau_{xy} = 0$ along the x-axis, this gives

$$\sigma_y = \phi_+' + \phi_-' \tag{16}$$

where the subscripts + and − denote the limit of ϕ' as z approaches the x-axis from above and below, respectively, i.e.,

$$\phi_\pm \equiv \lim_{y \to 0} \phi\,(x \pm iy), \quad y > 0$$

The displacements u and v in the x and y-directions, in plane stress, can be expressed as

$$E\frac{\partial}{\partial x}(u + iv) = (3 - \nu)\phi'(z) - (1 + \nu)\,[\bar{\phi}'(z) - (z - \bar{z})\,\bar{\phi}''(\bar{z})] \tag{17}$$

Since u is continuous across the x-axis, this implies that

$$\frac{d\delta}{dx} \equiv \left(\frac{\partial v}{\partial x}\right)_+ - \left(\frac{\partial v}{\partial x}\right)_- = \frac{4}{iE}\,[\phi_+' - \phi_-'] \tag{18}$$

Using the solution in (8) and (16), we can show that the stress σ_y in $(-\infty, -\gamma)$ for the problem in Fig. 3(b) is given by

$$\sigma_y/\sigma_Y = -\sqrt{\frac{\alpha - \xi}{-\xi - \gamma}}\left[\int_0^\alpha \sqrt{\frac{\alpha - \tau}{\tau + \gamma}}\,\frac{d\tau}{\tau - \xi} + \frac{4}{\pi^2}\int_\alpha^1 \sqrt{\frac{\tau - \alpha}{\tau + \gamma}}\,\frac{f_1(\tau)d\tau}{\tau - \xi}\right] \text{ for } \xi < -\gamma$$

For completeness, we restate the problem during reloading at K ($> K_{\text{open}}$). The boundary conditions along the x-axis are

$$\begin{aligned} F_+ + F_- &= 0 && \text{for} \quad \xi < 0 \\ F_+ + F_- &= \pi^2 && 0 < \xi < \beta \\ F_+ - F_- &= 2\pi i\, f_2(\xi) && \beta < \xi < \alpha \\ F_+ - F_- &= 2\pi i\, f_1(\xi) && \beta < \xi < 1 \end{aligned} \tag{19}$$

and $F \sim \pi(K/K_{\max}{}^U)/\sqrt{\zeta}$ as $|\zeta| \to \infty$ and β denotes the extent of tensile yielding ahead of the crack tip. The solution is given by

$$\sqrt{\zeta - \beta}\,F(\zeta) = \frac{\pi}{2}\int_0^\beta \frac{\sqrt{\beta - \xi}\,d\xi}{\xi - \zeta} - \int_\beta^\alpha \frac{\sqrt{\xi - \beta}}{\xi - \zeta} f_2(\xi)d\xi - \int_\alpha^1 \frac{\sqrt{\xi - \beta}}{\xi - \zeta} f_1(\xi)d\xi + \pi\,K/K_{\max}{}^U \tag{20}$$

Imposing $F(\beta) = 0$ and setting $\beta = 0$ gives (12).

With the use of (18) and (20), the displacement gradient in $(0, \alpha)$ is given by

$$\begin{aligned} -f_2(\xi) &\equiv \frac{d}{d\xi}(\delta_M/\delta_o{}^U) \\ &= -\frac{1}{\pi}\sqrt{\frac{\xi}{(\alpha - \xi)(\xi + c)(\xi + d)}} \\ &\times \left\{ s - \frac{\pi}{2}\,\mathcal{J}_0^\alpha \frac{dt}{t - \xi}\sqrt{\frac{(\alpha - t)(t + c)(t + d)}{t}} - \int_\alpha^1 \frac{f_1(t)dt}{t - \xi}\sqrt{\frac{(t - \alpha)(t + c)(t + d)}{t}} \right. \\ &\left. - \frac{\pi}{2}\int_{-b}^{-c} \frac{dt}{t - \xi}\sqrt{\frac{(\alpha - t)(-t - c)(t + d)}{-t}} \right\} \end{aligned} \tag{21}$$

A. F. Mak
L. M. Keer
Professor.
Mem. ASME

Department of Civil Engineering,
The Technological Institute,
Northwestern University,
Evanston, Ill. 60201

A No-Slip Edge Crack on a Bimaterial Interface[1]

A solution for an interface edge crack whose surfaces experience no tangential slip is presented. At the crack tip, K_I is found to be identical to that for a surface crack in a homogeneous medium. Moreover, the shear stress is singular both at the surface and at the crack tip, with K_{II} at the crack tip being equal to $-\beta K_I$, where β is Dundurs' constant. The crack opening displacements were calculated using a successive approximation scheme, which shows remarkable speed of convergence in this case.

Introduction

Because of the increasing use of adhesive structures in different areas of engineering, interface fracture mechanics has become an important research subject. Several interface crack models have been recently suggested [1, 2] to overcome the mathematical difficulty that the conventional treatment possesses, namely, the oscillating singularities and material overlapping at the crack tip [4, 5]. Also, there are compelling motivations from physical applications that require studies such as the present one [3].

We note in many actual engineering applications interfaces are very rough and interdigitated. The presence of such interdigitations can prevent relative slip at the interface or a portion thereof and thus enable shear loads to be sustained there even when the interface has failed in tension. A typical example is the interface between cancellous bone and PMMA, and in [3], an investigation of such a no-slip interface crack was solved for two bonded half planes. This paper continues the study and presents a solution for such an interface surface crack between two bonded quarter planes. In this investigation we use the work of Sneddon [6] as a point of departure. It will be seen that the method used in reference [6] provides a technique that most efficiently determines the pertinent physical quantities.

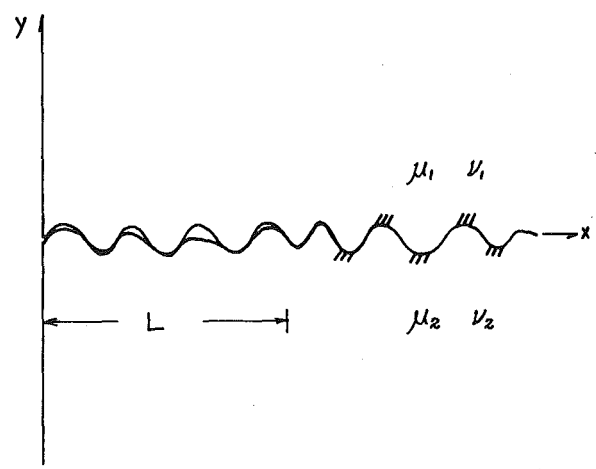


Fig. 1 Geometry and coordinate system for interface edge crack

Problem Formulation

Consider two elastic quarter planes of different materials that are bonded along the interface $y = 0$ except for a region $0 < x < L$ where adhesion perpendicular to the interface is lost (Fig. 1). The symbols μ and ν are, respectively, the shear modulus and Poisson's ratio. The derivation below is for plane strain; for plane stress replace ν by $\nu/(1+\nu)$. Note that due to interdigitations, relative shear motions are prohibited even within the region, $0 < x < L$. The boundary conditions are as follows:

$$\sigma^1_{yy}(x,0) = \sigma^2_{yy}(x,0) \tag{1a}$$

$$\sigma^1_{xy}(x,0) = \sigma^2_{xy}(x,0)\ x \geq 0 \tag{1b}$$

$$\sigma^1_{yy}(x,0) = \sigma^2_{yy}(x,0) = -P_0,\ 0 < x < L \tag{1c}$$

$$u^1_x(x,0) = u^2_x(x,0),\ x \geq 0 \tag{1d}$$

$$u^1_y(x,0) = u^2_y(x,0),\ x > L \tag{1e}$$

$$\sigma^1_{xx}(0,y) = \sigma^2_{xx}(0,y) = 0 \tag{1f}$$

$$\sigma^1_{xy}(0,y) = \sigma^2_{xy}(0,y) = 0\ y > 0, y < 0 \tag{1g}$$

[1] Part of this work was supported by Grant NIHR 23P-55898 and the Army Research Office under Grant DAAG 29-78-G-0199.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, April, 1980. Paper No. 80-WA/APM-14.

where σ^i_{xx}, σ^i_{yy}, σ^i_{xy} and u^i_x, u^i_y are stresses and displacements, respectively, (i = 1,2). Superscripts are used to distinguish the field quantities in their respective quarter planes (Fig. 1).

The displacements, which automatically satisfy the field equations of elasticity, can be expressed in terms of their suitably chosen Fourier sine and cosine transforms as follows [6]:

$$2\mu_1 u^1_x(x,y) = (2/\pi)^{1/2} \int_0^\infty [A_1(\xi) = y\xi B_1(\xi)]e^{-\xi y} \sin \xi x \, d\xi + (2/\pi)^{1/2} \int_0^\infty \zeta^{-1} C_1(\zeta)\,(2 - 2\nu_1 + \zeta x)e^{-\zeta x} \cos \zeta y \, d\zeta \quad (2a)$$

$$2\mu_2 u^2_x(x,y) = (2/\pi)^{1/2} \int_0^\infty [A_2(\xi) + y\xi B_2(\xi)]e^{\xi y} \sin \xi x \, d\xi + (2/\pi)^{1/2} \int_0^\infty \zeta^{-1} C_2(\zeta)\,(2 - 2\nu_2 + \zeta x)e^{-\zeta x} \cos \zeta y \, d\zeta \quad (2b)$$

$$2\mu_1 u^1_y(x,y) = (2/\pi)^{1/2} \int_0^\infty [A_1(\xi) + (3 - 4\nu_1 + y\xi)B_1(\xi)]e^{-\xi y} \cos \xi x \, d\xi - (2/\pi)^{1/2} \int_0^\infty \zeta^{-1} C_1(\zeta) \times (1 - 2\nu_1 - \zeta x)e^{-\zeta x} \sin \zeta y \, d\zeta, \quad (2c)$$

$$2\mu_2 u^2_y(x,y) = (2/\pi)^{1/2} \int_0^\infty [-A_2(\xi) + (3 - 4\nu_2 - y\xi)B_2(\xi)]e^{\xi y} \cos \xi x \, d\xi - (2/\pi)^{1/2} \int_0^\infty \zeta^{-1} C_2(\zeta) \times (1 - 2\nu_2 - \zeta x)e^{-\zeta x} \sin \zeta y \, d\zeta. \quad (2d)$$

The corresponding interfacial displacements and stresses are thus written as follows:

$$2\mu_1 u^1_x(x,0) = (2/\pi)^{1/2} \int_0^\infty A_1(\xi) \sin \xi x \, d\xi + (2/\pi)^{1/2} \int_0^\infty \zeta^{-1} C_1(\zeta)(2 - 2\nu_1 + \zeta x)e^{-\zeta x} \, d\zeta, \quad (3)$$

$$2\mu_1 u^1_y(x,0) = (2/\pi)^{1/2} \int_0^\infty [(3 - 4\nu_1)B_1(\xi) + A_1(\xi)]\cos \xi x \, d\xi, \quad (4)$$

$$2\mu_2 u^2_x(x,0) = (2/\pi)^{1/2} \int_0^\infty A_2(\xi) \sin \xi x \, d\xi + (2/\pi)^{1/2} \int_0^\infty \zeta^{-1} C_2(\zeta)(2 - 2\nu_2 + \zeta x)e^{-\zeta x} \, d\zeta, \quad (5)$$

$$2\mu_2 u^2_y(x,0) = (2/\pi)^{1/2} \int_0^\infty [(3 - 4\nu_2)B_2(\xi) - A_2(\xi)]\cos \xi x \, d\xi, \quad (6)$$

$$\sigma^1_{yy}(x,0) = -(2/\pi)^{1/2} \int_0^\infty [2(1 - \nu_1)B_1(\xi) + A_1(\xi)]\xi \cos \xi x \, d\xi - (2/\pi)^{1/2} \int_0^\infty C_1(\zeta)(1 - \zeta x)e^{-\zeta x} \, d\zeta, \quad (7)$$

$$\sigma^1_{xy}(x,0) = -(2/\pi)^{1/2} \int_0^\infty [(1 - 2\nu_1)B_1(\xi) + A_1(\xi)]\xi \sin \xi x \, d\xi, \quad (8)$$

$$\sigma^2_{yy}(x,0) = (2/\pi)^{1/2} \int_0^\infty [2(1 - \nu_2)B_2(\xi) - A_2(\xi)]\xi \cos \xi x \, d\xi - (2/\pi)^{1/2} \int_0^\infty C_2(\zeta)(1 - \zeta x)e^{-\zeta x} \, d\zeta, \quad (9)$$

and

$$\sigma^2_{xy}(x,0) = -(2/\pi)^{1/2} \int_0^\infty [(1 - 2\nu_2)B_2(\xi) - A_2(\xi)]\xi \sin \xi x \, d\xi. \quad (10)$$

The stress components on the edge $x = 0$ are

$$\sigma^1_{xx}(0,y) = (2/\pi)^{1/2} \int_0^\infty [(\xi y - 2\nu_1)\xi B_1(\xi) + \xi A_1(\xi)]e^{-\xi y} \, d\xi - (2/\pi)^{1/2} \int_0^\infty C_1(\zeta) \cos \zeta y \, d\zeta, \quad (11a)$$

$$\sigma^2_{xx}(0,y) = (2/\pi)^{1/2} \int_0^\infty [(\xi y + 2\nu_1)\xi B_2(\xi) + \xi A_2(\xi)]e^{\xi y} \, d\xi - (2/\pi)^{1/2} \int_0^\infty C_2(\zeta) \cos \zeta y \, d\zeta, \quad (11b)$$

and, clearly,

$$\sigma^1_{xy}(0,y) = \sigma^2_{xy}(0,y) \equiv 0. \quad (11c)$$

Thus boundary condition (1g) is satisfied automatically.

Writing

$$\begin{aligned} \xi B_1(\xi) &= D_1(\xi) + E(\xi) \\ \xi A_1(\xi) &= -2(1 - \nu_1)E(\xi) - (1 - 2\nu_1)D_1(\xi) \\ \xi B_2(\xi) &= D_2(\xi) + E(\xi) \\ \xi A_2(\xi) &= 2(1 - \nu_2)E(\xi) + (1 - 2\nu_2)D_2(\xi) \end{aligned} \quad (12)$$

the interfacial shear stress can be reduced to a Fourier sine transform as

$$\sigma^1_{xy}(x,0) = \sigma^2_{xy}(x,0) = (2/\pi)^{1/2} \int_0^\infty E(\xi) \sin \xi x \, d\xi \quad (13)$$

which satisfies boundary condition (1b) automatically.

Imposing boundary condition (1f) and taking the inverse cosine transform for equations (11a) and (11b), $C_1(\zeta)$ and $C_2(\zeta)$ can be expressed explicitly as

$$C_1(\zeta) = -\int_0^\infty \{[3K_2(\xi,\zeta) + K_1(\xi,\zeta)]E(\xi) + 2K_2(\xi,\zeta)D_1(\xi)\}d\xi, \quad (14a)$$

$$C_2(\zeta) = -\int_0^\infty \{[3K_2(\xi,\zeta) + K_1(\xi,\zeta)]E(\xi) + 2K_2(\xi,\zeta)D_2(\xi)\}d\xi, \quad (14b)$$

where

$$K_1(\xi,\zeta) = (2/\pi)\xi^3/(\zeta^2 + \xi^2)^2 \quad (15a)$$

$$K_2(\xi,\zeta) = (2/\pi)\xi\zeta^2/(\zeta^2 + \xi^2)^2. \quad (15b)$$

Summing equations (7) and (9) and using boundary condition (1c), we find

$$(2/\pi)^{1/2} \int_0^\infty [D_2(\xi) - D_1(\xi)]\cos \xi x \, d\xi - (2/\pi)^{1/2} \int_0^\infty \int_0^\infty 2K_2(\xi,\zeta)[D_2(\xi) - D_1(\xi)](1 - \zeta x)e^{-\zeta x} d\xi d\zeta = -2P_0, \quad x < L. \quad (16)$$

Boundary condition (1a) implies

$$(2/\pi)^{1/2} \int_0^\infty [D_1(\xi) + D_2(\xi)]\cos \xi x d\xi = (2/\pi)^{1/2} \int_0^\infty [C_2(\zeta) - C_1(\zeta)](1 - \zeta x)e^{-\zeta x} d\zeta. \quad (17)$$

Substituting C_1, C_2 from equations (14a), (14b) and using the definitions of the kernels below, which are found upon taking the indicated inverse cosine transforms

$$L_{22}(\eta,\xi) \equiv \int_0^\infty K_2(\zeta,\xi)K_2(\eta,\zeta)d\zeta = \frac{4}{\pi^2}\frac{\eta\xi^2}{(\xi^2 - \eta^2)^2}\left\{\frac{\xi^2 + \eta^2}{\xi^2 - \eta^2}\log\left(\frac{\xi}{\eta}\right) - 1\right\}, \quad (18a)$$

$$L_{12}(\eta,\xi) \equiv \int_0^\infty K_1(\zeta,\xi)K_2(\eta,\zeta)d\zeta$$

$$= \frac{4}{\pi^2}\frac{\eta}{(\xi^2-\eta^2)^2}\left\{\frac{-2\eta^2\xi^2}{\xi^2-\eta^2}\log\left(\frac{\xi}{\eta}\right)+\frac{\xi^2+\eta^2}{2}\right\}, \quad (18b)$$

$$L_{11}(\eta,\xi) \equiv \int_0^\infty K_1(\zeta,\xi)K_1(\eta,\zeta)d\zeta$$

$$= \frac{4}{\pi^2}\frac{\eta^3}{(\xi^2-\eta^2)^2}\left\{\frac{\xi^2+\eta^2}{\xi^2-\eta^2}\log\left(\frac{\xi}{\eta}\right)-1\right\}, \quad (18c)$$

equation (17) becomes

$$D_1(\xi)+D_2(\xi) = 4\int_0^\infty \{[3L_{22}(\eta,\xi)+L_{12}(\eta,\xi)]E(\eta) + L_{22}(\eta,\xi)[D_1(\eta)+D_2(\eta)]\}d\eta. \quad (19)$$

Similarly, boundary condition (1d) leads to

$$\left(\frac{1-\nu_1}{\mu_1}+\frac{1-\nu_2}{\mu_2}\right)E(\xi)+\frac{1-2\nu_2}{2\mu_2}D_2(\xi)+\frac{1-2\nu_1}{2\mu_1}D_1(\xi)$$
$$= \int_0^\infty \left\{\left[\frac{2(1-\nu_2)}{\mu_2}L_{12}(\eta,\xi)-\frac{2\nu_2}{\mu_2}L_{22}(\eta,\xi)\right]D_2(\eta) + \left[\frac{2(1-\nu_1)}{\mu_1}L_{12}(\eta,\xi)-\frac{2\nu_1}{\mu_1}L_{22}(\eta,\xi)\right]D_1(\eta) + \left[\left(\frac{3-4\nu_2}{\mu_2}+\frac{3-4\nu_1}{\mu_1}\right)L_{12}(\eta,\xi) - 3\left(\frac{\nu_2}{\mu_2}+\frac{\nu_1}{\mu_1}\right)L_{22}(\eta,\xi)+\left(\frac{1-\nu_2}{\mu_2}+\frac{1-\nu_1}{\mu_1}\right)L_{11}(\eta,\xi)\right]E(\eta)\right\}d\eta. \quad (20)$$

By defining

$$b(x) \equiv (\partial/\partial x)(u_y^1(x,0)-u_y^2(x,0)), \quad (21)$$

boundary condition (1e) can be rewritten as

$$-\left(\frac{1-2\nu_1}{2\mu_1}-\frac{1-2\nu_2}{2\mu_2}\right)E(\xi)-\frac{1-\nu_1}{\mu_1}D_1(\xi)+\frac{1-\nu_2}{\mu_2}D_2(\xi) = (2/\pi)^{1/2}\int_0^L b(\eta)\sin\xi\eta\, d\eta. \quad (22)$$

We let

$$\eta = \xi z, \quad (23)$$

$$\alpha = \frac{\mu_2(1-\nu_1)-\mu_1(1-\eta_2)}{\mu_2(1-\nu_1)+\mu_1(1-\nu_2)}, \quad (24a)$$

$$\beta = \frac{\mu_2(1-2\nu_1)-\mu_1(1-2\nu_2)}{2[\mu_2(1-\nu_1)+\mu_1(1-\nu_2)]}, \quad (24b)$$

where α and β are Dundurs' constants [7], and further define

$$2\mathcal{D}_2(\xi) \equiv D_2(\xi)-D_1(\xi) \quad (25a)$$

and

$$2\mathcal{D}_1(\xi) \equiv D_2(\xi)+D_1(\xi). \quad (25b)$$

The system of integral equations which needs to be solved becomes

$$(2/\pi)^{1/2}\int_0^\infty \mathcal{D}_2(\xi)\cos\xi x\, d\xi - (2/\pi)^{1/2}\int_0^\infty\int_0^\infty 2K_2(\xi,\zeta)\mathcal{D}_2(\xi)(1-\zeta x)e^{-\zeta x}d\xi d\zeta = -P_0, x<1 \quad (26)$$

$$\mathcal{D}_1(\xi) = 2\int_0^\infty \{[3L_3(z)+L_2(z)]E(\xi z) + 2L_3(z)\mathcal{D}_1(\xi z)\}dz, \quad \xi>0 \quad (27)$$

$$E(\xi) = \int_0^\infty \{[L_2(z)-L_3(z)]2\mathcal{D}_1(\xi z) + [2L_2(z)-3L_3(z)+L_1(z)]E(\xi z)\}dz \quad (28)$$

$$+ \int_0^\infty \{(\alpha-2\beta)L_3(z)-\alpha L_2(z)\}\,2\mathcal{D}_2(\xi z)dz + \beta\mathcal{D}_2(\xi), \quad \xi>0 \quad (28)$$

(*Cont.*)

$$[\mathcal{D}_2(\xi)-\alpha\mathcal{D}_1(\xi)-\beta E(\xi)]\cdot\left[\frac{1-\nu_1}{\mu_1}+\frac{1-\nu_2}{\mu_2}\right] = (2/\pi)^{1/2}\int_0^1 b(\zeta)\sin\xi\zeta d\zeta, \quad \xi>0 \quad (29)$$

where x has been nondimensionalized with respect to the crack length L, and

$$L_1(z) \equiv -\frac{4}{\pi^2}z^3\left\{\frac{(1+z)\log z}{(1-z^2)^3}+\frac{1}{(1-z^2)^2}\right\}, \quad (30a)$$

$$L_2(z) \equiv \frac{4}{\pi^2}z\left\{\frac{2z^2\log z}{(1-z^2)^3}+\frac{1+z^2}{2(1-z^2)^2}\right\}, \quad (30b)$$

$$L_3(z) \equiv -\frac{4}{\pi^2}z\left\{\frac{(1+z)\log z}{(1-z^2)^3}+\frac{1}{(1-z^2)^2}\right\}. \quad (30c)$$

We define the Fourier sine transforms of $E(\xi)$, $\mathcal{D}_1(\xi)$, $\mathcal{D}_2(\xi)$ as

$$e(s) = (2/\pi)^{1/2}\int_0^\infty E(\xi)\sin\xi s\, d\xi, \quad (31a)$$

$$d_1(s) = (2/\pi)^{1/2}\int_0^\infty \mathcal{D}_1(\xi)\sin\xi s\, d\xi, \quad (31b)$$

$$d_2(s) = (2/\pi)^{1/2}\int_0^\infty \mathcal{D}_2(\xi)\sin\xi s\, d\xi. \quad (31c)$$

In order to be consistent with equation (29), we require these transforms be identically zero for $s>1$. Then, after some changes of variables, equations (26)–(28) can be rewritten as the following system of Fredholm integral equations:

$$\frac{1}{\pi}\int_0^1 d_2(s)\frac{1}{s-x}ds - \frac{1}{\pi}\int_0^1 d_2(s)\left[\frac{2s(s-x)}{(s+x)^3}-\frac{1}{s+x}\right]ds = -P_0, \quad 0<x<1 \quad (32)$$

$$d_1(x) = 2\int_0^1\frac{1}{s}\left\{e(s)\cdot\left[3L_3\left(\frac{x}{s}\right)+L_2\left(\frac{x}{s}\right)\right] + 2d_1(s)L_3\left(\frac{x}{s}\right)\right\}ds, \quad 0<x<1 \quad (33)$$

$$e(x) = \int_0^1\frac{1}{s}\left\{e(s)\cdot\left[2L_2\left(\frac{x}{s}\right)-3L_3\left(\frac{x}{s}\right)+L_1\left(\frac{x}{s}\right)\right] + 2d_1(s)\cdot\left[L_2\left(\frac{x}{s}\right)-L_3\left(\frac{x}{s}\right)\right]\right\}ds + \int_0^1 2d_2(s)\cdot\frac{1}{s}\left[(\alpha-2\beta)L_3\left(\frac{x}{s}\right)-\alpha L_2\left(\frac{x}{s}\right)\right]ds + \beta d_2(x), \quad 0<x<1. \quad (34)$$

The remarkable feature of this system of equations is that $d_2(s)$, which will give the Mode-I stress-intensity factor at $x=L$, only depends upon the load and not upon the elastic constants. Equation (25) is immediately recognized as that associated with an edge crack perpendicular to the surface of a homogeneous elastic half plane [8–10].

Solution

Equation (32) is first solved independently for $d_2(s)$. After $d_2(s)$ is obtained, it is used in equations (33) and (34), which are then solved by successive approximations. Solutions to the edge crack problem have been given by many authors, see, e.g. [8–10], and we recalculate only to obtain $d_2(s)$, which is to be used later in the successive approximation scheme.

Changing the variables by

$$r = 2x - 1, \quad t = 2s - 1, \tag{35}$$

and writing $d_2(t)$ of the form

$$d_2(t) = d_2^*(t)/(1 - t^2)^{1/2}, \tag{36}$$

equation (32) was solved using the Gauss-Chebyshev integration formula [11] with 40 integrating points. The values of $d_2^*(t)$ obtained are shown in Table 1.

We obtain for K_{I} the value

$$K_{\mathrm{I}} = \lim_{x \to 1^+} [2(x-1)]^{1/2}\, \sigma_{yy}(x,0) = -d_2^*(1)/\sqrt{2} = 1.1213, \tag{37}$$

which is in agreement with the references just cited.

To solve equations (26) and (27) by successive approximations, substitute

$$d_2(x) = d_2^*(2x-1)/2[x(1-x)]^{1/2} \tag{38}$$

back into equation (27). For the first approximations of $d_1(x)$ and $e(x)$, denoted by ${}_1d_1(x)$, ${}_1e(x)$, respectively, the contributions from the terms with kernels L's are assumed to be small compared to $d_2(x)$. Hence

$${}_1d_1(x) = 0 \tag{39a}$$

$${}_1e(x) = \beta \cdot d_2^*(2x-1)/2[x(1-x)]^{1/2}. \tag{39b}$$

From equation (13), one can write

$$\sigma_{xy}(x,0) = e(x)\mathrm{H}(1-x), \tag{40a}$$

where $\mathrm{H}(x)$ is the Heaviside unit step function and $e(x)$ is related to $E(\xi)$ by

$$E(\xi) = \sqrt{\frac{2}{\pi}} \int_0^1 e(\rho) \sin(\rho\xi) d\rho. \tag{40b}$$

The first approximation of K_{11} is obtained as

$$K_{\mathrm{II}}\Big|_{x=1^-} = \lim_{x \to 1^-} [2(1-x)]^{1/2} \cdot \sigma_{xy}(x,0) = \beta d_2^*(1)/\sqrt{2} = -1.1213\,\beta, \tag{41a}$$

$$K_{\mathrm{II}}\Big|_{x=0} = \lim_{x \to 0^+} [2x]^{1/2} \cdot \sigma_{xy}(x,0) = \beta d_2^*(-1)/\sqrt{2} = -0.0932\,\beta. \tag{41b}$$

It can be shown that higher approximations do not contribute to K_{II}. Note that at the crack tip,

$$K_{\mathrm{II}}\Big|_{x=1^-} = -\beta K_{\mathrm{I}}\Big|_{x=1^+}. \tag{42}$$

To determine the first approximation of the crack opening displacements $\delta(x)$, recall

$$b(x) = \left[\frac{1-\nu_1}{\mu_1} + \frac{1-\nu_2}{\mu_2}\right] [d_2(x) - \alpha d_1(x) - \beta e(x)], \quad 0 < x < 1. \tag{43}$$

Substituting ${}_1e(x)$ and ${}_1d_1(x)$ into equation (43), the first approximation ${}_1b(x)$ is readily obtained as

$${}_1b(x) = (1-\beta^2)\left[\frac{1-\nu_1}{\mu_1} + \frac{1-\nu_2}{\mu_2}\right] d_2^*(2x-1)/2[x(1-x)]^{1/2}. \tag{44}$$

The corresponding crack opening displacements $\delta(x)$ can be found from

$${}_1\delta(x) = -\int_x^1 {}_1b(\gamma) d\gamma, \quad 0 < x < 1. \tag{45}$$

To calculate the second approximation, ${}_1e(x)$ and ${}_1d_1(x)$ were substituted back into the right-hand side of equations (33) and (34), the resultant ${}_2e(x)$ and ${}_2d_1(x)$ obtained on the left-hand side were then put into equations (43) and (45) to determine ${}_2\delta(x)$. These procedures can be carried out successively until the maximum difference between ${}_i\delta(x)$ and ${}_{i-1}\delta(x)$ is found to be less than a preset value.

Table 1 The values of $d_2^*(t)$

t	$d_2^*(t)$	t	$d_2^*(t)$
0.9993	-1.585	0.0000	-0.937
0.9934	-1.581	-0.0766	-0.896
0.9817	-1.572	-0.1526	-0.855
0.9643	-1.559	-0.2279	-0.816
0.9411	-1.543	-0.3017	-0.778
0.9125	-1.522	-0.3738	-0.740
0.8785	-1.497	-0.4437	-0.704
0.8394	-1.469	-0.5110	-0.667
0.7953	-1.438	-0.5753	-0.630
0.7466	-1.404	-0.6362	-0.592
0.6934	-1.367	-0.6934	-0.553
0.6362	-1.328	-0.7466	-0.511
0.5753	-1.287	-0.7953	-0.467
0.5110	-1.244	-0.8394	-0.420
0.4437	-1.201	-0.8785	-0.370
0.3738	-1.157	-0.9125	-0.317
0.3017	-1.112	-0.9411	-0.262
0.2279	-1.068	-0.9643	-0.203
0.1526	-1.023	-0.9817	-0.143
0.0766	-0.980	-0.9934	-0.079

Table 2 The convergence of $10^8 \cdot \delta(x)$ at $x = 0$ and $1/2$

x	${}_1\delta$	${}_3\delta$	${}_5\delta$	${}_7\delta$	${}_9\delta$	${}_{11}\delta$	${}_{13}\delta$
0.	0.5907	0.5748	0.5783	0.5805	0.5813	0.5817	0.5818
0.5	0.4272	0.4210	0.4220	0.4228	0.4230	0.4232	0.4232

Table 2 shows the convergence of ${}_i\delta(x)$ at two particular locations of x. The interface was taken to be that between cancellous bone and PMMA, the material properties of which are listed in Table 3.

One should note that the first approximation already gives a very good estimate of the actual crack opening displacements. If only the first approximation is used, one is in error by about 1.5 percent at maximum. The shape of the crack opening is shown in Fig. 2.

Discussion and Conclusions

It is interesting to note that the K_{I} found at the tip of this type of

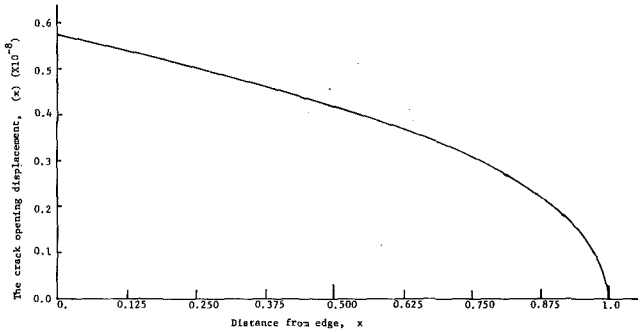


Fig. 2 The shape of the crack opened under uniform pressure; the interface was taken to be that between PMMA and cancellous bone with material properties given in Table 3

interface crack is identical to that for an edge crack in a homogeneous material. This feature is implicit in equation (32), which is identical to the singular integral equation associated with the homogeneous edge crack. The same analog has been demonstrated for the interior crack in the full plane [3].

Though the relative shear displacements were prescribed to be zero along the whole interface, the resultant shear stresses have been shown to be singular at the crack tip and only zero outside the crack region. More interesting, K_{II} as $x \to 1^-$ is equal to $-\beta K_I$, where β is Dundurs' constant. Again, this feature was also demonstrated in the case of interior crack [3].

The method of successive approximations which was used to calculate the crack opening displacements showed excellent convergence in this application. For relatively little computational effort, excellent estimates of the crack opening displacements can be obtained. The first approximation alone is within 1.5 percent of the final solution.

The extension of this method to solve the problem with nonuniform pressure as a function of x is obvious. One can just replace $-P_0$ in equation (32) with $-P(x)$ and proceed accordingly in exactly the same manner.

For a linear loading of the form

$$P(x) = -P_0(1 - x/L), \quad 0 < x < L, \tag{46}$$

the resultant stress-intensity factors are shown in Table 4, which also lists, for the sake of comparison, the corresponding stress-intensity factors due to uniform pressure as obtained in the foregoing.

As a final remark, it is interesting to note that for linear pressure loading, K_{II} at $x = 0$ is twice as large as that for the uniform case.

Table 3 The elastic properties of PMMA and cancellous bone

material	properties	
	$\mu(N/m^2)$	ν
(1) PMMA	2.0×10^8	0.4
(2) Cancellous Bone	1.3×10^8	0.32

Table 4 Comparison of the stress-intensity factors due to uniform loading and linear loading

	Uniform Loading	Linear Loading
K_I $x \to 1^+$	1.1213	0.8768
K_{II} $x \to 0^+$	$-0.0932\ \beta$	$-0.1864\ \beta$
K_{II} $x \to 1^-$	$-1.1213\ \beta$	$-0.8768\ \beta$

References

1 Comninou, M., "The Interface Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 631–636.

2 Achenbach, J. D., et al., "Loss of Adhesion at the Tip of an Interface Crack," *Journal of Elasticity* Vol. 9, 1979, pp. 397–424.

3 Mak, A. F., et al., "A No-Slip Interface Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 347–350.

4 Williams, M. L., "The Stresses Around a Fault or Crack in Dissimilar Media," *Bulletin of the Seismological Society of America,* Vol. 49, 1959, pp. 199–204.

5 England, A. H., "A Crack Between Dissimilar Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, 1965, pp. 400–440.

6 Sneddon, I. N., "Fourier Transform Solutions of Quarter-Plane Problems in Elasticity," File No. PSR-99/6, Applied Mathematics Research Group, North Carolina State University, May 1971.

7 Dundurs, J., Discussion, ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, 1967, pp. 650–652.

8 Sneddon, I. N., and Das, S. C., "The Stress-Intensity Factor at the Tip of an Edge Crack in an Elastic Half Plane," *International Journal of Engineering Science,* Vol. 9, 1971, pp. 25–36.

9 Koiter, W. T., Discussion, ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, 1965, p. 237.

10 Koiter, W. T., "On the Flexural Rigidity of a Beam, Weakened by Transverse Saw Cuts, II," *Proceedings, Kon. Ned. Akad. Wet.,* Amsterdam, Vol. 59B, 1956, pp. 365–374.

11 Erdogan, F., Gupta, G. D., Cook, T. S., "Numerical Solution of Singular Integral Equations," *Mechanics of Fracture,* Vol. 1, ed., Sih, G., Noordhoff International Publishing, Leydon, 1973.

G. R. Speich
A. J. Schwoeble
B. M. Kapadia
Research Laboratory,
United States Steel Corporation,
Monroeville, Pa. 15146

Elastic Moduli of Gray and Nodular Cast Iron

The elastic moduli of both gray and nodular cast iron were measured by a pulse-echo elastic-wave technique at ambient and elevated temperatures up to 760°C. When compared at similar graphite contents, the elastic moduli of gray cast iron were much lower than those of nodular cast iron. These results are satisfactorily explained by theoretical models for the elastic moduli of two-phase solids which take into account not only the volume fraction but also the shape of the graphite particles. The temperature-dependence of the elastic moduli of both gray and nodular cast iron can also be correctly predicted from these same models.

Introduction

The general problem of predicting the elastic moduli of a two-phase solid from the elastic moduli and volume fractions of each of the two phases has been studied by a number of investigators [1–5]. The solution of the problem requires that isotropic elastic behavior be assumed for both phases, and that interaction effects between the particles can be ignored. In many polycrystalline metals that contain large, widely spaced particles these approximations appear to be reasonable.

Hashin [1] has solved the problem for the case of spherical particles. Wu [2, 3] and Rossi [4] have given solutions for the case of needle or disk-shaped particles. Wu and Rossi have both shown that the elastic moduli of the two-phase solid are more strongly affected by disk-shaped particles than by spherical particles. Budiansky and O'Connel [5] have solved the problem for the special case of a cracked solid (second phase is a disk-shaped hole).

The elastic moduli of gray and nodular cast-iron offer an interesting example of the effect of particle shape. The elastic moduli of both gray and nodular cast iron decrease with increasing graphite content. However, when compared at similar graphite contents, the elastic moduli of gray cast iron, in which the graphite particles are disk-shaped (flakes), are much lower than those of nodular cast iron, in which the graphite particles are nearly spherical [6]. In the present study, the elastic moduli of several gray cast irons and one nodular cast iron were measured at ambient and elevated temperatures. The theories of Hashin [1] of Wu [2, 3] and of Rossi [4] were used to explain the large effect of graphite particle shape on the elastic moduli of cast irons.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. *Manuscript received by ASME Applied Mechanics Division, January, 1979; final revision, June, 1980.*

Table 1 Chemical compositions of gray and nodular cast irons—percent

Gray Cast Iron I

C	Graphite	Mn	P	S	Si	Cr	Mo
3.00	2.60	0.70	0.089	0.02	1.90	0.17	0.06

Gray Cast Iron II

C	Graphite	Mn	P	S	Si	Ti
4.30	4.1	0.53	0.076	0.025	1.52	0.040

Nodular Cast Iron

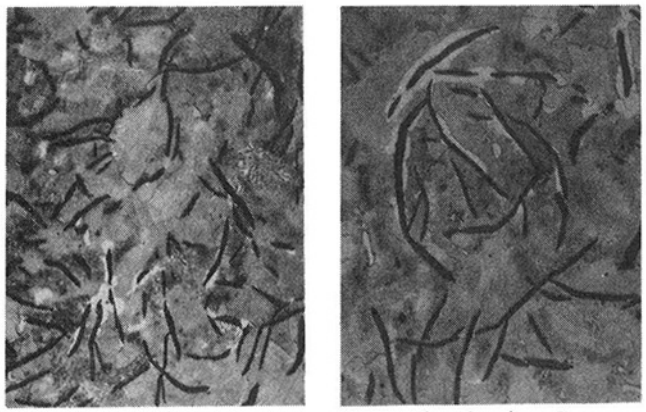
C	Graphite	Mn	P	S	Si	Ni	Ti	Mg
3.69	2.9	0.67	0.055	0.005	2.33	0.70	0.024	0.043

Experimental Procedure

The chemical compositions of the two gray cast irons and the nodular cast iron are given in Table 1. Specimens of the lower carbon gray cast iron I were were obtained from the jamb of a coke-oven door. Specimens of the higher carbon gray cast iron II were obtained from an ingot mold. Because of variations in graphite flake size and amount with cooling rate during casting, specimens from two representative locations (designated *A* and *B*) in the casting were obtained from each of the gray cast irons. Specimens of the nodular cast iron were obtained from a test bar cast from a Mg-innoculated heat.

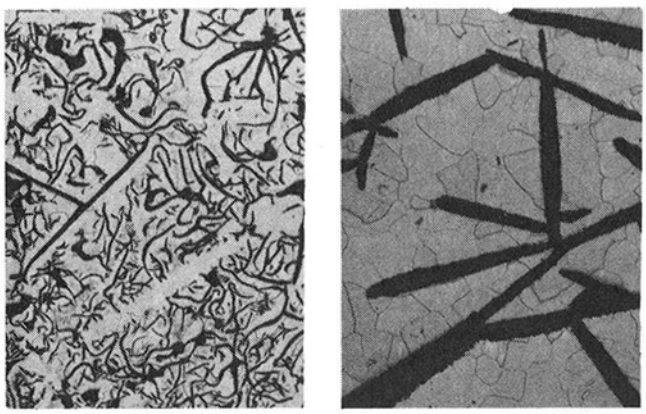
Rod specimens with a diameter of 3 mm and a length of 76 mm were machined from gray cast iron I and from the nodular cast iron for a dynamic determination of the elastic moduli. For the more brittle gray cast iron II, 12.7-mm-thick plate specimens were surface-ground and used for a dynamic determination of the elastic moduli.

Although specimens of both the gray and nodular cast irons were examined in the as-cast condition, several specimens of the nodular cast iron were also examined in the annealed and normalized-and-tempered conditions to determine the possible effects of matrix microstructure (ferrite or pearlite) on the elastic moduli. The annealing



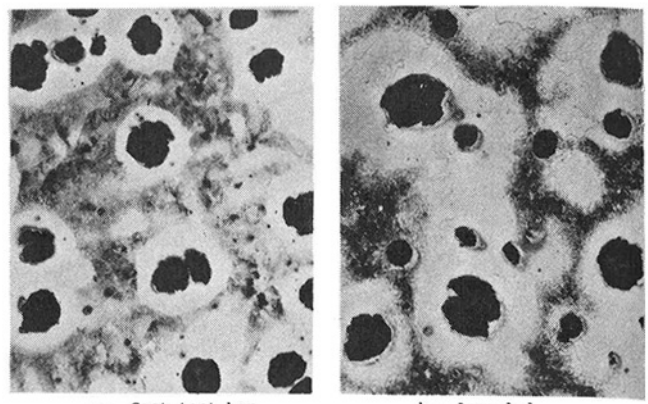
a. Specimen A. b. Specimen B.

Fig. 1 Microstructure of gray cast iron I. X100.

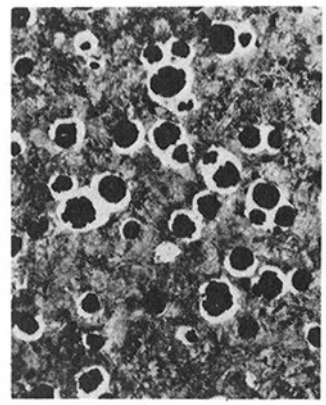


a. Specimen A. b. Specimen B.

Fig. 2 Microstructure of gray cast iron II. X100.



a. Cast test bar b. Annealed



c. Normalized and tempered

Fig. 3 Microstructure of nodular cast iron. X200.

treatment consisted of heating the specimen to 732°C, holding for 1 hr, and furnace cooling. The normalizing and tempering treatment consisted of heating the specimen to 898°C for 1.5 hr, followed by air cooling and then tempering at 537°C for 1.5 hr.

The elastic moduli of both gray cast iron I and the nodular cast iron were measured by the thin-rod, pulse-echo technique (100 kHz) described earlier [7]. For elevated-temperature measurements, the thin rods were placed in an inert-gas atmosphere furnace and slowly heated from ambient to temperatures of 760°C. A tungsten wire spot-welded to the rods served as a waveguide to transmit and receive elastic pulses from the specimen. The transit times for both dilatational and shear waves were measured at 100°C intervals on a Tektronix 7704 oscilloscope, previously calibrated with a time-mark generator.

The elastic moduli of gray cast iron II were measured on the plate specimens using 1 MHz dilatational and shear waves generated by thin quartz crystals with a pulse-echo-overlap technique [8]. These measurements were restricted to ambient temperature.

Because densities are required for the determination of the elastic moduli from the elastic wave velocities, the densities of all the specimens were measured at ambient temperature by a standard weight-loss technique [9]. Specimen lengths and densities were corrected for thermal expansion at elevated temperatures by using the data of Nix and McNair for pure iron [10].

All the specimens were examined metallographically after polishing and etching in 2 percent nital, and the volume fraction, mean size, and axial ratio of the graphite flakes or nodules were determined by quantitative metallographic techniques [11,12].

Results

Microstructural Observations. Typical microstructures of gray cast iron I are shown in Figs. 1(*a*, *b*). Specimen *A*, which was taken from a thinner section of the casting, had slightly smaller graphite flakes than Specimen *B*, because of its higher cooling rate during solidification. The matrix phase of both specimens was pearlite.

Typical microstructures of gray cast iron II are shown in Figs. 2(*a*, *b*). The graphite flakes in specimen *A* were much smaller than those in specimen *B* because specimen *A* was located closer to a chill casting stool. The matrix phase of both specimens was primarily ferrite.

The microstructures of the nodular cast iron in the as-cast condition, after annealing, and after normalizing and tempering are shown in Figs. 3(*a*, *b*, *c*), respectively. The as-cast specimen contained spheroidal graphite nodules which were surrounded by a ferrite rim; the remainder of the matrix being pearlite. The annealed specimen had a similar microstructure but contained a larger amount of ferrite. The normalized and tempered specimen contained a relatively small amount of ferrite and a correspondingly larger amount of pearlite. The smaller graphite nodules in the normalized-and-tempered specimen are not a result of the heat treatment but are believed to result from variations in the original as-cast specimens caused by differences in solidification rate.

The density of the nodular and gray cast irons is given in Table 2. The densities of both gray cast irons and of the nodular cast iron were all considerably lower than that of iron because the density of graphite (2.265 g/cm^3) is much less than that of iron (7.870 g/cm^3). Likewise, the density of gray cast iron II was much less than that of gray cast iron I because of its much higher graphite content.

The volume fractions of graphite in the nodular and gray cast irons are given in Table 3. The volume fraction of graphite was slightly different in specimens *A* and *B* for each of the gray cast irons, presumably because of differences in cooling rate during solidification.

Table 2 Density of gray and nodular cast irons

Type Cast Iron	Specimen	Density, g/cm^3
Gray (I)	A	7.232
	B	7.224
Gray (II)	A	6.880
	B	6.938
Nodular	As-cast	7.113
	Annealed	7.056
	Normalized and tempered	7.084

Table 3 Volume fraction and particle size of graphite in gray and nodular cast irons

Type Cast Iron	Specimen	Volume Fraction Graphite	Mean Particle Size: Diameter, $\bar{d}$, mm	Mean Particle Size: Axial Ratio, $(\bar{d}/t)$
Gray (I)	A	0.082	0.120	14.6
	B	0.071	0.194	15.2
Gray (II)	A	0.127	0.14	23.9
	B	0.108	0.44	24.6
			D, mm	
Nodular	As-cast	0.092	0.050	
	Annealed	0.105	0.058	
	Normalized and Tempered	0.092	0.034	

As expected from its higher carbon content, gray cast iron II contained a larger volume fraction of graphite than gray cast iron I. The slightly higher volume fraction of graphite in the nodular iron in the annealed condition as compared with the as-cast or normalized and tempered conditions was probably related to slower cooling during the annealing treatment.

The mean diameter and mean axial ratio of the graphite flakes were calculated by the method of Fullman [12]. If the distribution of particle sizes is not large, and the shapes of the graphite flakes are approximated by disks, Fullman indicates that the mean diameter, $\bar{d}$, the mean thickness, $\bar{t}$, and the mean axial ratio, $(\overline{d/t})$, can be determined from measurements of the length and width of the intersections of the graphite flakes with a random-two-dimensional surface (the plane of polish). The values of $\bar{d}$ and $(\overline{d/t})$ for the gray cast irons are given in Table 3.

For nodular graphite, the mean diameter of the graphite nodules, $\bar{D}$, was calculated in a manner completely analogous to that used for graphite flakes, assuming that the nodules have a spherical shape [12]. The values of $\bar{D}$ for the nodular cast iron are given in Table 3.

Elastic Moduli Results. The elastic moduli of gray cast iron I and of the nodular cast iron were calculated from the formulas for propagation of elastic waves in thin rods [13]:

$$E = \rho V_l^2 \tag{1}$$

$$G = \rho V_s^2 \tag{2}$$

where E is the Young's modulus, G is the shear modulus, V_l is the velocity of longitudinal (dilatational) waves, V_s is the velocity of shear waves, and ρ is the density.

The elastic moduli of gray cast iron II, for which plate rather than rod specimens were used, were determined from the formulas for propagation of elastic waves in bulk solids [13]:

$$G = \rho V_s^2 \tag{3}$$

$$E = \frac{G(3\lambda + 2G)}{(G + \lambda)} \tag{4}$$

Table 4 Young's modulus of gray and nodular cast irons

Type Cast Iron	Specimen	Young's Modulus, 10^4 MPa (10^6 psi)
Gray (I)	A	12.8 (18.6)
	B	13.4 (19.5)
Gray (II)	A	9.72 (14.1)
	B	8.34 (12.1)
Nodular	As-cast	17.6 (25.6)
	Annealed	16.9 (24.5)
	Normalized and tempered	16.9 (24.6)

Table 5 Shear modulus of gray and nodular cast irons

Type Cast Iron	Specimen	Shear Modulus, 10^4 MPa (10^6 psi)
Gray (I)	A	5.16 (7.48)
	B	5.11 (7.42)
Gray (II)	A	3.93 (5.70)
	B	3.38 (4.90)
Nodular	As-cast	6.60 (9.58)
	Annealed	6.14 (9.44)
	Normalized and tempered	6.56 (9.52)

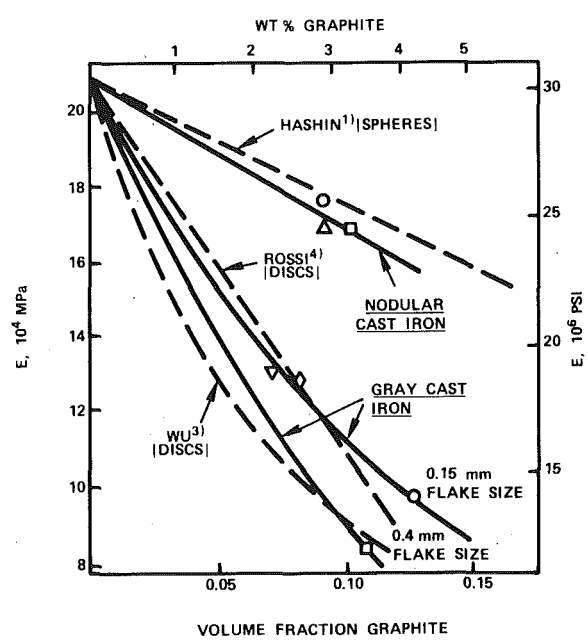


Fig. 4 Young's modulus of nodular and gray cast iron at 25°C

$$\lambda + 2G = \rho V_l^2 \tag{5}$$

where λ is the Lamé constant.

The values of E and G for the gray and nodular cast irons at room temperature are given in Tables 4 and 5, respectively. These values are also plotted versus graphite content in Figs. 4 and 5. The curves

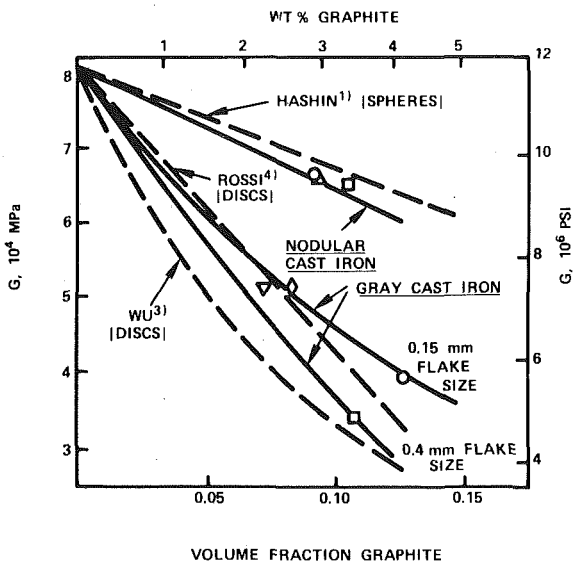


Fig. 5 Shear modulus of nodular and gray cast iron at 25°C

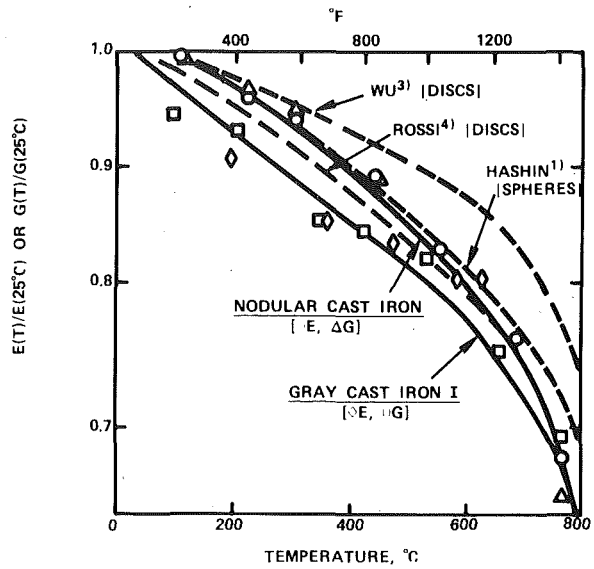


Fig. 6 Temperature-dependence of elastic moduli of nodular and gray cast iron

have been drawn through the data points so that they extrapolate to a value of Young's modulus of 20.9×10^4 MPa (30.4×10^6 psi) and a value of the shear modulus of 8.13×10^4 MPa (11.8×10^6 psi) typical of polycrystalline iron [15]. The elastic moduli of both the nodular and gray cast irons decreased with increasing graphite content. However, the rate of decrease was much greater in the gray cast irons than in the nodular cast iron. In addition to graphite content, graphite flake size also has a slight influence on the elastic moduli. Three of the gray iron specimens with an average flake size of 0.15 mm had uniformly higher elastic moduli (for a given volume fraction of graphite) than the fourth specimen which had a significantly larger flake size of 0.44 mm.

The temperature-dependence of the elastic moduli of specimen A of gray cast iron I and of the as-cast specimen of the nodular cast iron are shown in Fig. 6 in terms of the normalized values $E(T)/E(25°C)$ or $G(T)/G(25°C)$. The elastic moduli of both nodular and gray cast iron decreased in a similar manner with increasing temperature. The same values of E and G were obtained upon heating and cooling.

Discussion

The effect of graphite on the elastic moduli of nodular and gray cast irons can be understood in terms of theoretical models for the elastic moduli of two-phase solids [1–4]. The required input information is the elastic moduli of the matrix phase, the elastic moduli of the graphite particles, and the volume fraction and shape of the graphite particles.

The matrix phase in cast irons consists of ferrite, pearlite, or mixtures of both phases. Because the elastic moduli of cementite are essentially the same as those of pure iron [14], it is not expected that differences in the amount of pearlite in the matrix phase can significantly affect the elastic moduli. This conclusion is supported by the present results because normalizing and tempering, or annealing did not significantly change the elastic moduli of the nodular cast iron, even though the pearlite contents were changed by these heat treatments (Fig. 3). Similar results for gray cast iron have been reported by Plenard [6].

The ferrite and pearlite phases in cast iron usually contain up to several percent of alloying elements, principally silicon and manganese. However, a few percent of substitutional alloying elements can alter the elastic moduli of iron by only a few percent [7]. Because of the relatively minor effects of pearlite content and of substitutional alloying elements on the elastic moduli, only a small error results if the matrix phase of cast iron is considered to have the same values of Young's modulus and shear modulus as polycrystalline iron.

The second-phase graphite particles are polycrystalline in the case of nodular graphite but are nearly single crystals in the case of flake graphite [16]. Therefore, the elastic moduli for polycrystalline graphite, which are applicable for nodular graphite, would not be appropriate for flake graphite. However, because of the large difference in the elastic moduli of iron and of graphite, it is not believed that a significant error is introduced if both the nodules and flakes are assumed to have a Young's modulus of 0.85×10^4 MPa (1.2×10^6 psi) and a shear modulus of 0.33×10^4 MPa (0.48×10^6 psi) typical of polycrystalline graphite [15].

Because of the much lower elastic moduli of graphite, the elastic moduli of both gray and nodular cast iron decrease as the graphite content is increased, as shown in Figs. 4 and 5. However, the decrease in elastic moduli of gray cast iron, in which the graphite particles are disk-shaped (flakes), is much greater than the corresponding decrease in elastic moduli of nodular cast iron, in which the graphite particles are nearly spherical.

Nodular Cast Iron. Hashin [1] has given solutions for the elastic moduli of a two-phase solid which contains a random distribution of spherical particles. Hashin obtained these solutions by considering the changes in the strain energy in a loaded homogeneous body caused by the insertion of inhomogeneities. Two geometric approximations were made by Hashin:

1 The particles are spherical.
2 The action of the two-phase material on any one particle is transmitted through a spherical shell which lies wholly in the matrix phase.

Upper and lower bounds of the bulk and shear modulus of the two-phase solid were then calculated by Hashin using the principle of minimum strain energy. The upper and lower bounds of both the bulk and shear moduli of nodular cast iron with a low volume fraction of spherical graphite particles are nearly identical and can be approximated by the following expressions:

$$K = K_{Fe}\left[1 - \frac{3(1-\nu_{Fe})(1-K_c/K_{Fe})}{2(1-2\nu_{Fe}) + (1+\nu_{Fe})K_c/K_{Fe}} f_c\right] \quad (6)$$

$$G = G_{Fe}\left[1 - \frac{15(1-\nu_{Fe})(1-G_c/G_{Fe})}{(7-5\nu_{Fe}) + 2(4-5\nu_{Fe})G_c/G_{Fe}} f_c\right] \quad (7)$$

where K and G are the bulk and shear moduli of the cast iron, respectively, K_{Fe} and G_{Fe} are the bulk and shear moduli of the iron matrix, respectively, K_C and G_C are the bulk and shear moduli of the graphite particles, respectively, ν_{Fe} is the Poisson's ratio of the iron matrix, and f_c is the volume fraction of the graphite particles. To

calculate the Young's modulus of the nodular cast iron, equations (6) and (7) and the standard relationship between the isotropic elastic constants can be used [13]:

$$E = \frac{9KG}{(3K + G)} \tag{8}$$

Values of E and G for nodular cast iron were calculated as a function of the volume fraction of graphite from equations (6)–(8) using known values of the elastic moduli for pure iron and for graphite [15]. The values of E and G calculated from the Hashin theory were in excellent agreement with the observed values as shown in Figs. 4 and 5. Good agreement with the Hashin theory was also obtained for noduluar cast iron by Plenard [6].

Gray Cast Iron. For disk-shaped particles, the solutions must take into account the stress-concentration factor at the tip of the disk. Since the stress-concentration factor is dependent on the relative orientation between the stress axis and the plane of the disk, the problem is much more complicated than for the case of spheres, and solutions are only possible if some simplifying assumptions are made. In the theory of Wu [3], the strain field given by Eshelby [17] about an ellipsoidal inclusion was used, and the inclusions were assumed to be embedded in a matrix whose gross properties were identical to those of the composite. The final solution was then obtained by performing a spatial average over all possible disc orientations.

The equations given by Wu are quite complex but can be simplified for the case where the graphite particles, the iron matrix, and the cast iron composite all have a value of ν equal to 0.2. Although the actual values of ν for graphite, iron, and the cast iron composite are 0.27, 0.28, and 0.27, respectively [15], it is believed that this does not cause a significant error. The Wu equations then reduce to the simpler form

$$E = E_{\mathrm{Fe}} \frac{[2 - (1 - E_c/E_{\mathrm{Fe}})f_c]}{[2 - (1 - E_{\mathrm{Fe}}/E_c)f_c]} \tag{9}$$

with a similar expression for G. Using the known elastic moduli for iron and graphite [15] the values of E and G for various volume fractions of graphite were calculated from equation (9) and are given in Figs. 4 and 5, respectively. The agreement between the calculated and experimental values is reasonable considering that the theory is only approximate.

Rossi [4] has also derived expressions for the elastic moduli of a two-phase solid with small concentrations of disk-shaped particles. By analogy with equations (6) and (7), Rossi derives equations of the form

$$E = E_{\mathrm{Fe}}\,(1 - Mf_c) \tag{10}$$

where M is a stress-concentration factor and is determined by the axial ratio of the disk-shaped particle and by the ratio of the Young's modulus of the particle to that of the matrix. The value of M for randomly oriented particles is derived by assuming that the stress-concentration factor is primarily a function of the radius of curvature at the point of tangency of the particle with the stress axis. For the special case where ν has a value of 0.2 for the particle, matrix, and composite, Rossi derives values of M for various axial ratios of the particles and for various ratios of the Young's modulus of the particle to that of the matrix. For the case of gray cast iron where E_c/E_{Fe} has a value of 0.039, and where $(\overline{d/t})$ has a value of about 20, M has a value of 4.8.

The values of E calculated from equation (10) for various volume fractions of graphite are given in Fig. 4. The corresponding values of G which were calculated from the values of E by using a value of ν equal to 0.2 and the standard relationship

$$G = \frac{E}{2(1 + \nu)} \tag{11}$$

are given in Fig. 5. The values of E and G from the Rossi solution are in slightly better agreement with the observed values of E and G for gray cast iron than those calculated from the Wu solution. However, both the Wu and Rossi solutions clearly indicate that, for a similar volume fraction of particles, disk-shaped graphite particles lower the elastic moduli of the cast iron much more than spherical graphite particles.

Neither the Wu nor the Rossi solutions for disk-shaped particles predict an effect of particle size on the elastic moduli. Yet, several studies on gray cast iron have shown that larger graphite flakes produce lower elastic moduli than smaller graphite flakes [6, 18]. A similar effect was observed in this investigation. As shown in Figs. 4 and 5, the elastic moduli for three gray cast iron specimens having an average flake size of 0.15 mm fall on a curve corresponding to higher elastic moduli than those for gray cast iron having a flake size of 0.44 mm. One possible explanation of the effect of particle size is that the stress fields of individual particles begin to overlap and reinforce each other at larger particle sizes, resulting in lower elastic moduli but this problem has not been treated theoretically. The effect of particle size is small, however, compared with the large effects of particle volume fraction and particle shape.

Temperature-Dependence. The temperature-dependence of the elastic moduli of nodular and gray cast iron can be interpreted in terms of the temperature-dependence of the elastic moduli of iron and of graphite by using equations (6) and (7) for nodular iron and either equation (9) or (10) for gray cast iron. Changes in f_c with temperature that might result from decomposition of cementite or changes in $(\overline{d/t})$ that might result from Ostwald ripening of the graphite particles are small and can be neglected. Also, the temperature dependence of ν_{Fe} [19], and of K_c or G_c [20] is small and can be ignored. As a result, the only temperature-dependent terms in equations (6), (7), and (9) are K_{Fe}, G_{Fe}, and E_{Fe}. In equation (10), the value of M depends on the ratio E_c/E_{Fe} and therefore is temperature-dependent.

The values of E_{Fe} at temperatures between ambient temperature and 1000°C have been determined by Köster [21]. From these values and a constant value of ν_{Fe} equal to 0.28, values of K_{Fe} and G_{Fe} were calculated at various temperatures by using equations (8) and (11). These values were then used in equations (6)–(10) to calculate values of E at various temperatures for the case of spherical or disk-shaped graphite particles. The normalized values $E(T)/E(25°\mathrm{C})$ are shown in Fig. 6. Because of the assumption of constant ν, and the relationship given in equation (11), the calculated temperature dependences of G and E are identical.

The calculated value of $E(T)/E(25°\mathrm{C})$ from the Hashin theory for spherical graphite particles and from the Rossi theory for disk-shaped graphite particles agree quite closely with the observed values of $E(T)/E(25°\mathrm{C})$ for nodular and gray cast iron, respectively. However, the calculated values of $E(T)/E(25°\mathrm{C})$ from the Wu theory for disk-shaped graphite particles are considerably higher than the observed values of $E(T)/E(25°\mathrm{C})$ for gray cast iron.

Conclusions

The elastic moduli of both nodular and gray cast iron were measured at ambient and elevated temperatures up to 760°C. These results were then analyzed by using the available theories for the elastic moduli of two-phase solids and the measured size, shape, and volume fraction of graphite particles. The following conclusions were reached:

1 The elastic moduli of both nodular and gray cast iron decrease when the graphite content is increased. However, the decrease is much greater for gray cast iron in which the graphite particles are disk-shaped (flakes) than for nodular cast iron in which the graphite particles are spherical.

2 The effect of graphite on the elastic moduli of both nodular and gray cast irons is satisfactorily predicted by existing theories for the elastic moduli of two-phase solids when both the volume fraction and shape of the graphite particles are considered.

3 Graphite flake size also has a small effect on the elastic moduli, larger graphite flakes resulting in lower elastic moduli. The size effect is not predicted by existing theories.

4 The temperature-dependence of the elastic moduli of both nodular and gray cast iron can be correctly predicted from the existing theories for the elastic moduli of two-phase solids and the known temperature-dependence of the elastic moduli of the polycrystalline iron matrix and of the graphite particles.

References

1 Hashin, Z., ASME JOURNAL OF APPLIED MECHANICS, Vol. 29, 1962, pp. 43–50.

2 Wu, T. T., USAEC AT (30–1)—2394, Technical Report No. 10, Brown University, 1964.

3 Wu, T. T., *International Journal of Powder Metallurgy,* 1965, Vol. 1., pp. 8–12.

4 Rossi, R. C., *Journal of the American Ceramic Society,* 1968, Vol. 51, pp. 433–439.

5 Budiansky, B., and O'Connel, R. J., *International Journal of Solids and Structures,* 1976, Vol. 12, pp. 81–97.

6 Plenard, E., *Recent Research on Cast Iron,* Gordon and Breach, New York, 1968, pp. 707–792.

7. Speich, G. R., Schwoeble, A. J., and Leslie, W. C., *Met. Trans.,* 1972, Vol. 3, pp. 2031–2037.

8 Papadakis, E. P., *Journal of the Acoustical Society of America,* 1967, Vol. 42, pp. 1045–1051.

9 Garafalo, F., and Wriedt, H. A., *Acta Metall.,* 1962, Vol. 10, pp. 1007–1012.

10 Nix, F. C., and McNair, D., *Physical Review,* 1941, Vol. 60, pp. 596–605.

11 Dehoff, R. T., and Rhines, F. N., *Quantitative Microscopy,* McGraw-Hill, New York, 1968.

12 Fullman R., *Trans. Am. Inst. Min Metall. Eng.,* 1953, Vol. 197, pp. 447–52.

13 Kolsky, K., *Stress Waves in Solids,* Clarendon Press, London, 1953.

14 Laszlo, F., and Nolle, A., *Journal of the Mechanics and Physics of Solids,* 1959, Vol. 7, pp. 193–208.

15 Gschneider, K. A., *Solid State Physics,* 1964, Vol. 16, pp. 275–307.

16 Oron, M., and Minkoff, I., *Recent Research on Cast Iron,* Gordon and Breach, New York, 1968, pp. 173–194.

17 Eshelby, J. D., *Proceedings of the Royal Society,* London, 1957, Series A, Vol. 241, p. 376.

18 Sisco, F. T., *The Alloys of Iron and Carbon,* Vol. II, Properties, McGraw-Hill, New York, 1937, pp. 291–295.

19 Koster, W., *Metallurgical Reviews,* 1961, Vol. 6, pp. 1–56.

20 Faris, F. E., Green, L., and Smith, C. A., *Journal of Applied Physics,* 1952, Vol. 23, pp. 89–95.

21 Köster, W., *Arch. Eisenhuttenwes,* 1940, Vol. 6, pp. 271–278.

Y. A. Bahei-El-Din[1]
Visiting Assistant Professor,
Department of Civil Engineering,
Duke University,
Durham, N.C.

G. J. Dvorak
Professor and Chairman,
Department of Civil Engineering,
University of Utah,
Salt Lake City, Utah 84112
Mem. ASME

Plastic Deformation of a Laminated Plate With a Hole

The elastic-plastic behavior of a (0/90) symmetric FP-Al plate containing a circular hole is investigated using the finite-element method. Of principal concern are plastic yielding at the circular hole and fracture of the plate caused by failure of the fiber at the hole. The results illustrate the significance of plasticity in deformation of metal-matrix composites. The behavior of the laminated plate is compared with that of a geometrically similar plate made of an unreinforced matrix material, for uniaxial loading/unloading/reloading sequences. The comparison reveals significant differences in the role of plastic deformation in these two materials. Specifically, plastic yielding in the matrix of the laminated plate at the circular hole leads to a substantial increase in the local stress concentration in the elastic fibers adjacent to the hole boundary. Also, it is found that the fiber reinforcement causes a large increase in the stiffness and strength of the composite, but only a minor elevation of its yield strength. Therefore, to take advantage of the mechanical properties of the metal matrix composite material, it is necessary to admit working loads which exceed its elastic limit.

Introduction

Metal-matrix composites reinforced by continuous fibers may exhibit an appreciable amount of elastic-plastic deformation depending on the state of stress and temperature. Since most of the practically used fibers, such a boron, graphite, silicon carbide, and the FP fiber, remain elastic until failure, the inelastic component of the overall deformation is caused by plastic flow of the matrix. Although the fibers strengthen the matrix substantially, and are the principal source of composite stiffness, they have a relatively small effect on the overall stress level which causes the onset of plastic flow. Indeed, the presence of the reinforcing fibers may be the very cause of plastic deformation, as in the case of heat-treatment of unidirectional materials [1]. In any event, to take advantage of the high strength of the metal-matrix composite materials, it is necessary to admit working loads which exceed its elastic limit.

The significance of plasticity in deformation behavior of metal-matrix composites can be illustrated by the following evaluation of matrix plastic strain magnitudes in typical material systems. Consider a unidirectional specimen loaded by simple tension in the fiber direction. We denote the maximum fiber strain at failure as ϵ_f^*, the corresponding normal plastic strain in the matrix as $\epsilon_m{}^p$, and the elastic part of the total strain in the matrix as $\epsilon_m{}^e$. Compatibility in the fiber direction requires that at fiber failure $\epsilon_f^* = \epsilon_m{}^e + \epsilon_m{}^p$. The magnitudes of $\epsilon_m{}^p$, derived from this simple failure criterion, for a group of aluminum-matrix composites reinforced by different fibers are shown in Table 1.

Table 1 Matrix plastic strain magnitudes in aluminum-matrix composites

Fiber	Young's modulus MPa	Strength MPa	Failure strain ϵ_f^*	Elastic strain in Al-matrix[1] $\epsilon_m{}^e$	Plastic strain in Al-matrix $\epsilon_m{}^p$
Boron	400	3.4	0.0085	0.0010	0.0075
FP	350	1.4	0.0040	0.0008	0.0032
Graphite	250	3.4	0.0136	0.0012	0.0124
Silicon Carbide	428	3.1	0.0072	0.0009	0.0063

[1] Derived from Fig. 2.

The axial loading of a unidirectional lamina is clearly the most restrictive case, much larger plastic strains may develop in other than axial directions, e.g., in the case of interlaminar shear of laminated plates in bending, or at free edges of laminated structures. Yet, the plastic strain magnitudes which can be attained in the matrix prior to fiber failure are about one order of magnitude larger than the elastic strain magnitudes. This suggests that it is necessary to consider the existence of axial plastic strains in the development of constitutive equations for fibrous composites.

These aspects of plastic deformation of fibrous composites have not been accounted for in the earlier formulations of elastic-plastic

[1] On leave from the Division of Structures, Department of Civil Engineering, Cairo University, Giza, Egypt.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, February, 1980; final revision, May, 1980.

constitutive equations which adopted the assumption of purely elastic or even rigid response of the material in the axial direction [2–5]. Of course, in applications to laminated plates and other structures this assumption implies elastic or rigid behavior in each fiber direction, and thus severely restricts practical usefulness of the theory.

A constitutive theory for elastic-plastic deformation of unidirectional fibrous composites has been recently developed in a six-dimensional stress space [6–8]. In order to analyze composites with complex geometries and loading regimes, the new constitutive equations were incorporated in a finite-element scheme [7, 9]. In this paper, we use the finite-element code to examine the behavior of a FP-Al crossply plate with a circular hole. First, we briefly describe the new elastic-plastic material model of the composite and the finite-element routine, the details of which are discussed elsewhere [6–9]. Next, we present the results obtained for the laminated plate with a hole and illustrate the significance of the matrix plastic yielding in metal-matrix composites. Finally, the behavior of the laminated plate is compared with that of a geometrically similar plate made of an unreinforced matrix materials, for uniaxial loading/unloading/reloading sequences.

Material Model

New constitutive equations for the elastic-plastic behavior of unidirectional composites have been recently derived from a composite model [6–8]. The material model represents the mechanical properties of the constituents, and the essential features of their mutual constraint. Specifically, the composite lamina is modeled as an elastic-plastic matrix with a unidirectional constraint imposed by the fiber. Any stable matrix material may be used in this model. A Mises-type matrix exhibiting kinematic hardening has been used for the present work to represent the metal matrix, and an elastic brittle fiber to represent the reinforcement.

In the material model, each of the fibers is assumed to be of very small diameter, so that although the fibers occupy a finite volume fraction of the composite they do not interfere with matrix deformation in the transverse and longitudinal directions. As a result, the transverse tension and shear as well as longitudinal shear response of the composite are derived from that of the matrix, except when there is an axial prestrain which causes coupling of axial and transverse plastic strain components. Fig. 1 shows a schematic drawing of the material model. There, σ_{ij} and ϵ_{ij} represent the stress and strain tensors, respectively, v_f and v_m are the volume fractions of fiber and matrix such that $v_f + v_m = 1$. Stress and strain components with top bar are overall composite quantities, while superscripts f and m denote fiber and matrix, respectively. The model can be represented by parallel fiber and matrix bars or plates with axial coupling and is similar to the model used for nonlinear viscoelastic problems in composites by Lou and Schapery [10].

Since the fibers are elastic up to failure, the inelastic strains of the lamina are caused by matrix deformation. The elastic constraint imposed on the matrix by the fiber affects the shape of the lamina yield surface, it leads to additional kinematic components in the hardening rule of the lamina, and it has an influence on the magnitude of overall plastic strains. All aspects of this behavior are examined and accounted for in the formulation of the lamina constitutive equations which are explicitly described in [7, 8]. The constitutive equations are generated in matrix form and, therefore, can be readily incorporated in a numerical scheme.

Finite-Element Analysis of Composites

A finite-element code for elastic-plastic analysis of fibrous composite structures has been developed [7, 9] using the new constitutive equations of the composite lamina. The program, named PAC78, is based on the finite-element procedure in conjunction with the displacement method of analysis. The nonlinearities caused by the elastic-plastic behavior of the composite are handled by a modified Newton-Raphson iteration procedure.

The laminated structure of most fibrous composites makes it necessary to conduct a three-dimensional analysis even for in-plane loads.

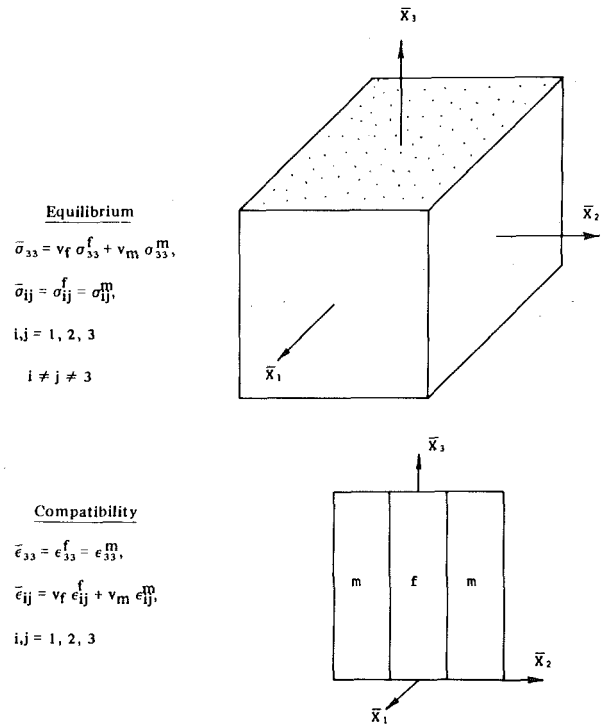


Fig. 1 Material model of the elastic-plastic lamina

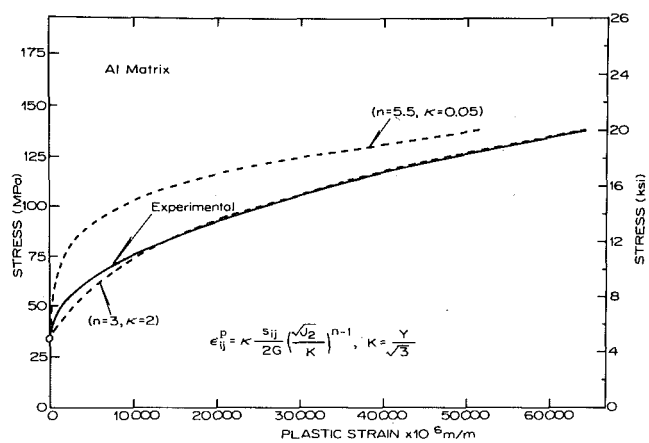


Fig. 2 Ramberg-Osgood approximations of matrix stress-strain curves

The out-of-plane stresses, e.g., interlaminar shear and normal stresses, are significant in evaluating the behavior of laminates. In the present finite-element analysis, the eight-node hexahedral element has been used. The element represents a unidirectional composite material whose fiber is arbitrarily orientated in the structural coordinates. Laminated structures with layers in several directions are built up by joining together individual unidirectional elements.

Each unidirectional composite element is treated as macroscopically homogeneous and anisotropic. The elastic response, initial yielding, and kinematic hardening and flow rules of the composite are derived from the elastic-plastic material model described earlier. The computational aspects of the PAC78 program together with its features and capabilities are described in [7, 9].

Laminated Plate With a Circular Hole

The behavior of a (0/90) symmetric laminated plate containing a circular hole has been investigated using the PAC78 finite-element code. The plate is made of a 6061 aluminum matrix reinforced by 50

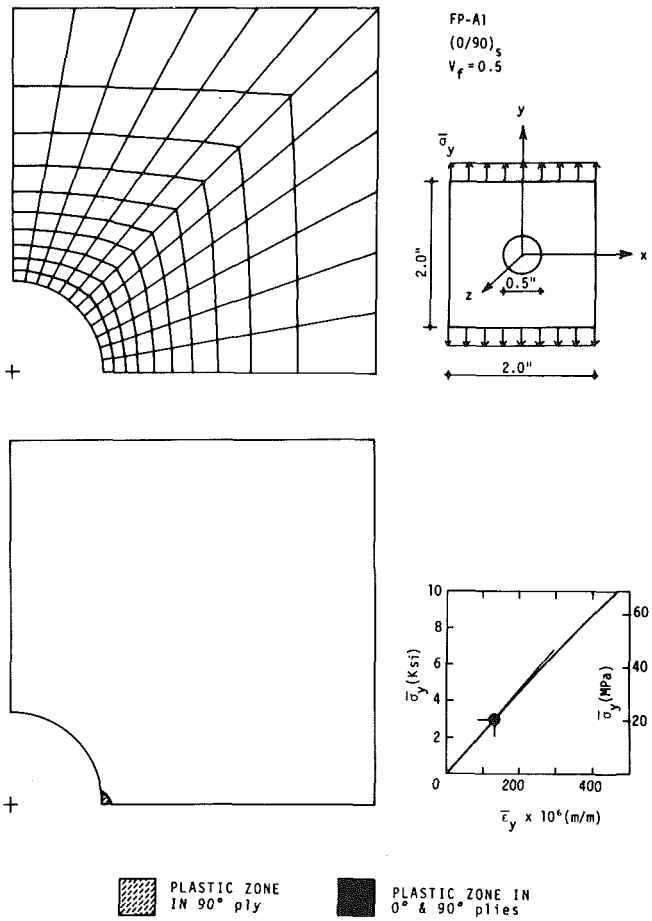


Fig. 3 Geometry and initial yielding of a crossply FP-Al plate with a hole

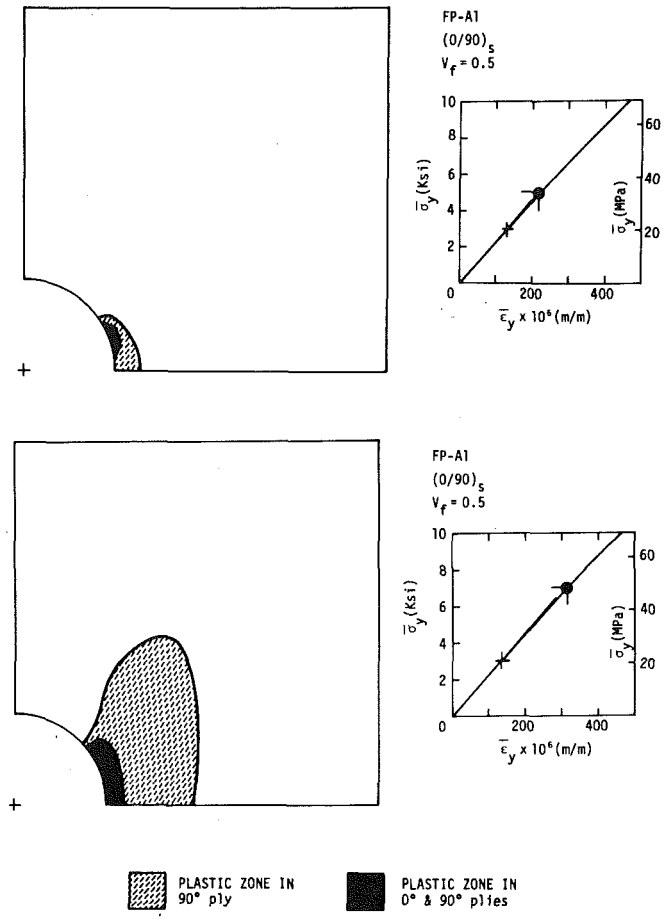


Fig. 4 Plastic zones in a crossply FP-Al plate with a hole

volume percent of FP fibers. The material in Fig. 2, with Ramberg-Osgood parameters $n = 5.5$, $\kappa = 0.05$ was used to define the properties of the aluminum matrix. Although the particular choice of the parameters does not appear to agree well with the experimental response of the matrix, it provides satisfactory approximation to the behavior of a number of laminated plates, as indicated in references [6–8]. The reasons for this discrepancy are believed to be related to the effect of fibers on matrix hardening in the transverse plane, to *in situ* hardening characteristics of the matrix, and to other factors not accounted for in the material model.

The plate is loaded in uniaxial in-plane tension. A loading/unloading/reloading sequence is employed. The analyzed specimen is 2 in. × 2 in. and 0.4 in. thick, and contains a 0.5 in. hole. The finite-element mesh used in the analysis is shown in Fig. 3. Each layer of the plate is represented by one layer of unidirectional composite elements. The fiber direction in the outside layers (0-deg layers) coincides with the loading direction, while the fiber direction in the inside layers (90-deg layers) is perpendicular to the loading direction. We assume that the plate is free from initial stresses. This may be difficult to achieve in practice, because of cooling from fabrication temperatures. However, the effect of initial stresses will be quite small after more extensive plastic deformation of the plate occurs.

First yielding is observed relatively early in the loading process at about 21 MPa (3 ksi). This takes place in the 90-deg layers at the edge of the circular hole, Fig. 3. Figs. 4 and 5 show the development of plastic zones as loading continues. The plastic zone in the 90-deg layers spreads rapidly. We note that the 90-deg layers have yielded almost completely at about 69 MPa (10 ksi). On the other hand, yielding in the 0-deg layers is somewhat delayed, and when it develops the plastic zone tends to remain small and contained to the immediate vicinity of the hole.

We have unloaded the specimen from 69 MPa (10 ksi) to zero applied load and reloaded the specimen to an overall stress of 166 MPa (24 ksi). Fig. 6 shows the continued development of plastic zones in the layers of the plate for different loading levels up to and including the maximum load of 166 MPa. The plastic zone in the 0-deg layers spreads upward, in the loading direction. At 166 MPa, Fig. 6, all layers of the plate have yielded almost completely.

Fig. 7 shows the distribution of local σ_{yy} stress along the x-axis for different overall stress levels. We note here that by disregarding plastic deformation, the stress concentration at the edge of the hole at maximum applied load will be underestimated by about 25 percent for the 0-deg layers, and overestimated by over 50 percent for the 90-deg layers. Accordingly, the overall strength of the plate will be overestimated. This is particularly obvious from Fig. 8 which shows the axial stress in the fibers of the 0-deg layers at the edge of the hole along the x-axis as a function of applied load $\bar{\sigma}_y$. The curve in Fig. 8 is quite nonlinear due to plastic deformation of the composite. According to our analysis, failure occurs at the hole under overall stress of about 166 MPa (24ksi) due to failure of the fibers in the 0-deg layers. At this stress level, the FP fibers attain their ultimate tensile strength of about 1400 MPa (200 ksi). If one disregards plastic deformation of the composite, the overall stress at failure is estimated as 214 MPa (31 ksi), an overestimation of about 30 percent. If the plates were reinforced by boron fibers, which have ultimate strength double that of the FP fibers, the errors encountered in an elastic analysis would be even larger.

Fig. 8 indicates that plastic yielding in the matrix of the laminated plate at the circular hole leads to a substantial increase in the local stress concentration in the elastic fibers adjacent to the hole boundary.

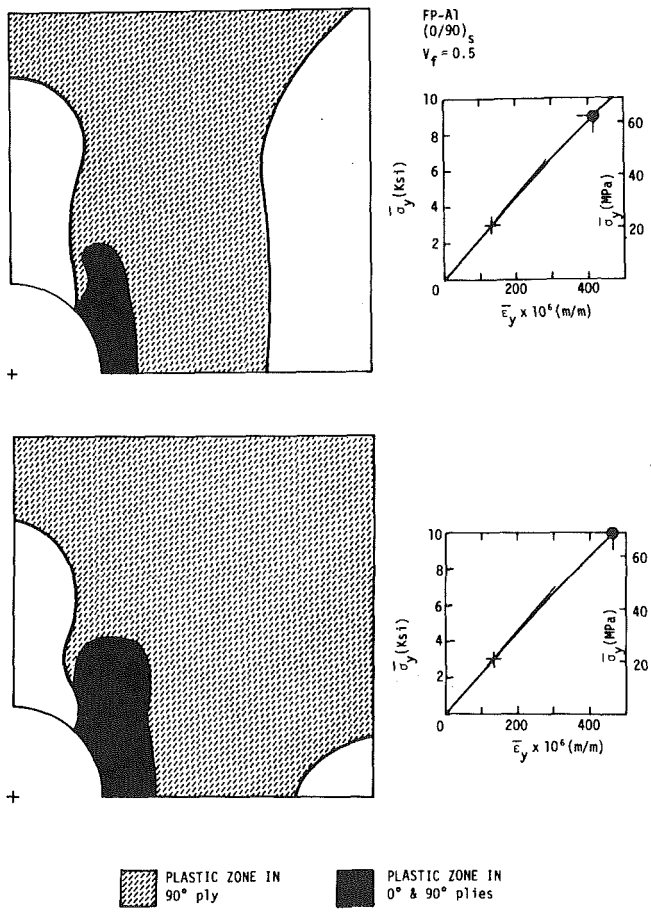


Fig. 5 Plastic zones in a crossply FP-Al plate with a hole

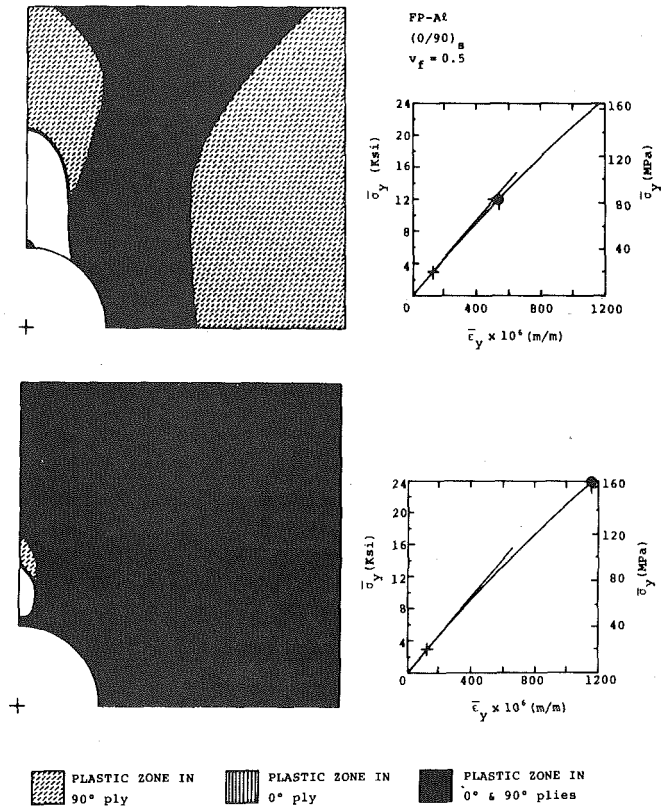


Fig. 6 Plastic zones in a crossply FP-Al plate with a hole

This result is not unexpected; analogous conclusions will probably be reached in studies of other large imperfections in laminated metal matrix structures. However, it is in contrast with the reduction of stress concentration caused by localized yielding in unreinforced plates.

To illustrate the differences in plastic yielding at circular holes and in overall behavior of reinforced and unreinforced plates, we analyzed a geometrically similar aluminum plate made of the matrix material used in the composite plate. Comparison of the development of plastic zones in the aluminum plate and in the $(0/90)_s$ plate is shown in Fig. 9 at different levels of overall stress normalized with respect to the yield stress of the aluminum matrix. As expected, one finds that the unreinforced plate yields at lower loads and experiences more extensive yielding. We note that the unreinforced plate yields almost completely at $\bar{\sigma}_y/Y = 1.0$ ($Y = 34.5$ MPa (5 ksi)). At the same level of loading, the plastic zones in all layers of the reinforced plate are very small and contained. However, the yield pattern of the 90-deg layers in the laminated plate is somewhat similar to that in the aluminum plate. In fact, the initial yield stresses of the two plates are relatively close (21 MPa and 12.5 MPa).

The overall stress-strain response of both the $(0/90)_s$ plate and the aluminum plate is shown in Fig. 10. The gain in stiffness due to the fiber reinforcement is obvious. The aluminum plate has lost almost 85 percent of its elastic stiffness at about 35 MPa (5 ksi) and is expected to fail at a somewhat higher load. The stiffness of the reinforced plate, on the other hand, does not change much in the initial part of loading. At failure, $\bar{\sigma}_y = 166$ MPa (24 ksi), the initial stiffness of the reinforced plate has been reduced by about 30 percent. It is also seen from Fig. 10 that plastic straining takes place in the laminated plate during unloading, which has been also observed in experiments on laminated plates [11].

It is of interest to note in passing the implication of the existence of different plastic zones in the individual laminate layers. This

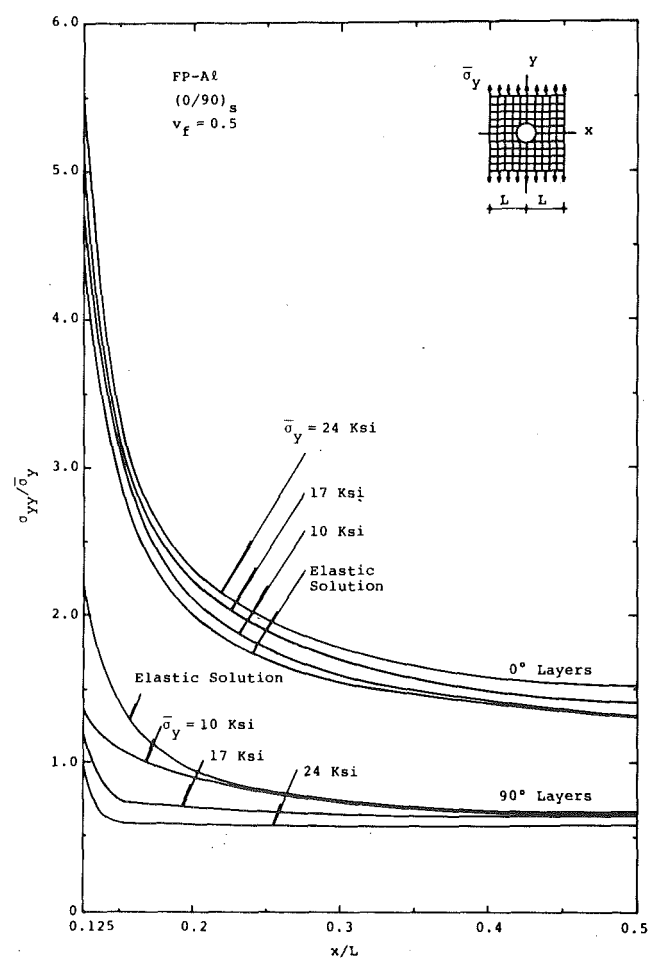


Fig. 7 Stress distribution in a crossply FP-Al plate with a hole

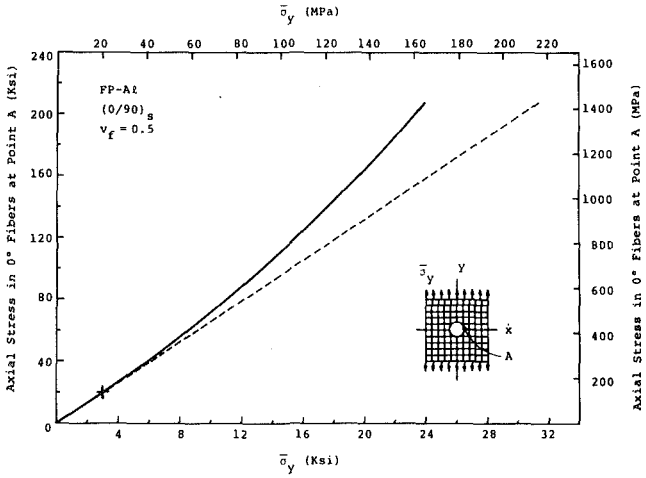


Fig. 8 Axial stress in 0° fibers at the edge of a hole in crossply FP-Al plate

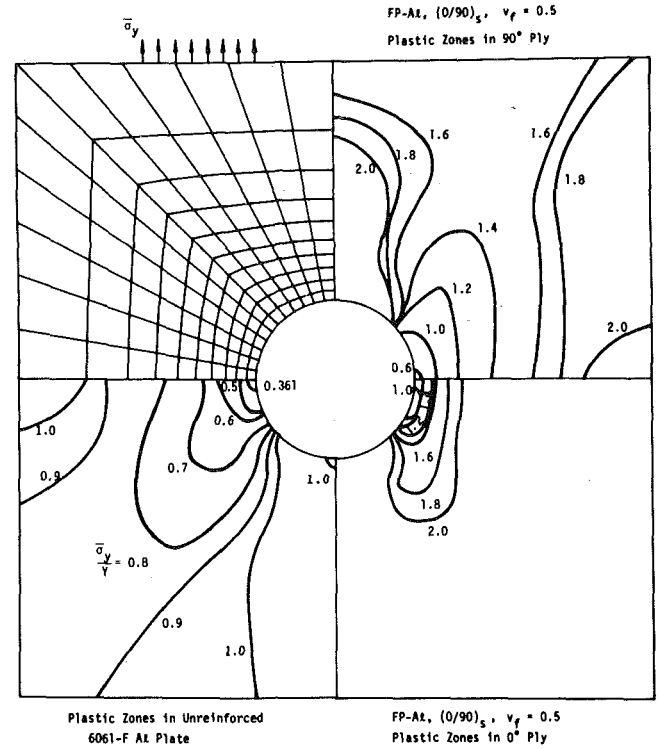


Fig. 9 Comparison of the development of plastic zones around a hole in reinforced and unreinforced plates; numbers indicate the ratio of current load to matrix yield stress, $\bar{\sigma}_y/Y$, (Y = 34.5 MPa)

suggests that the loading point in the local stress space does not coincide with the corner of the local yield surface. As shown in references [6–8], the overall yield surface of the crossply plate has two elliptical branches which intersect at four corners. The loading point resides on the 90-deg layer branch at all points in the plate domain where the 90-deg layer has yielded before the 0-deg layer. The loading point coincides with the corner in the overall yield surface at all locations where both layers experience plastic straining.

Conclusions

It is evident that plastic yielding of the matrix is the dominant mode of deformation in most applications of metal-matrix composites. Elastic behavior is of limited significance in applications which hope to utilize the high strength of the composite. A simple calculation reveals that plastic strain magnitudes in the matrix are about one order of magnitude larger than elastic strain magnitudes before failure of the fiber in unidirectional composites.

The significance of plasticity in fibrous composites has been also demonstrated for laminated plates. Misleading results and predictions of stiffness and strength of laminated plates may occur if plastic deformation of the composite is disregarded. It has been shown that the strength of a FP-Al crossply plate containing a hole can be overestimated by about 30 percent if one does not account for plastic deformation of the matrix. Since plastic deformation in aluminum matrix composites reinforced by the FP fiber is less pronounced than in those reinforced by other, stronger and/or more compliant fibers (c.f. Table 1), even greater errors may be experienced in composites reinforced by boron or SiC fibers.

Plastic straining of the matrix has been also shown to take place during cyclic loading which is important in studying fatigue and local damage in composites [12].

Acknowledgments

This work has been supported by the U.S. Army Research Office.

References

1 Dvorak, G. J., and Rao, M. S. M., "Thermal Stresses in Heat-Treated Fibrous Composites," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, 1976, pp. 619–624.

2 Mulhern, J. F., Rogers, T. G., and Spencer, A. J. M., "A Continuum Model for Fiber-Reinforced Plastic Materials," *Proceedings of the Royal Society*, London, Vol. 301, 1967, pp. 473–492.

3 Mulhern, J. F., Rogers, T. G., and Spencer, A. J. M., "A Continuum Theory of a Plastic-Elastic Fiber-Reinforced Material," *International Journal of Engineering Science*, Vol. 7, 1969, pp. 129–152.

4 Prager, W., "Plastic Failure of Fiber-Reinforced Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, 1969, pp. 542–544.

5 Hashin, Z., Bagchi, D., and Rosen, B. W., "Nonlinear Behavior of Fiber Composite Laminates," NASA CR-2313, 1974.

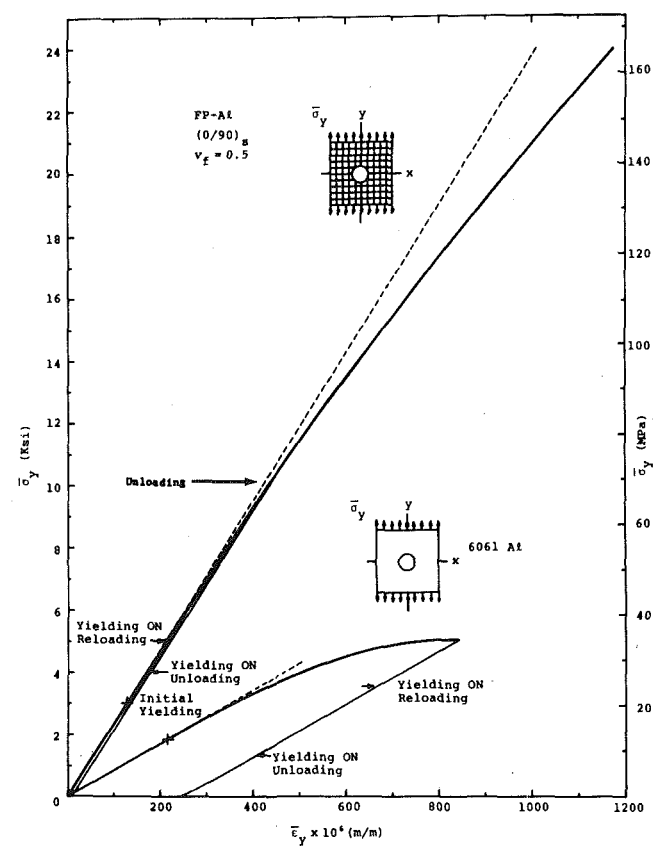


Fig. 10 Stress-strain curves of a crossply FP-Al plate and an unreinforced Al plate

6 Dvorak, G. J., and Bahei-El-Din, Y. A., "Plasticity of Composite Laminates," *Proceedings of the Research Workshop on Mechanics of Composite Materials,* Dvorak, G. J., ed., Duke University, Oct. 1978, pp. 32–54.

7 Bahei-El-Din, Y. A., "Plastic Analysis of Metal Matrix Composite Laminates," PhD Dissertation, Duke University, 1979.

8 Bahei-El-Din, Y. A., and Dvorak, G. J., "Plastic Yielding at a Circular Hole in a Laminated FP-Al Plate," *Modern Developments in Composite Materials and Structures,* Vinson, J. R., ed., The American Society of Mechanical Engineers, 1979, pp. 123–147.

9 Bahei-El-Din, Y. A., Dvorak, G. J., and Utku, S., "Finite-Element Analysis of Elastic-Plastic Fibrous Composite Structures," to be presented at the George Washington University and NASA Langley Research Center Symposium on Computational Methods in Nonlinear Structural and Solid Mechanics, October 6–8, 1980; to appear in *Computers and Structures.*

10 Lou, Y. C., and Schapery, R. A., "Viscoelastic Characterization of Nonlinear Fiber-Reinforced Plastics," *Journal of Composite Materials,* Vol. 5, 1971, pp. 208–234.

11 Sova, J. A., and Poe, C. C., Jr., "Tensile Stress-Strain Behavior of Boron/Aluminum Laminates," NASA TP-1117, 1978.

12 Dvorak, G. J., and Johnson, W. S., "Fatigue of Metal Matrix Composites," to appear in a special issue on Fatigue of Composites and Nonmetals, *International Journal of Fracture.*

R. L. Weaver
Research Associate.

Yih-Hsing Pao
Professor.
Mem. ASME

Department of Theoretical and Applied Mechanics,
Cornell University,
Ithaca, N. Y. 14853

Multiple Scattering of Waves in Irregularly Laminated Composites

The transition matrix formulation of multiple scattering is applied to the problem of wave propagation in a one-dimensional layered medium. The effect of geometrical irregularity in an otherwise periodic layered structure is investigated in detail for the case of elastic waves propagating normally to elastic layers embedded in elastic, or in viscoelastic matrix media. The irregularity is found to widen and diminish the stop bands and soften the sharp band features characteristic of a fully periodic structure, and to generate scattering losses with a consequent increase in the attenuation of the coherent wave field.

1 Introduction

The propagation of elastic waves normal to a laminate composite which is modeled by an infinite periodic array of solid layers has been analyzed by applying variational techniques [1, 2], and by use of the transfer matrix [3]. Mathematically, the problem is similar to the propagation of one-dimensional waves through an infinite periodic array of potentials [4]. It illustrates in a simple manner the phenomenon of band structure, which in three dimensions is of great interest to the understanding of the physics of crystals and semiconductors. But the methods which have been so successful in treating the wave mechanics of perfect, crystalline structures are not so effective when applied to quasi-crystalline arrays. The problem is analogous to a composite of an infinite array of layers with irregularity in its otherwise periodic spacing.

For fibers randomly and sparsely distributed in an elastic matrix medium, the theory of multiple scattering of elastic waves has been applied to analyze the dispersion of waves in such a medium [5, 6]. Results in [6] showed that the waves are not only dispersed but also attenuated, especially at higher frequencies. The attenuation is presumably caused by the random scattering of waves by the fibers out of the average, or coherent, field. This loss mechanism is generally known as scattering losses, to distinguish it from losses due to viscous processes [7].

Many composite materials are made of alternating layers of different lamina. Each layer is considered here to be effectively homogeneous, i.e., the structure of any inhomogeneity in an individual layer is assumed to be much finer than the wavelengths to be treated. The lamina are all nearly equal in thickness when compressed together. Scattering losses occur in such a structure because of the nonuniform thicknesses of the lamina. This will result in attenuation and dispersion of waves propagating through such a medium. If the propagation vector is normal to the laminate, the problem is one-dimensional and it is analogous to the aforementioned problem of one-dimensional waves in a quasi-periodic potential.

Imperfections in such a periodic structure can be either substitutional (materials of the layers) or geometrical (spacing between the layers). Both cases have been treated previously by the method of transfer matrix [8, 9], but only for elastic materials, and only for an unusual distribution of lamina thickness [8]. In this paper, the theory of multiple scattering is applied to investigate the dispersion and attenuation of waves in a laminated composite with imperfect geometry, and with viscoelastic matrix materials.

The theory and the concept of the configurational average along with the method of the transition matrix as presented in reference [6] using the Lax quasi-crystalline assumption are summarized in the next section.

Following that, we develop a one-dimensional polar coordinate system, and the solution of the one-dimensional Helmholtz equation in such a coordinate system. The purpose of this development is to render the method of transition matrix which was developed by Waterman [10] primarily for two and three-dimensional scattering of waves readily applicable to one-dimensional problems. A simplified derivation of the transition matrix was given later by Pao [11]. A Gaussian probability distribution function is then assumed to represent a one-dimensional periodic or nearly periodic (quasi-regular) layered structure. Attenuation and dispersion of waves are then determined through the configurational average of the total field.

Section 3 presents numerical results and compares the dispersion and attenuation of the field for periodic layers, and of the coherent, average, field for the case of nearly periodic layers. Attenuation in a perfectly regular periodic composite manifests in the form of stop bands in certain frequency intervals. True scattering losses occur only for layers with nonuniform properties or thicknesses.

In the last section we assume the material for the matrix is visco-

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, November, 1979; final revision, May, 1980. Paper No. 80-WA/APM-22.

elastic, and then compare the scattering losses with the losses due to viscosity.

2 Multiple Scattering in One Dimension

The problem of one-dimensional waves in periodically layered media can be treated by a closed-form analysis when the periodicity is exact. If, however, there exists slight irregularity in the periodic spacing or in the material properties, we find that it is necessary to apply the theory of multiple scattering of waves to determine the average wave quantities. At the beginning of this section, we briefly state the theory of multiple scattering using the transition matrix as given in reference [6]. We then develop one-dimensional wave functions, and apply them toward the solutions of multiple scattering of scalar waves in periodic layers with small irregularities.

Transition Matrix Formulation of Multiple Scattering. Consider an infinite number of arbitrarily shaped and constituted scatterers embedded in the right-hand side of an infinitely extended space filled with matrix material. For a steady-state wave field at frequency ω, the matrix is characterized by wave number k, not necessarily real, and the spatial part of the wave field in the matrix material satisfies the Helmholtz equation,

$$\nabla^2\phi(\mathbf{r}) + k^2\phi(\mathbf{r}) = 0 \tag{1}$$

Under the excitation of an externally applied field $\phi^A(\mathbf{r})$ the total field at any point, $\mathbf{r}$, in the matrix material may be written

$$\phi(\mathbf{r}) = \phi^A(\mathbf{r}) + \sum_j^N \phi_j^S(\mathbf{r}) \tag{2}$$

where ϕ_j^S is the wave scattered by the inclusion at $\mathbf{r}_j$ and the summation over j is from 1 to the total number of scatterers, N. The total field exciting the ith scatterer is obtained from $\phi(\mathbf{r})$ by omitting the field scattered from itself.

$$\begin{aligned}\phi_i^E(\mathbf{r}) &= \phi(\mathbf{r}) - \phi_i^S(\mathbf{r}) \\ &= \phi^A(\mathbf{r}) + \sum_{j\neq i}^N \phi_j^S(\mathbf{r})\end{aligned} \tag{3}$$

where the summation from $j = 1$ to N excludes $j = i$.

The applied field can be expanded in an infinite series of basis functions $\hat{\psi}_n(\mathbf{r} - \mathbf{r}_i)$ which are linearly independent solutions of equation (1), and are regular at $\mathbf{r}_i$ [10, 11],

$$\phi^A(\mathbf{r}) = \sum_n^\infty b_n^i \hat{\psi}_n(\mathbf{r} - \mathbf{r}_i) \tag{4}$$

The coefficients b_n^i can be calculated from the applied field and the orthogonal properties of the basis functions. Similarly, the excitation field $\phi_j^E(\mathbf{r})$ on the scatterer at $\mathbf{r}_j$ can be expanded into a series with the same basis functions and unknown coefficients a_n^j,

$$\phi_j^E(\mathbf{r}) = \sum_n^\infty a_n^j \hat{\psi}_n(\mathbf{r} - \mathbf{r}_j). \tag{5}$$

The field scattered from the jth scatterer in equation (3) is expressible in terms of another set of basis functions $\psi_m(\mathbf{r} - \mathbf{r}_j)$ ($m = 1, 2, \ldots$)

$$\phi_j^S(\mathbf{r}) = \sum_m^\infty d_m^j \psi_m(\mathbf{r} - \mathbf{r}_j) \tag{6}$$

The functions $\psi_m(\mathbf{r} - \mathbf{r}_j)$ are outgoing waves from the scatterer. They are singular at $\mathbf{r} = \mathbf{r}_j$ and satisfy the radiation condition of the Helmholtz equation. The $\hat{\psi}_m(\mathbf{r} - \mathbf{r}_j)$ must be regular at $\mathbf{r} = \mathbf{r}_j$; they are taken as the regular part of the functions ψ_m.

As the ψ_m and $\hat{\psi}_m$ serve to discreetly represent the scattered and exciting fields, we have a matrix representation for the transition operator which relates the coefficients d_m to the a_n [10, 11],

$$d_m^j = \sum_n^\infty T_{mn}^j a_n^j \tag{7}$$

where T_{mn}^j is an infinite matrix. The elements of $\mathbf{T}^j$ for homogeneous scatterers may be expressed in terms of integrals of the basis functions over the boundary of the jth scatterer.

Furthermore, ψ_m in equation (6), which is the mth outgoing partial wave from the jth scatterer when evaluated near the ith scatterer, can be expressed as a linear combination of partial waves regular at $\mathbf{r}_i$.

$$\psi_m(\mathbf{r} - \mathbf{r}_j) = \sum_l^\infty D_{lm}(\mathbf{r}_j - \mathbf{r}_i)\hat{\psi}_l(\mathbf{r} - \mathbf{r}_i) \tag{8}$$

$D(\mathbf{r}_j - \mathbf{r}_i)$ is a shifting operator which shifts a field from a representation in terms of outgoing waves near $\mathbf{r}_j$ to a representation in terms of regular functions near $\mathbf{r}_i$. This shifting matrix, which is a function of $\mathbf{r}_j - \mathbf{r}_i$, may be worked out in detail in two or three dimensions by using addition theorems for the Bessel functions [6, 15]. For one-dimensional scalar waves $\mathbf{D}$ is given in the next section.

Substituting equations (7) and (8) into (6), and then (4)–(6) into equation (3), we obtain

$$\begin{aligned}\sum_n a_n^i \hat{\psi}_n(\mathbf{r} - \mathbf{r}_i) &= \sum_n b_n^i \hat{\psi}_n(\mathbf{r} - \mathbf{r}_i) \\ &\quad + \sum_{j\neq i}^N \sum_n \sum_m \sum_l D_{lm}(\mathbf{r}_j - \mathbf{r}_i) T_{mn}^j a_n^j \hat{\psi}_l(\mathbf{r} - \mathbf{r}_i).\end{aligned} \tag{9}$$

As the $\psi_n(\mathbf{r} - \mathbf{r}_i)$ form a linearly independent set of functions on the region near $\mathbf{r}_i$, we may equate term by term in equation (9) and obtain

$$a_n^i = b_n^i + \sum_{j\neq i}^N \sum_m \sum_l D_{nm}(\mathbf{r}_j - \mathbf{r}_i) T_{ml}^j a_l^j. \tag{10}$$

Equation (10) represents N coupled systems of an infinite number of equations for the unknown coefficients a_n^i. The solution to equation (10) clearly depends on the positions of all scatterers in the configuration via the dependence of b_n^i on $\mathbf{r}_i$ and of $\mathbf{D}(\mathbf{r}_j - \mathbf{r}_i)$ on $\mathbf{r}_j - \mathbf{r}_i$. These positions, in a realistic composite, are stochastic variables. The $\mathbf{T}^j$ may also be stochastic, as the quality of the scatterers may not be uniform. In multiple scattering calculations, attention is usually directed toward determining the average of all solutions to equation (10).

Let the probability density of finding the first scatterer at $\mathbf{r}_1$, the second at $\mathbf{r}_2, \ldots$, and the Nth at $\mathbf{r}_N$ be $p(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_N)$. The probability density for finding the ith scatterer at $\mathbf{r}_i$ is then

$$p(\mathbf{r}_i) \equiv \int p(\mathbf{r}_1, \ldots, \mathbf{r}_N) d\mathbf{r}_1\, d\mathbf{r}_2 \ldots' \ldots d\mathbf{r}_N. \tag{11}$$

The prime indicates the omission of $d\mathbf{r}_i$ from the integration. The conditional expectations or conditional averages of a configuration dependent quantity a are then defined as

$$p(\mathbf{r}_i)\langle a\rangle_i \equiv \int a\, p(\mathbf{r}_i, \ldots, \mathbf{r}_N) d\mathbf{r}_1 d\mathbf{r}_2 \ldots' \ldots d\mathbf{r}_N$$

By taking the conditional averages of equation (10) and applying the Lax quasi-crystalline assumption, which is a statistical independence assumption introduced to truncate a hierarchy of coupled equations on conditional averages of all orders, we obtain a closed set of integral equations for the $\langle a_n^i\rangle_i$,

$$\langle a_n^i\rangle_i = b_n^i + \sum_{j\neq i} \sum_m \sum_l \int d\mathbf{r}_j p(\mathbf{r}_j|\mathbf{r}_i) D_{mn}(\mathbf{r}_j - \mathbf{r}_i) T_{ml} \langle a_l^j\rangle_j \tag{12}$$

In these equations, the shifting operator $\mathbf{D}$ is presumably known, and the transition matrix $\mathbf{T}$ can be determined by the nature of the scatterer. The conditional probability density $p(\mathbf{r}_i|\mathbf{r}_j)$ is defined as

$$p(\mathbf{r}_j)p(\mathbf{r}_i|\mathbf{r}_j) \equiv \int p(\mathbf{r}_1, \ldots, \mathbf{r}_N) d\mathbf{r}_1 \ldots' \ldots' \ldots d\mathbf{r}_N.$$

It describes the probability of finding a scatterer at $\mathbf{r}_i$, given another scatterer at $\mathbf{r}_j$. We assume identical scatterers in (12) and $\mathbf{T}^j = \mathbf{T}$.

The application of the Lax approximation, first introduced by Lax [13], has been a point of prolonged discussion in the literature [5, 12, 14], and is discussed critically in the article by Frisch [16]. The approximation is exact if the configuration is nonstochastic, hence if it has no random irregularity. It was an implicit assumption of Lax [13] that under nearly crystalline circumstances, the assumption is reasonable. We shall refrain from making any further justification of its application here and accept it as an assumption.

One-Dimensional Wave Functions. As the discussion of scat-

tering requires the concepts of incident and outgoing waves, which in turn are best expressed in polar coordinates, we here present a system of one dimensional polar coordinates and discuss the corresponding solutions to the Helmholtz equation. These solutions will then be used to construct the transition matrix and the **D** operator in one dimension. Equation (12) will then be rendered into a form amenable to numerical solution.

One-dimensional space is usually designated by a Cartesian coordinate x $(-\infty < x < \infty)$. The polar coordinates of any point along the x-axis are $(r, 0)$ or (r, π). They are more conveniently expressed by (r, σ) where r takes on any value from zero to infinity and σ takes on the discreet values $+1$ (cos 0) and -1 (cos π). In terms of the Cartesian coordinate x, one dimensional polar coordinates are defined by

$$r = |x|, \quad \sigma = \text{sgn}\,(x) \tag{13}$$

There are two linearly independent solutions to the equation $d^2\phi/dx^2 = -k^2\phi$. The following regular solutions are chosen as the basis functions (even or odd) for exciting wave fields:

$$\hat{\psi}_e(\mathbf{r}) = \cos kr = \cos kx$$
$$\hat{\psi}_0(\mathbf{r}) = \sigma \sin kr = \sin kx \tag{14}$$

For outgoing waves, the basis functions are

$$\psi_e(\mathbf{r}) = e^{ikr}$$
$$\psi_0(\mathbf{r}) = -i\,\sigma\, e^{ikr} \tag{15}$$

Note that $\hat{\psi}_n$ is the regular part of ψ_n, as ψ_n has a discontinuous derivative at $r = 0$.

Consider a one-dimensional scatterer, a layer of thickness $2h$, which is confined to the region $r < h$ (Fig. 1). Any exciting or incident wave, impinging from the right, the left or from both directions, can be expanded in a series of regular basis functions as in equation (5)

$$\phi^E(\mathbf{r}) = a_e\hat{\psi}_e(\mathbf{r}) + a_0\hat{\psi}_0(\mathbf{r}) \tag{16}$$

In particular, for an incident rightward plane wave of unit amplitude

$$e^{ikx} = \hat{\psi}_e(\mathbf{r}) + i\hat{\psi}_0(\mathbf{r}). \tag{17}$$

The infinite series expansion is reduced to a sum of two terms in the one-dimensional case. The wave scattered by the layer is, according to equation (6),

$$\phi^S(\mathbf{r}) = d_e\psi_e(\mathbf{r}) + d_0\psi_0(\mathbf{r}), \quad r > h \tag{18}$$

The total wave outside the layer is

$$\phi = \phi^E + \phi^S, \quad r > h \tag{19}$$

The transition matrix as defined by equation (7) is thus a 2×2 matrix,

$$\begin{bmatrix} d_e \\ d_0 \end{bmatrix} = \begin{bmatrix} T_{ee} & T_{e0} \\ T_{0e} & T_{00} \end{bmatrix} \begin{bmatrix} a_e \\ a_0 \end{bmatrix} \tag{20}$$

Reciprocal properties of scattering insure the symmetry of the T matrix

$$\mathbf{T} = \mathbf{T}', \quad \text{or} \quad T_{e0} = T_{0e}.$$

Energy considerations, valid in the absence of loss mechanisms in the scatterer or embedding medium, insure the additional property of the T matrix

$$\mathbf{T}^*\mathbf{T} = -Re\ \mathbf{T}.$$

Furthermore, if the scatterer is symmetric about $r = 0$, then its transition matrix must be diagonal.

Transition Matrix for One Homogeneous Layer. For plane elastic waves propagating in the x-direction, the displacement u and stress τ can be calculated from the wave potential ϕ by

$$u = \partial\phi/\partial x$$
$$\tau = \beta\, \partial u/\partial x = \beta\, \partial^2\phi/\partial x^2 = -\rho\omega^2\phi \tag{21}$$

In the case of longitudinal waves, u in the foregoing represents displacement in the direction of wave propagation and β represents $\lambda + 2\mu = \kappa + 4\mu/3$ where λ and μ are the Lamé constants, μ the shear modulus and κ the bulk modulus of the material. In this case, the wave speed is $c = [(\lambda + 2\mu)/\rho]^{1/2}$. For a fluid medium we set $\mu = 0$ and note that u now represents longitudinal velocity. In the case of transverse waves (shear waves), u represents displacement in the y or z-direction and $\beta = \mu$. The corresponding wave speed is $c = [\mu/\rho]^{1/2}$.

For either longitudinal or transverse waves the boundary conditions in terms of one dimensional polar coordinates are continuity of displacements

$$\frac{d\phi(r, \sigma)}{dr} = \frac{d\phi^f(r, \sigma)}{dr} \quad \text{at} \quad r = h, \quad \sigma = \pm 1 \tag{22}$$

and continuity of stresses

$$\rho\phi(r, \sigma) = \rho_f\phi^f(r, \sigma) \quad \text{at} \quad r = h, \quad \sigma = \pm 1 \tag{23}$$

The super or subscript f refers to the inclusion material. The regular basis functions for waves inside the inclusion are

$$\hat{\psi}_e^f(\mathbf{r}) = \cos k_f r, \quad \hat{\psi}_0^f(\mathbf{r}) = \sigma \sin k_f r \tag{24}$$

in terms of which the wave refracted in the scattering layer may be expressed

$$\phi^f(\mathbf{r}) = f_e\hat{\psi}_e^f(\mathbf{r}) + f_0\hat{\psi}_0^f(\mathbf{r}), \quad r < h \tag{25}$$

Substituting equations (16) and (18) into (19), and equation (19) and (25) into the boundary conditions, one may solve algebraically for the unknown coefficients d_e, d_0, f_e, f_0 in terms of the known a_e and a_0. The elements of the transition matrix in equation (20) thus found are

$$T_{ee} = \frac{(\rho c/\rho_f c_f)\cos kh \sin k_f h - \sin kh \cos k_f h}{-(\rho c/\rho_f c_f)e^{ikh}\sin k_f h - ie^{ikh}\cos k_f h}$$
$$T_{00} = \frac{(\rho c/\rho_f c_f)\sin kh \cos k_f h - \cos kh \sin k_f h}{(\rho c/\rho_f c_f)ie^{ikh}\cos k_f h + e^{ikh}\sin k_f h} \tag{26}$$
$$T_{e0} = T_{0e} = 0$$

where $k = \omega/c$, $c^2 = \beta/\rho$, $k_f = \omega/c_f$, $c_f^2 = \beta_f/\rho_f$. Note that the foregoing results can be expressed compactly as

$$T_{sq} = -\delta_{sq}\left.\frac{(\rho c/\rho_f c_f)\hat{\psi}_s\hat{\psi}_q^{f\prime} - \hat{\psi}_s'\hat{\psi}_q^f}{(\rho c/\rho_f c_f)\psi_s\hat{\psi}_q^{f\prime} - \psi_s'\hat{\psi}_q^f}\right|_{\substack{r=h\\ \sigma=+1}} \tag{27}$$

where a prime indicates a derivative with respect to argument.

Substitution of the aforementioned results into equation (20) and then equation (18) completes the calculation of the scattered wave ϕ^S exterior to the layer $-h < x < h$. Note that for an incident wave moving in the $+x$-direction and impinging the layer from the left, we take $a_0 = i$ and $a_e = 1$ for ϕ^E in equation (16). Thus ϕ^S in the region $x < -h$ is the wave reflected by the layer, ϕ^f is the standing wave inside the layer, and $\phi^S + \phi^E$ in the region $x > h$ is the wave transmitted through the layer. The conventional amplitude transmission and reflection coefficients $\mathcal{T}$ and $\mathcal{R}$ are thus given by

$$\mathcal{T} = 1 + T_{ee} + T_{00} + iT_{e0} - iT_{0e}$$
$$\mathcal{R} = T_{ee} - T_{00} + iT_{e0} + iT_{0e} \tag{28}$$

The shifting operator **D**, defined in equation (8) is given by inspection of the identities, valid for $|x| < |\xi|$,

$$e^{ik|x-\xi|} = e^{ik|\xi|}[\cos kx - i\ \text{sgn}\,(\xi)\sin kx],$$
$$-i\ \text{sgn}\,(x - \xi)e^{ik|x-\xi|} = e^{ik|\xi|}[i\ \text{sgn}\,(\xi)\cos kx + \sin kx]$$

as

$$\begin{bmatrix} D_{ee}(\xi) & D_{e0}(\xi) \\ D_{0e}(\xi) & D_{00}(\xi) \end{bmatrix} = \begin{bmatrix} \psi_e(\xi) & -\psi_0(\xi) \\ \psi_0(\xi) & \psi_e(\xi) \end{bmatrix} \tag{29}$$

Multiple Scattering by Nearly Periodic Layers. To apply the transition matrix for a single layer as given by equation (26) to multiple scattering by N identical layers (Fig. 1), we first suppress the subscripts l, m, n in equation (12) and rewrite it compactly as

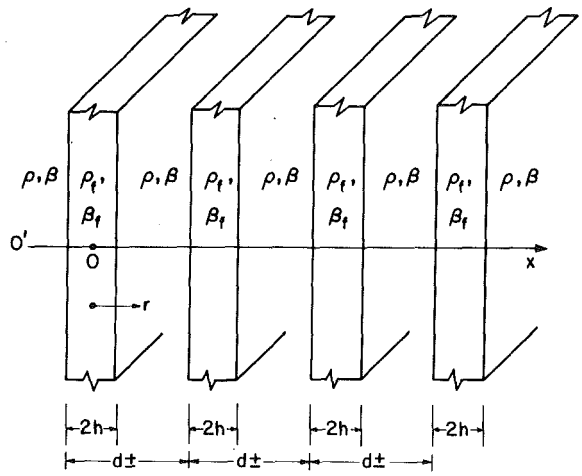


Fig. 1 Geometry of a laminated composite

$$\langle \mathbf{a}^i \rangle_i = \mathbf{b}^i + \sum_{j \neq i} \int d\mathbf{r}_j p(\mathbf{r}_j | \mathbf{r}_i) \mathbf{D}(\mathbf{r}_j - \mathbf{r}_i) \mathbf{T} \langle \mathbf{a}^j \rangle_j \tag{30}$$

where **a** and **b** are vectors in a two-dimensional vector space. We will refer to them as "excitors." **DT** is an excitor operator, a matrix in the same space. In the case of a two or three-dimensional composite this vector space becomes infinite dimensional. However, the sums which ran to infinity in Section 2 here run from 1 to 2, that is, the addition of the even (e) and odd (0) parts. Care must be taken to distinguish vectors such as $\mathbf{r}_i$ from excitors such as $\mathbf{a}^i$.

If the medium is statistically homogeneous one is permitted to seek an average total wave field which is an eigenfunction of translation and therefore of the form

$$\langle \phi(\mathbf{r}) \rangle = \phi_0 e^{i\mathbf{K}\cdot\mathbf{r}} \tag{31}$$

where **K** may be complex, and exp ($i\mathbf{K}\cdot\mathbf{t}$) is the eigenvalue of a translation of coordinates by a vector **t**. One is also permitted to seek an average excitor field which has the same eigenvalue, thus

$$\langle \mathbf{a}^i \rangle_i = e^{i\mathbf{K}\cdot\mathbf{r}_i} \mathbf{a}. \tag{32}$$

where **a** is an unknown two component constant excitor. For $\phi^A(\mathbf{r})$, we choose a plane wave of given wave length $2\pi/k$ incident from the left so that $\mathbf{b}^i = \mathbf{b} \exp(i\mathbf{k}\cdot\mathbf{r}_i)$. Substituting equation (32) and the assumed $\mathbf{b}^i$ into equation (30), we obtain

$$\mathbf{a}\, e^{i\mathbf{K}\cdot\mathbf{r}_i} = \mathbf{b}\, e^{i\mathbf{k}\cdot\mathbf{r}_i} + (N-1)[\int d\mathbf{r}\, p(\mathbf{r}|\mathbf{r}_i) \mathbf{D}(\mathbf{r} - \mathbf{r}_i) \mathbf{T} e^{i\mathbf{K}\cdot\mathbf{r}}] \mathbf{a} \tag{33}$$

From equation (17), the excitor **b** is given by $b_e = 1, b_0 = i$. Equation (33) can be rewritten as

$$[\mathbf{I} - \int d\mathbf{r}\, \rho(\mathbf{r}|\mathbf{r}_i) \mathbf{D}(\mathbf{r} - \mathbf{r}_i) \mathbf{T} e^{i\mathbf{K}\cdot(\mathbf{r}-\mathbf{r}_i)}] \mathbf{a} = e^{i(\mathbf{k}-\mathbf{K})\cdot\mathbf{r}_i} \mathbf{b} \tag{34}$$

where $\rho(\mathbf{r}|\mathbf{r}_i) = (N-1)p(\mathbf{r}|\mathbf{r}_i)$ is a conditional number density equal to the average number density as $|\mathbf{r} - \mathbf{r}_i| \to \infty$ in the absence of long range order, and **I** is the 2×2 identity matrix.

In a parallel configuration of planar scatterers confined to a half space equation (34) becomes after a change of variables $\xi = x - x_i$ for one-dimensional **r** and $\mathbf{r}_i$

$$\left[\mathbf{I} - \int_{-x_i}^{\infty} d\xi\, \rho(\xi|0)\, \mathbf{D}(\xi) \mathbf{T} e^{iK\xi}\right] \mathbf{a} = \mathbf{b}\, e^{i(k-K)x_i} \tag{35}$$

In equation (35), the integrand is another 2×2 matrix; **a** is an unknown two component excitor and **b** is given. The range of the integration is overall ξ consistent with $\xi = x - x_i$ where x is restricted to the half space occupied by the scatterers. As the edge of the scatterer distribution is taken to be at $x = 0$, the integral runs from $-x_i$ to $+\infty$.

We now assume $\rho(\xi|0)$ to be a two-point correlation function describing a quasi-regular array of layers at distance d apart with a spread in positions which grows with distance,

$$\rho(\xi|0) = \sum_{n=-n_i}^{\infty} e^{-(\xi-nd)^2/\epsilon^2|n|} \frac{1}{(\pi\epsilon^2|n|)^{1/2}} \tag{36}$$

where n_i = integer (x_i/d) and $n \neq 0$. Thus nearest neighbors are at a separation $d \pm \epsilon$; second nearest neighbors are at $2d \pm \sqrt{2}\epsilon$, and so on. There is no long range order. Equation (36) would describe a laminar composite constructed of alternating layers of a uniform thickness of one material and a matrix thickness with an error of $\pm\epsilon$. Such a distribution would result from a construction process employing many identical scattering layers sequentially and alternatingly stacked with another type of material layer of slightly varying thicknesses. Alternatively one could omit the occurrences of $|n|$ in equation (36) and have a conditional density distribution which admits of long range order. Such a distribution would result from a construction process which placed many identical scattering layers on predetermined regular positions, each to a tolerance of $\pm\epsilon$, and then poured in a filler material between the layers.

Either choice for $\rho(\epsilon|0)$, or another different one, could be used in equation (35). Our choice of equation (36) for purposes of illustration is arbitrary. Note that equation (36) allows some unphysical scatterer overlap, as the probability for $|\xi| < 2h$ is nonzero. But for small values of ϵ/d the error is small of order $1 - \mathrm{erf}\{(d-2h)/\epsilon\}$, a quantity of order 10^{-5} or smaller in the composites which will be considered.

In using this conditional density one will encounter integrals like

$$S_q = \sum_{n \neq 0} \int_{-x_i}^{\infty} d\xi\, e^{-(\xi-nd)^2/\epsilon^2|n|} \frac{1}{(\pi\epsilon^2|n|)^{1/2}} \psi_q(\xi) e^{iK\xi}, \qquad q = e, 0 \tag{37}$$

For $\epsilon/d \lesssim 1/2$ we obtain by substituting equation (15) into the foregoing,

$$\begin{aligned}
\begin{bmatrix} S_e \\ S_0 \end{bmatrix} &\cong \begin{bmatrix} 1 \\ -i \end{bmatrix} \sum_{n=1}^{\infty} e^{iknd} e^{iKnd} \int_{-\infty}^{+\infty} d\eta \frac{e^{-\eta^2/\epsilon^2 n}}{(\pi\epsilon^2 n)^{1/2}} e^{i(K+k)\eta} \\
&+ \begin{bmatrix} 1 \\ i \end{bmatrix} \sum_{n=-n_i}^{-1} e^{ik|n|d} e^{iKnd} \int_{-\infty}^{+\infty} d\eta \frac{e^{-\eta^2/\epsilon^2|n|}}{(\pi\epsilon^2|n|)^{1/2}} e^{i(K-k)\eta} \\
&= \begin{bmatrix} 1 \\ -i \end{bmatrix} \sum_{n=1}^{\infty} e^{i(K+k)nd}\, e^{-(K+k)^2\epsilon^2 n/4} \\
&+ \begin{bmatrix} 1 \\ -i \end{bmatrix} \sum_{n=1}^{n_i} e^{i(k-K)nd}\, e^{(K-k)^2\epsilon^2 n/4}
\end{aligned} \tag{38}$$

These sums converge for suitable k and K.

$$\begin{bmatrix} S_e \\ S_0 \end{bmatrix} = \begin{bmatrix} 1 \\ -i \end{bmatrix} L + \begin{bmatrix} 1 \\ i \end{bmatrix} M - \begin{bmatrix} 1 \\ i \end{bmatrix} P \tag{39}$$

with

$$L = e^{i(k+K)d-(k+K)^2\epsilon^2/4} [1 - e^{i(k+K)d-(k+K)^2\epsilon^2/4}]^{-1}$$

$$M = e^{i(k-K)d-(k-K)^2\epsilon^2/4} [1 - e^{i(k-K)d-(k-K)^2\epsilon^2/4}]^{-1}$$

$$P = e^{[i(k-K)d-(k-K)^2\epsilon^2/4]x_i/d} [1 - e^{i(k-K)d-(k-K)^2\epsilon^2/4}]^{-1}$$

Making use of all previous results we find that equation (35) becomes

$$\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} L + M - P & iM - iL - iP \\ -iM + iL + iP & L + M - P \end{bmatrix} \times \begin{bmatrix} T_{ee} & 0 \\ 0 & T_{00} \end{bmatrix}\right\} \begin{bmatrix} a_e \\ a_0 \end{bmatrix} = \begin{bmatrix} 1 \\ i \end{bmatrix} e^{i(k-K)x_i} \tag{40}$$

If $\epsilon = 0$ the terms in P on the left-hand side of equation (40) have the same x_i dependence as the right-hand side. Their equality follows from the validity of equation (40) for all x_i and from the extinction theorem that the applied field (the right-hand side) is extinguished by waves induced at the boundary of the composite. Note that P is due to the boundary ($x = 0, \xi = -x_i$) of the composite. Though we are aware of no published proofs of the extinction theorem in one dimension, this use of that theorem differs in no way from that employed in references [6, 12]. The interested reader may construct a proof of the extinction theorem in one dimension along the standard lines; see, for example, Born and Wolf [17].

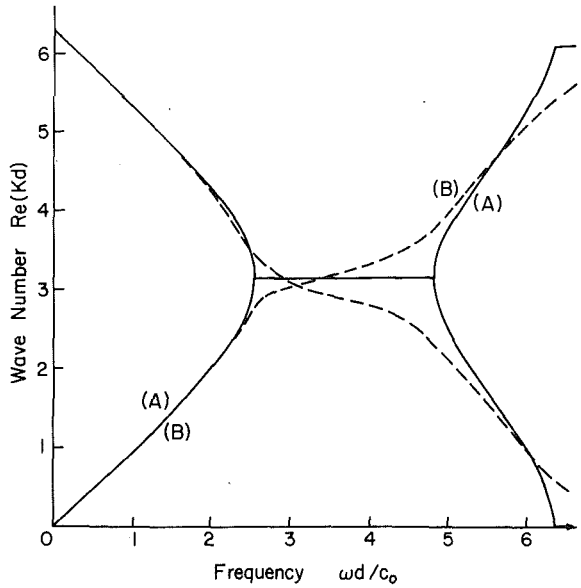


Fig. 2 Dispersion relation of Model 1 composite with elastic matrix; (*A*) Solid line, exactly periodic layers; (*B*) Dashed lines, periodic layers with 20 percent irregularity in spacing

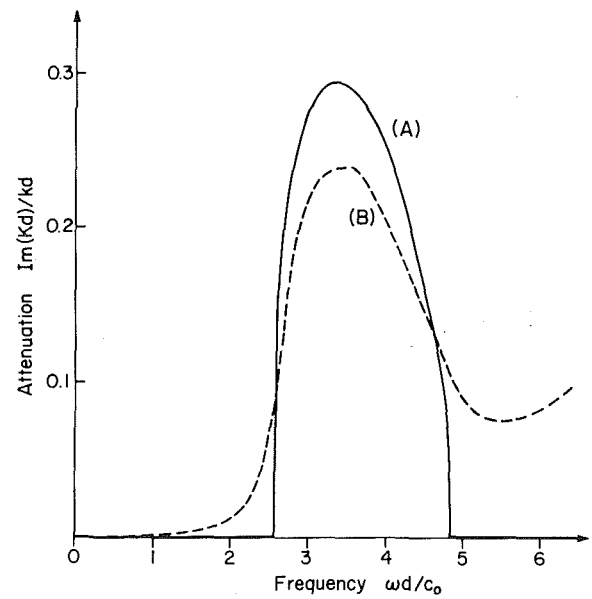


Fig. 3 Attenuation of waves in Model 1 composite with elastic matrix; (*A*), (*B*); same as in Fig. 2

When the x_i dependencies are dropped we are left with a homogeneous equation for **a**.

$$\left\{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} - \begin{bmatrix}L+M & iM-iL\\-iM+iL & L+M\end{bmatrix}\begin{bmatrix}T_{ee} & 0\\0 & T_{00}\end{bmatrix}\right\}\begin{bmatrix}a_e\\a_0\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \quad (41)$$

The inhomogeneous (dropped) part would serve to fix the complex magnitude of **a** as proportional to the applied field.

If $\epsilon \neq 0$ then the x_i dependencies in equation (40) are not identical. This is due to the error in having given distant scatterers a wide spread in position whereas in reality they are confined to the region to the right of the distribution boundary at $\mathbf{r} = 0$. For $\epsilon \neq 0$ we also drop the terms in P and the right-hand side, again obtaining equation (41). This is justified, in spite of the discrepancy in form, by the extinction theorem.

Equation (41) has nontrivial solution only if the determinant of the matrix multiplying **a** is zero; this requirement results in a transcendental equation for K.

$$[1 - T_{ee}(L+M)][1 - T_{00}(L+M)] - T_{ee}T_{00}(L-M)^2 = 0 \quad (42)$$

Its roots can be found numerically.

3 Results for Elastic Composites

The solutions of Kd in equation (42) as a function of $kd = \omega d/c_0$ were obtained interactively on a PDP 11/60 computer. Newton's method rather than Muller's method was chosen because the latter does not easily locate multiple roots.

We consider two models of laminated composites. For Model 1, the density of the scattering layer (inclusion) is twice that of the matrix, $\rho_f = 2\rho$, and the shear modulus is eight times larger, $\mu_f = 8\mu$. The spacing between the center lines of two inclusions is five times the half width of the inclusion, $d = 5h$. The Model 2 with softer scattering layers will be discussed in the next section, $c_0^2 = \mu/\rho$ in both cases.

The real and imaginary parts of Kd are listed in Table 1 for two cases: (A) When the periodicity is exact ($\epsilon = 0$). (B) When the periodic spacing is slightly irregular ($\epsilon = 0.2d$). The real part of the roots of both cases are shown in Fig. 3 with solid line for (A) and dashed lines for (B). They define the dispersion relation of the "average wave" in the composite medium. The normalized attenuation coefficient which is defined as Im $(Kd)/kd$ where $d = \omega/c_0$ is shown in Fig. 4, again solid line for (A) and dashed lines for (B). Note that

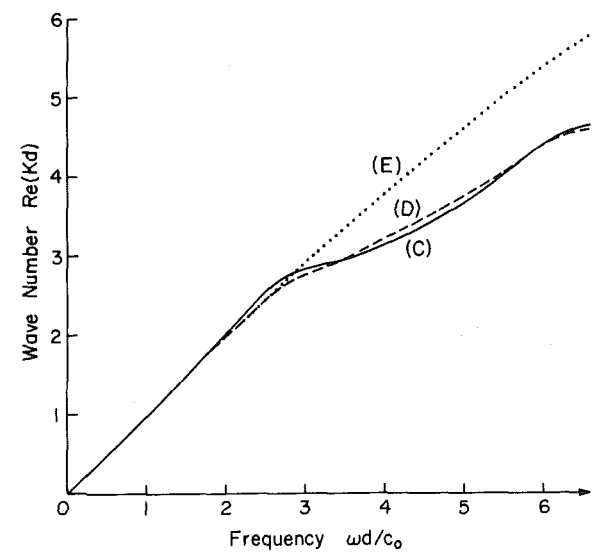


Fig. 4 Dispersion relation of Model 1 composite with viscoelastic matrix; (*C*) Solid line, exactly periodic layers; (*D*) Dashed lines, periodic layers with 20 percent irregularity in spacing; (*E*) Dotted lines, homogeneous Voigt solid

for perfectly periodic layers, the exact dispersion equation is available [18],

$$\cos Kd = \cos[k(d-2h) + 2k_f h] - \frac{(1-p)^2}{2p}\sin[k(d-2h)]\sin(2k_f h) \quad (43)$$

where $p^2 = \rho_f\mu_f/\rho\mu$. The roots in Table 1, Case (A), agree with those obtained from the exact equation.

The solid curve in Fig. 2 clearly exhibits the stop band, $2.55 < \omega d/c_0 < 4.85$, for the exactly periodic structure. Within the stop band, the attenuation as shown by the solid curve in Fig. 3 is very large. The attenuation is caused by the blocking of propagation of waves through the periodic layers, rather than by a dissipation of energy or mode conversion (S-wave to P-wave, or vice versa) in the medium. The situation is analogous to the forcing of an oscillator or a system of oscillators at an unnatural frequency, i.e., at a frequency at which the system does not oscillate in the absence of external forces. There is no net transfer of energy from the forcing mechanism to the oscillator.

Table 1 Real and imaginary wave numbers of the average wave in laminated composite, Model No. 1 ($d = 5h$, $\rho_f = 2\rho$, $\mu_f = 8\mu_0$)

$\frac{\omega d}{c_o}$ \ Kd	(A) Regular Elastic	(B) Irregular Elastic	(C) Regular Viscous	(D) Irregular Viscous	(E) Homogeneous Voigt Matrix
0.5	0.478	0.478 + 0.000i	0.478 + 0.011i	0.478 + 0.011i	0.500 + 0.012i
1.0	0.964	0.964 + 0.002i	0.960 + 0.045i	0.960 + 0.047i	0.996 + 0.050i
2.0	2.024	2.022 + 0.023i	1.979 + 0.219i	1.968 + 0.235i	1.971 + 0.195i
2.2	2.281	2.276 + 0.044i	2.202 + 0.293i	2.177 + 0.314i	2.161 + 0.235i
2.4	2.601	2.567 + 0.103i	2.425 + 0.404i	2.376 + 0.420i	2.350 + 0.278i
2.6	3.142 + 0.304i	2.851 + 0.282i	2.617 + 0.564i	2.547 + 0.549i	2.537 + 0.324i
2.8	3.142 + 0.644i	2.992 + 0.496i	2.743 + 0.735i	2.680 + 0.681i	2.722 + 0.374i
3.0	3.142 + 0.821i	3.066 + 0.646i	2.823 + 0.876i	2.785 + 0.798i	2.905 + 0.426i
3.2	3.142 + 0.932i	3.120 + 0.746i	2.883 + 0.984i	2.873 + 0.895i	3.086 + 0.482i
3.4	3.142 + 1.000i	3.167 + 0.808i	2.937 + 1.063i	2.956 + 0.973i	3.264 + 0.540i
3.6	3.142 + 1.033i	3.214 + 0.839i	2.993 + 1.120i	3.038 + 1.035i	3.440 + 0.600i
3.8	3.142 + 1.033i	3.265 + 0.841i	3.055 + 1.157i	3.123 + 1.084i	3.614 + 0.663i
4.0	3.142 + 1.000i	3.324 + 0.815i	3.125 + 1.179i	3.213 + 1.124i	3.785 + 0.729i
4.2	3.142 + 0.930i	3.397 + 0.762i	3.208 + 1.190i	3.310 + 1.159i	3.953 + 0.797i
4.4	3.142 + 0.810i	3.496 + 0.083i	3.309 + 1.194i	3.413 + 1.190i	4.119 + 0.866i
4.6	3.142 + 0.610i	3.632 + 0.589i	3.412 + 1.197i	3.523 + 1.222i	4.283 + 0.938i
4.8	3.142 + 0.092i	3.812 + 0.500i	3.538 + 1.203i	3.638 + 1.255i	4.444 + 1.011i
5.0	3.806	4.024 + 0.439i	3.673 + 1.217i	3.755 + 1.292i	4.602 + 1.086i
5.2	4.135	4.249 + 0.409i	3.814 + 1.243i	3.875 + 1.333i	4.758 + 1.163i
5.4	4.423	4.473 + 0.405i	3.956 + 1.279i	3.993 + 1.377i	4.911 + 1.241i
5.6	4.701	4.691 + 0.421i	4.097 + 1.325i	4.111 + 1.426i	5.061 + 1.321i
5.8	4.986	4.901 + 0.455i	4.232 + 1.380i	4.226 + 1.476i	5.209 + 1.401i
6.0	5.296	5.098 + 0.504i	4.362 + 1.440i	4.383 + 1.529i	5.353 + 1.483i
6.2	5.686	5.280 + 0.561i	4.485 + 1.504i	4.446 + 1.583i	5.497 + 1.566i
6.4	6.283 + 0.420i	5.445 + 0.623i	4.601 + 1.571i	4.552 + 1.637i	5.637 + 1.649i

Table 2 Real and imaginary wave numbers for the average wave in laminated composite, Model No. 2 ($d = 6h$, $\rho_f = \rho/2$, $\mu_f = \mu_0/8$)

$\frac{\omega d}{c_o}$ \ Kd	(A)	(B)	(C)	(D)	(E)
0.50	0.843	0.843 + 0.004i	0.843 + 0.004i	0.843 + 0.008i	0.050 + 0.012i
1.00	1.763	1.762 + 0.028i	1.761 + 0.015i	1.762 + 0.043i	0.996 + 0.050i
1.20	2.221	2.212 + 0.070i	2.218 + 0.021i	2.213 + 0.091i	1.194 + 0.071i
1.35	2.723	2.632 + 0.196i	2.718 + 0.036i	2.628 + 0.225i	1.341 + 0.090i
1.37	2.844	2.692 + 0.235i	2.835 + 0.046i	2.687 + 0.263i	1.360 + 0.093i
1.38	2.914	2.716 + 0.254i	2.901 + 0.057i	2.711 + 0.282i	1.370 + 0.094i
1.39	3.019	2.740 + 0.274i	2.983 + 0.084i	2.734 + 0.301i	1.380 + 0.095i
1.40	3.142 + 0.179i	2.769 + 0.301i	3.073 + 0.185i	2.762 + 0.328i	1.390 + 0.097i
1.50	3.142 + 0.635i	2.295 + 0.538i	3.131 + 0.835i	2.921 + 0.558i	1.488 + 0.111i
1.65	3.142 + 0.942i	3.023 + 0.806i	3.148 + 0.944i	3.030 + 0.824i	1.633 + 0.134i
1.80	3.142 + 1.131i	3.075 + 0.985i	3.164 + 1.135i	3.095 + 1.006i	1.779 + 0.159i
2.00	3.142 + 1.287i	3.122 + 1.140i	3.188 + 1.298i	3.165 + 1.169i	1.971 + 0.195i
2.40	3.142 + 1.383i	3.195 + 1.242i	3.260 + 1.417i	3.305 + 1.306i	2.350 + 0.278i
2.70	3.142 + 1.294i	3.255 + 1.163i	3.353 + 1.371i	3.441 + 1.287i	2.630 + 0.349i
3.00	3.142 + 1.033i	3.350 + 0.925i	3.512 + 1.217i	3.637 + 1.185i	2.905 + 0.426i
3.30	3.142 + 0.299i	3.643 + 0.469i	3.804 + 0.985i	3.428 + 1.035i	3.175 + 0.510i
3.33	3.332	3.702 + 0.422i	3.844 + 0.962i	3.963 + 1.020i	3.202 + 0.519i

The existence of the stop band has been confirmed experimentally for a precisely constructed layered model [19].

As can be seen from the dashed curves in both figures, the effect of the 20 percent irregularity in spacing has been to soften the sharp band features present in the perfectly regular case. The dispersion Curve B in Fig. 2 is now continuous. A small amount of attenuation has appeared at moderate frequencies $kd < 2.5$. This is attributed to the "scattering loss" which represents the part of energy not carried by the coherent average wave. Attenuation within the stop band has fallen. This is because the irregularity has hindered the ability of the periodic structure to exclude certain frequencies.

Note also that this 20 percent irregularity has essentially destroyed the higher pass bands. The Curve (B) in Fig. 3 shows a local minimum in attenuation of about 0.4 nepers per periodicity d at $kd \simeq 5.2$. The pass band still exists but for practical purposes cannot carry coherent information.

The group velocity v_g for the average wave is defined as Re $(dK/d\omega)^{-1}$ when ω is a real variable. Thus, at $kd = 2.55$ and 4.85, the group velocity for the average wave in Case (A) is zero. Within the stop band, $2.55 < \omega d/c_0 < 4.85$, the group velocity is also zero because $dK/d\omega$ is pure imaginary, and the waves at these frequencies are standing waves. However, in Case (B), waves at the same frequencies are of propagating mode. The group velocity at $kd = 2.55$ is about one half of that at low frequencies.

In Fig. 2 are shown two branches; the first one runs from the origin to upper right with positive group velocities, and the second one runs from upper left to lower right with negative group velocities. In Case (A), the negative branch is associated with waves moving toward negative direction of the x-axis. The interpretation for the negative branch in Case (B) is not clear because the calculated K's are such that the analytic sum performed in equation (38) did not converge as $n \to \infty$. We therefore suspect this branch of being unphysical. In any case, negative group velocity modes are ruled out by most ultrasonic experiments.

Of a more problematic nature are branches not included in Fig. 3. These are solutions for Re $(Kd) > 2\pi$, $0 < kd < 7$ with positive group velocity. In Case (A), these solutions are merely duplicates of the first branch, and there is no physical distinction between a solution Kd and a solution $Kd + 2\pi$ [20]. In Case (B), each solution still exists though its exact location has shifted. The difficulty is in that each solution shifts a different amount, hence they are now distinguishable.

4 Results for Viscoelastic Composites

The formalism developed in Sections 2 and 3 applies also in the case of a viscoelastic matrix material, a viscoelastic inclusion, or both, when k and μ, k_f and μ_f, or both are complex quantities. One need not assume any particular model for viscoelastic behavior, but rather merely insert observed complex values for k and μ as functions of real ω and material type. Nevertheless, for purposes of comparison in this paper, we assume a Voigt model for the matrix material.

Viscoelastic Model 1. Consider again a laminated composite as in the previous section and assume the inclusion material remains unchanged but the matrix material be modified in that k and μ are now complex functions of real frequency ω,

$$\mu = \mu_0[1 - i\tau\omega]$$
$$k = \omega/[c_0(1 - i\tau\omega)^{1/2}] \tag{44}$$

The μ_0 and $c_0 = (\mu_0/\rho)^{1/2}$ are the low frequency modulus and wave speed, respectively. We choose the same geometry and elastic properties as those for Model 1, and choose τ, the relaxation time, to equal $d/10c_0$. This value ensures that the viscous effects are not dominant in the domain of interest, nor are they negligible.

The complex wave number Kd of the average wave as calculated from equation (42) are shown again in Table 1 for three more cases: (C) An exactly periodic array of elastic layers sandwiching the Voigt matrix, (D) 20 percent irregularity in the spacing of the Voigt matrix. Thus (C) is the viscoelastic counterpart of (A) in the previous section, and (D) that of (B). In the table, we include the case (E) for a homogeneous Voigt solid. The real parts of Kd are shown in Fig. 4 with solid line for (C), dashed lines for (D), and dotted lines for (E). The corresponding imaginary parts, divided by $\omega d/c_0$, are shown in Fig. 5.

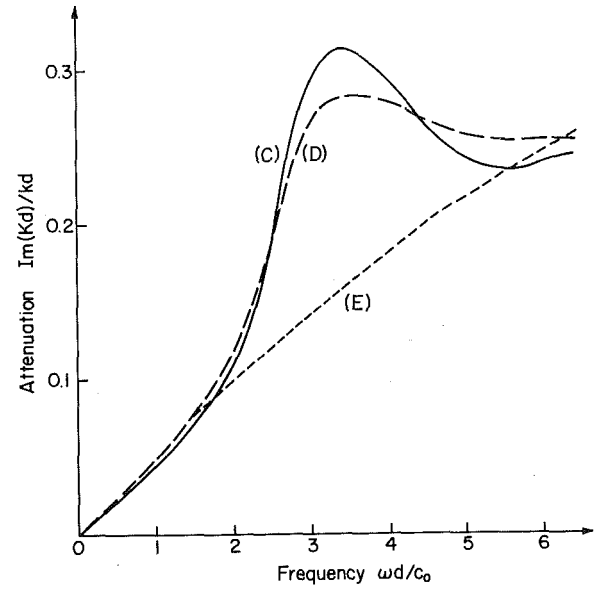


Fig. 5 Attenuation of waves in Model 1 composite with viscoelastic matrix (C), (D), (E); same as in Fig. 4

From the dispersion Curve (C) in Fig. 4, we see that the addition of viscosity in exactly periodic matrix layers has eliminated completely the standing wave mode within the stop band. Comparing Curve (C) with Curve (B) in Fig. 2, we find the effect of adding viscosity is similar to that of having irregularity in spacing. The addition of either irregularity or viscosity increases the group velocity in this frequency range.

The Curve (D) in Fig. 4 shows the combined effect of viscosity and irregular spacing. The addition of variations in periodicity changes only slightly the dispersion relation.

The effect of viscosity on the attenuation is pronounced. Outside the stop band, the attenuation has been greatly increased in Case (C) as shown in Fig. 5. Making the composite spacing irregular, thereby adding scattering losses, increases the attenuation more (Curve D in Fig. 6). Within the stop band ($\omega d/c_0 > 2.55$), the attenuation is more severe for the regular composite, rather than for the irregularly spaced layers. This is the same as in the elastic case treated in the preceding section. These results indicate that the attenuation in this region is due primarily to periodicity rather than either viscosity or irregularity.

Viscoelastic Model 2. For purposes of ascertaining the effects of viscosity and irregularity in a laminar composite independent of the choice of geometry and material parameters, we investigated another model for which the elastic inclusions have lighter density, $\rho_f = \rho/2$, and softer rigidity, $\mu_f = \mu_0/8$. The spacing between the center lines of two elastic layers is slightly increased, $d = 6h$. The viscoelastic matrix (Voigt body) has a complex modulus $\mu = \mu_0(1 - i\tau\omega)$ where $\tau c_0 = d/10$ and $c_0^2 = \mu_0/\rho$.

As was done for Model 1, we assume all four cases: (A), (B), (C), and (D). The calculated complex wave numbers for the average wave are listed in Table 2. Again, the Case (E) is for a homogeneous Voigt solid where $Kd = kd$. Note that the values for Case (E) in Table 2 are the same as those in Table 1. The dispersion relations, Re (Kd), for all cases are shown in Fig. 6; and the attenuations, Im $(Kd)/(\omega d/c_0)$, in Fig. 7.

As can be seen from Fig. 6, the sharpness of the band structure (Curve A) is again destroyed by the irregularity in spacing (Curve B). However, the viscosity has little effect on the band structure (Curve C, Fig. 6). Its significance is only in the increasing of attenuation in the pass band (Fig. 7).

One striking contrast between Model 1 and Model 2 is that in the

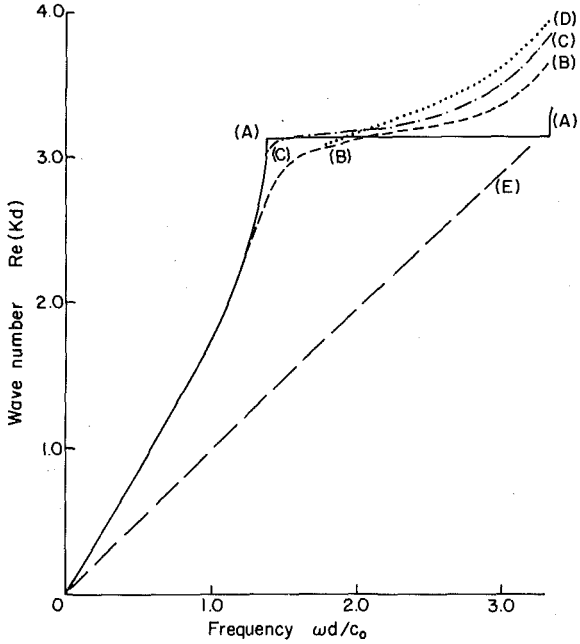


Fig. 6 Dispersion relation of Model 2 Composite; (*A*) Solid line, exactly periodic layers, all elastic; (*B*) Dashed line, irregularly periodic layers, all elastic; (*C*) Dash-dot line, exactly periodic layers, elastic-viscoelastic; (*D*) Dotted line, irregularly period layers, elastic-viscoelastic; (*E*) Long dashed lines, homogeneous Voigt solid

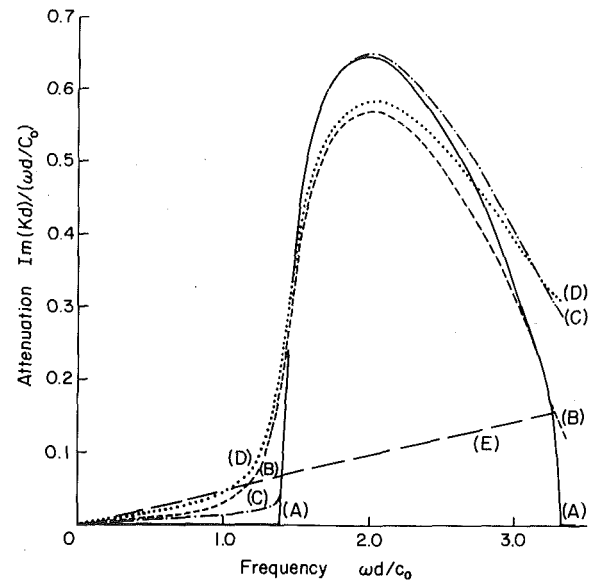


Fig. 7 Attenuation of waves in Model 2 composite; (*A*), (*B*), (*C*), (*D*), (*E*); same as in Fig. 6

former, the viscosity as well as irregular spacing destroys sharp band structure, whereas in Model 2, the viscosity is relatively ineffective. Furthermore, the low frequency attenuation of Model 1 is dominated by viscosity, and that of Model 2 by irregularity.

Conclusion. Based on the results for Model 1 (Table 1, Figs. 2–5) and those for Model 2 (Table 2, Figs. 6, 7), we conclude that for laminated composites, there exists a sharp stop band for exactly periodic, elastic layers. The values for the lower and upper limits of the stop band can easily be calculated from equation (43). Outside the stop band, the attenuation is zero; inside it, very large. The high attenuation within the stop band is attributed to the blocking of the transmission of waves through the layers. The waves are in standing mode and the group velocity is zero. There is, however, no dissipation of energy associated with the stop band.

The sharp band structure is destroyed by irregularity in the spacing of elastic layers, viscosity in the matrix material, or both. Within the stop band, the irregularity breaks up the symmetry responsible for standing waves, resulting in a propagating mode with reduced attenuation. Outside the stop band, the irregularity gives rise to a scattering loss in the form of small attenuation at low frequencies. This scattering loss is attributed to the incoherent scattering of waves out of the coherent average field.

Outside the stop band, the attenuation due to viscosity in a regular composite increases with frequency, and is much larger than the attenuation due to irregularity. It is approximately equal to the attenuation in a homogeneous body made of the same viscoelastic material when the viscosity is introduced in the less stiff and lighter material (Model 1), but it is smaller otherwise (Model 2). Inside the stop band, the addition of viscosity increases the attenuation.

Finally, we note from Tables 1 and 2 that at frequencies below the stop band the attenuation due to the combined effect of viscosity and irregularity is roughly equal to the sum of that due to each part. Within the stop band, the value for total attenuation is about equal to that for the regular viscoelastic composite.

Acknowledgment

This study was conducted under a contract from the Air Force Office of Scientific Research.

References

1 Kohn, W., Krumhansl, J. A., and Lee, E. H., "Variational Methods for Dispersion Relations and Elastic Properties of Composite Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, June 1972, pp. 327–336.

2 Mukherjee, S., and Lee, E. H., "Dispersion Relations and Mode Shapes for Waves in Laminated Viscoelastic Composites by Variational Methods," *International Journal of Solids and Structures,* Vol. 14, 1978, pp. 1–13.

3 Sun, D. T., Achenbach, J. D., and Herrmann, G., "Continuum Theory for a Laminated Medium," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, Sept. 1968, pp. 467–475.

4 Merzbacher, E., *Quantum Mechanics,* 2nd ed., Wiley, New York, 1970, Chapter 6.7.

5 Bose, S. K., and Mal, A. K., "Elastic Waves in a Fiber-Reinforced Composite," *Journal of the Mechanics and Physics of Solids,* Vol. 22, 1974, pp. 217–229.

6 Varadan, V. K., Varadan, V. V., and Pao, Y.-H., "Multiple Scattering of Elastic Waves by Cylinders of Arbitrary Cross Section. 1. SH Waves," *Journal of the Acoustical Society of America,* Vol. 63, 1978, pp. 1310–1319.

7 Truell, R., Elbaum, C., and Chick, B. B., *Ultrasonic Methods in Solid State Physics,* Academic Press, New York, 1969, Chapter 3.

8 Becus, G. A., "Wave Propagation in Imperfectly Periodic Structures: A Random Evolution Approach," *Journal of Applied Mathematics and Physics, (ZAMP),* Vol. 29, 1978, pp. 252–261.

9 Ziegler, F., "Wave Propagation in Periodic and Disordered Layered Composite Elastic Materials," *International Journal of Solids and Structures,* Vol. 13, 1977, pp. 293–305.

10 Waterman, P. C., "New Formulation of Acoustic Scattering," *Journal of the Acoustical Society of America,* Vol. 45, 1969, pp. 1417–1429.

11 Pao, Y.-H., "The Transition Matrix for the Scattering of Acoustic Waves and for Elastic Waves," *Modern Problems in Elastic Wave Propagation,* eds. Miklowitz, J., and Achenbach, J. D., Wiley and Interscience, New York, 1978, pp. 123–144.

12 Waterman, P. C., and Truell, R., "Multiple Scattering of Waves," *Journal of Mathematics and Physics,* Vol. 2, 1961, pp. 512–537.

13 Lax, M., "Multiple Scattering of Waves, II. The Effective Field in Dense Systems," *Physical Review,* Vol. 85, 1952, pp. 621–629.

14 Keller, J. B., "Stochastic Equations and Wave Propagation in Random Media," *Proceedings, Symposium of Applied Mathematics,* Vol. 16, 1964, pp. 145–170.

15 Peterson, B., "Multiple Scattering of Waves by an Arbitrary Lattice," *Physical Review,* Series A, Vol. 16, 1977, pp. 1363–1370.

16 Frisch, U., "Wave Propagation in Random Media," in *Probabilistic Methods in Applied Mathematics,* ed., Bharucha-Reid, A. T., Academic Press, 1968, p. 121.

17 Born, M., and Wolf, E., *Principles of Optics,* 5th ed., Pergamon Press, New York, p. 100.

18 Christensen, R. M., *Mechanics of Composite Materials,* Wiley, New York, 1979, Section 7.3.

19 Robinson, C. W., and Leppelmeier, C. W., "Experimental Verification of Dispersion Relations for Layered Composites," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, 1974, pp. 89–91.

20 Brillouin, L., *Wave Propagation in Periodic Structures,* Dover, 1953.

T.-C. Soong
Principal Engineer.

C. Li
Technical Specialist.

Xerox Corporation,
Rochester, N.Y. 14644

On the Unbonded Contact Between Plates and Layered Cylinders

An approximate solution is presented for the deflection of a thin plate and for the stresses and deformation of a layered circular cylinder in which the plate is pressed against the layered soft cylinder by a hard cylinder. The contact is considered as frictionless, and the plate has an initial curvature. Since the load on the plate is radial and the initial curvature is arbitrary, nonlinear beam theory is used. The state of the layered circular cylinder is described by a general stress function. The collocation method is used to relate loading and displacements between the contact surface of the plate and the soft substrate of the cylinder.

Introduction

For bodies in contact without bond, the region of contact is an important unknown which distinguishes it from some hard, punch-contact-type problems where stresses are sought. Among recent papers on unbonded contact which involve thin-walled members, Wu and Plunkett [1] studied thin, circular rings compressed between two rigid anvils of constant curvature, and Callegari and Keller [2] solved the contact of inflated membranes with rigid surfaces. These analyses used existing equations, but the solutions were new. Weitsman [3] presented an approximate solution for the radius of contact between a plate and a semi-infinite elastic half space. Hogg [4] solved the same problem but with the interface bonded. He found that tension existed in some regions of the bonded interface.

In the present paper, we solve a class of problems in which a plate, either initially curled or straight, is pressed between two circular cylinders. One cylinder is hard and the other has elastic layers bonded to a hard core. Because of the arbitrary curvature of the plate, the arc length of the contact zone between the plate and the soft cylinder could be quite large compared with the small indentation prescribed between the two.

For the contact of the layered soft roll, we use the same stress function approach as used by Hahn and Levinson [5, 6]. However, the numerical solution is made simpler—but of comparable accuracy—by using the collocation method instead of the Schmidt method of orthogonal function.

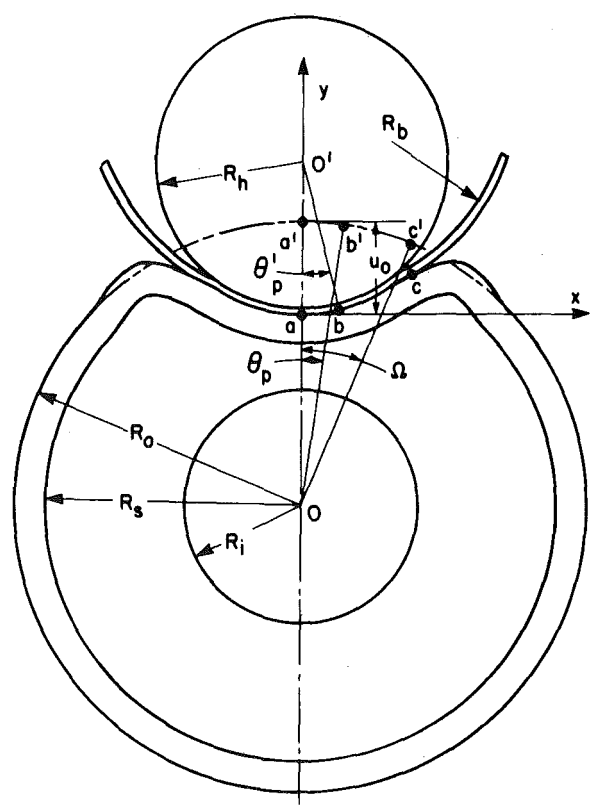


Fig. 1 Notations; equation derivation is only for a single-layer cylinder, but numerical results include a two-layer cylinder

Formulation of the Solution

Large Deflection of a Beam Under Distributed Loads. Fig. 1 shows a hard roll (circular cylinder) pressed on a hard-cored soft roll

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, December, 1978; final revision, April, 1980. Paper No. 80-WA/APM-11.

(layered cylinder) with a thin beam (plate) pressed in the nip. Assume the deflection is symmetric to the y-axis. The circular arc, $a'c'$, on the undeformed surface of the soft roll is now displaced to the contact interface, ac, where arc ab follows the contour of the hard roll and arc bc is the contact between the soft roll and the cantilevered part of the beam. The hard roll does not contact the beam between b and c.

Let s be the nondimensional curvilinear coordinate of the beam measured along its axis from a. It is the true length divided by the outer radius, R_0, of the soft roll. Let θ be the coordinate of a point originally on the undeformed arc of the soft roll. If the surface strain is not negligible compared with unity, then

$$\frac{d\theta}{ds} = (1 + \epsilon_\theta)^{-1} \cong 1 - \epsilon_\theta, \tag{1}$$

where ϵ_θ is the tangential surface strain of the soft roll.

Let ϕ be the slope of the deflected center line of the beam measured counterclockwise from the x-axis and σ_r, positive inward, be the radial stress acting on the beam from the soft roll. Then, the Euler-Bernoulli beam theory yields

$$\frac{d^2\phi}{ds^2} = \int_s \bar{\sigma}_r{}^* \cos(\phi + \theta) ds, \tag{2}$$

where $\bar{\sigma}_r{}^* = \sigma_r R_0{}^3/EI$, E is the Young's modulus, and I the sectional moment of inertia of the beam.

Using equation (1) to change coordinates from s to θ and dropping terms with $\epsilon_\theta{}^2$, equation (2) becomes

$$(1 - 2\epsilon_\theta)\frac{d^2\phi}{d\theta^2} - (1 - \epsilon_\theta)\frac{d\epsilon_\theta}{d\theta}\cdot\frac{d\phi}{d\theta} = \int_\theta^\Omega \bar{\sigma}_r{}^* \cos(\phi + \theta)(1 + \epsilon_\theta) d\theta, \theta_p \le \theta \le \Omega \tag{3}$$

We shall assume that the tangential surface strain ϵ_θ and the rate of change $d\epsilon_\theta/d\theta$ are both negligible and can be dropped in equation (3). Then we have

$$\frac{d^2\phi}{d\theta^2} = \int_\theta^\Omega \bar{\sigma}_r{}^* \cos(\phi + \theta) d\theta. \tag{4}$$

To solve equation (4), let

$$\{f_1(\theta), f_2(\theta)\} = \int_{\theta_p}^\theta \bar{\sigma}_r{}^* \{\cos\theta, \sin\theta\} d\theta. \tag{5a,b}$$

Then, equation (4) becomes

$$\frac{d^2\phi}{d\theta^2} = \int_{\theta_p}^\Omega \bar{\sigma}_r{}^* \cos(\phi + \theta) d\theta - \cos\phi f_1(\theta) + \sin\phi f_2(\theta), \quad \theta_p \le \theta \le \Omega. \tag{6}$$

Equation (6) may be decomposed into the following first-order set:

$$\frac{df_1}{d\theta} = \bar{\sigma}_r{}^* \cos\theta,$$
$$\frac{df_2}{d\theta} = \bar{\sigma}_r{}^* \sin\theta,$$
$$\frac{d\phi}{d\theta} = \bar{M}, \tag{7a–d}$$

and

$$\frac{d\bar{M}}{d\theta} = \int_{\theta_p}^\Omega \bar{\sigma}_r{}^* \cos(\phi + \theta) d\theta - \cos\phi f_1(\theta) + \sin\phi f_2(\theta).$$

The starting values of $f_1, f_2, \bar{M}$, and ϕ are their corresponding values at $\theta = \theta_p$, which are

$$f_1 = f_2 = 0,$$
$$\bar{M} = \bar{R}_h, \tag{8}$$

and

$$\phi = \theta_p' = \theta_p/\bar{R}_h,$$

where θ_p' is the subtended angle to the hard roll of θ_p as shown in Fig. 1. $\bar{R}_h$ is the hard roll cylinder radius, R_h, divided by R_0. The effect of the beam's thickness on the contact geometry may be approximated by adding its thickness to the hard roll radius R_h.

When σ_r of the soft roll is prescribed, there are only two unknowns, Ω and θ_p, in the beam solution. With a given indentation, u_0, which is defined as the interference between the two rolls, σ_r can be determined from the contact problem. The correct values of θ_p and Ω should be such that,

$$\sigma_{r\langle\theta=\Omega\rangle} = 0, \tag{9}$$

and

$$\frac{d\phi}{d\theta_{\langle\theta=\Omega\rangle}} = R_0/R_b, \tag{10}$$

where R_b is the initial radius of curvature of the beam, positive upward.

General Stress and Strain Expressions of the Layered Soft Roll

The radial surface stress, σ_r, required to produce a prescribed radial displacement at the surface of the cored soft roll and other strain and displacement expressions can be found from existing papers (reference [5]). Simple derivations and equations to be used for the present analysis are given as follows.

The well-known stress function (in cylindrical coordinates) which is symmetric to θ is used for the soft layer of the roll under external loads:

$$U = a_0 \ln r + \frac{1}{2} b_0 r^2 + (b_1 r^3 + c_1 r^{-1} + d_1' r \ln r)\cos\theta + \frac{1}{2} d_1 r\theta \sin\theta + \sum_{n=2}^{\infty} (a_n r^n + b_n r^{n+2} + c_n r^{-n} + d_n r^{-n+2})\cos n\theta, \tag{11}$$

where a term, $r^2 \ln r$, has been omitted so that displacements will be single-valued with respect to θ.

Let σ_r, σ_θ, $\tau_{r\theta}$ be the radial, tangential, and shear stresses, u and v be the radial and tangential displacements, and E^* and ν be the Young's modulus and Poisson's ratio, respectively, of the outer layer of the soft roll. The quantities are nondimensionalized as shown in the following:

$$\{\bar{\sigma}_r, \bar{\sigma}_\theta, \bar{\tau}_{r\theta}\} = \frac{1}{E^*}\{\sigma_r, \sigma_\theta, \tau_{r\theta}\} \tag{12a}$$

$$\{\bar{u}, \bar{v}, \bar{r}, \bar{R}_i\} = \frac{1}{R_0}\{u, v, r, R_i\}, \quad \bar{d}_1 = R_0{}^{-1} d_1'/E^* \tag{12b}$$

$$\bar{a}_n = R_0{}^{n-2} a_n/E^*, \quad \bar{b}_n = R_0{}^n b_n/E^*, \quad n = 0, 1, 2, \ldots, \tag{12c}$$

$$\bar{c}_n = R_0{}^{-n-2} c_n/E^*, \quad \bar{d}_n = R_0{}^{-n} d_n/E^*, \quad n = 1, 2, \ldots, \tag{12d}$$

The stress and strain expressions are well known, they are

$$\bar{\sigma}_r = \frac{1}{E^*}\left(\frac{1}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2}\frac{\partial^2 U}{\partial\theta^2}\right), \tag{13}$$

$$\bar{\sigma}_\theta = \frac{\partial^2 U}{E^* \partial r^2}, \tag{14}$$

$$\bar{\tau}_{r\theta} = -\frac{\partial}{E^* \partial r}\left(\frac{1}{r}\frac{\partial U}{\partial\theta}\right), \tag{15}$$

$$\epsilon_r = \frac{\partial u}{\partial r} = \kappa\bar{\sigma}_r - \beta\bar{\sigma}_\theta, \tag{16}$$

$$\epsilon_\theta = \frac{u}{r} + \frac{1}{r}\frac{\partial v}{\partial\theta} = \kappa\bar{\sigma}_\theta - \beta\bar{\sigma}_r, \tag{17}$$

and

$$\gamma_{r\theta} = \frac{1}{r}\frac{\partial u}{\partial\theta} + \frac{\partial v}{\partial r} - \frac{v}{r} = 2(1 + \nu)\bar{\tau}_{r\theta}. \tag{18}$$

where κ and β are functions of the Poisson's ratio, ν; and $\kappa = 1 - \nu^2$, $\beta = \nu(1 + \nu)$ in the case of plane strain, and $\kappa = 1, \beta = \nu$ in the case of plane stress.

The displacements u and v can be found by integration of equations (16) and (17), and the integration constants may be determined by substitution into equation (18).

Some relevant expressions, $\bar{\sigma}_r$, ϵ_θ, $\bar{u}$, etc, are given in Appendix A, which are also shown in reference [5], but a few typographic errors in the reference are corrected.

Boundary Conditions. The boundary conditions at the inner and outer radii, R_i and R_0, of the hard-cored soft roll with the radial stress at the outer radius expressed by a Fourier cosine series are

$$\bar{u} = \bar{v} = 0, \quad \text{on} \quad \bar{r} = \bar{R}_i, \tag{19a}$$

$$\bar{\tau}_{r\theta} = 0, \quad \text{on} \quad \bar{r} = 1, \tag{19b}$$

and

$$\bar{\sigma}_r = \sum_{n=0}^{\infty} A_n \cos n\theta, \quad \text{on} \quad \bar{r} = 1; \quad 0 \le \theta \le \pi \tag{19c}$$

Substitution of stresses and displacements expressions into the boundary equations will relate the undefined stress function coefficients in equation (11) to the Fourier series coefficients, A_n, which are prescribed or related to the plate deflection. These equations are given in Appendix B.

Solution Equations. Assume that the radial stress at the contact zone is expressed by a Fourier series with $M + 1$ terms:

$$\bar{\sigma}_{r\langle\bar{r}=1\rangle} = \sum_{m=0}^{M} \frac{\pi}{\Omega} h_m \cos \frac{m\pi\theta}{\Omega}, \; 0 \le \theta \le \Omega.$$
$$= 0, \qquad \Omega \le \theta \le \pi. \tag{20}$$

Since the Fourier series of equation (19c) covers the entire range of θ, coefficients A_n can be related to h_m by the the following equations:

$$A_0 = h_0, \tag{21a}$$

and

$$A_n = \sum_{m=0}^{M} h_m \left(\frac{\sin(m\pi - n\Omega)}{m\pi - n\Omega} + \frac{\sin(m\pi + n\Omega)}{m\pi + n\Omega} \right)$$
$$n = 1, 2, \ldots, \infty. \tag{21b}$$

The radial stress at the contact zone will produce a radial displacement, $\bar{u}(\theta)$, within the contact zone. Thus, for prescribed $\bar{u}(\theta)$, we have

$$\sum_{n=0}^{\infty} \alpha_n A_n \cos n\theta = \bar{u}(\theta), \quad 0 \le \theta \le \Omega, \tag{22}$$

where the α's are obtained from the $\bar{u}(\theta)$ expression of equation (35), together with equations (38)–(41). The α's are given in Appendix C.

Differentiation of equation (22) yields

$$\sum_{n=1}^{\infty} n\alpha_n A_n \sin n\theta = -\frac{d\bar{u}}{d\theta}, \quad 0 \le \theta \le \Omega. \tag{23}$$

For the part of the contact arc which is circular, the displacement is,

$$\bar{u}_1 \cong \bar{u}_0 + \frac{1}{2}(1 + \bar{R}_h)\,\theta^2/\bar{R}_h, \quad 0 \le \theta \le \theta_p, \tag{24}$$

where $\bar{u}_0$ is the indentation, u_0, divided by R_0.

For the part in contact with the beam, the radial displacement is

$$\bar{u}_2 = (1 + \bar{u}_{1\langle\theta=\theta_p\rangle}) \cos(\theta - \theta_p) + \int_{\theta_p}^{\theta} \sin(\theta + \phi(\xi))d\xi - 1,$$
$$\theta_p \le \theta \le \Omega, \tag{25}$$

where ϕ is the solution of equation (6).

Substitution of equations (24), (25), and (21) into equation (23) and dropping $\bar{u}_1$ term, yield

$$\sum_{m=0}^{M} \sum_{n=1}^{\infty} n h_m \alpha_n \sin n\theta \left(\frac{\sin(m\pi - n\Omega)}{m\pi - n\Omega} + \frac{\sin(m\pi + n\Omega)}{m\pi + n\Omega} \right)$$
$$= -(1 + \bar{R}_h)\theta/\bar{R}_h \quad \text{for} \quad 0 \le \theta \le \theta_p,$$
$$= (1 + \bar{u}_{\langle\theta=\theta_p\rangle}) \sin(\theta - \theta_p) - \sin(\theta + \phi(\theta))$$
$$- \int_{\theta_p}^{\theta} \cos(\theta + \phi(\xi))d\xi \quad \text{for} \quad \theta_p \le \theta \le \Omega. \tag{26}$$

Note that equation (26) does not contain the indentation, u_0.

Equation (26) may be further simplified by using equation (9), which yields

$$h_0 = \sum_{m=1}^{M} (-1)^{m+1} h_m. \tag{27}$$

Substitution into equation (26) yields

$$\sum_{m=1}^{M} \sum_{n=1}^{\infty} n h_m \alpha_n \sin n\theta \left\{ (-1)^{m+1}\, 2 \sin n\Omega/n\Omega \right.$$
$$\left. + \frac{\sin(m\pi - n\Omega)}{m\pi - n\Omega} + \frac{\sin(m\pi + n\Omega)}{m\pi + n\Omega} \right\}$$
$$= -(1 + \bar{R}_h)\theta/\bar{R}_h \quad \text{for} \quad 0 \le \theta \le \theta_p,$$
$$= \sin(1 + \bar{u}_{\langle\theta=\theta_p\rangle})(\theta - \theta_p) - \sin(\theta + \phi(\theta))$$
$$- \int_{\theta_p}^{\theta} \cos(\theta + \phi(\xi))d\xi \quad \text{for} \quad \theta_p \le \theta \le \Omega. \tag{28}$$

This is the final simultaneous equation set to solve for the h_m which defines the radial deflection pattern of the elastic-layered cylinder for a given indentation, u_0/R_0. Since u_0 does not appear in equation (28), the two unknown angles, θ_p and Ω, will determine the deflection pattern. We could assume a value for one of the unknowns, say Ω, and solve for θ_p iteratively. In each iteration, the load $\bar{\sigma}_r$ on the beam is found by inserting the h_m into equation (20). The correct θ_p for the assumed Ω must satisfy the remaining constraining equation, equation (10). The corresponding u_0 which produces the assumed Ω is obtained from equation (24)

$$\bar{u}_0 = \bar{u}_{1\langle\theta=\theta_p\rangle} - \frac{1}{2}(1 + \bar{R}_h)\theta_p{}^2/\bar{R}_h. \tag{29}$$

Problem With Multilayers. By assuming a different stress function for each layer and using the additional stress-function coefficients to satisfy the stress and displacement requirements at the interfaces, the multilayered soft roll can be reduced to an equivalent single layer roll. Formulas will not be presented here, but some results pertaining to a two-layer roll with a hard core will be given.

The Collocation Method

To solve for h_m in the simultaneous equation (28), the Schmidt orthogonal function method could be used (reference [5]). However, it will require a large number of integrations and tedious numerical manipulations. Instead, we choose a simpler collocation method for the solution of equation (28). We simply divide the angle Ω into M segments which, for convenience, may be of equal arc length. By consecutively letting $i = 1, 2, \ldots, M$ in

$$\theta = \theta_i = i\Omega/M$$

and substitution into equation (28), M linear equations with constant coefficients and with constant right-hand sides are obtained. Then M values of h_m can be solved directly.

That the collocation method is much simpler is obvious. Take a 7-term solution as an example. The orthogonal function method will require 6 matrix inversions, 6 summations, and 28 integrations to get the values of the seven h_m's, whereas the collocation method needs only a single matrix inversion. Fig. 2 is a numerical example, based on the geometry of reference [5], of the contact pressure σ_r at the crown obtained by the two methods with the same number of h_m coefficients. A crown stress of 0.182 is obtained in reference [5].

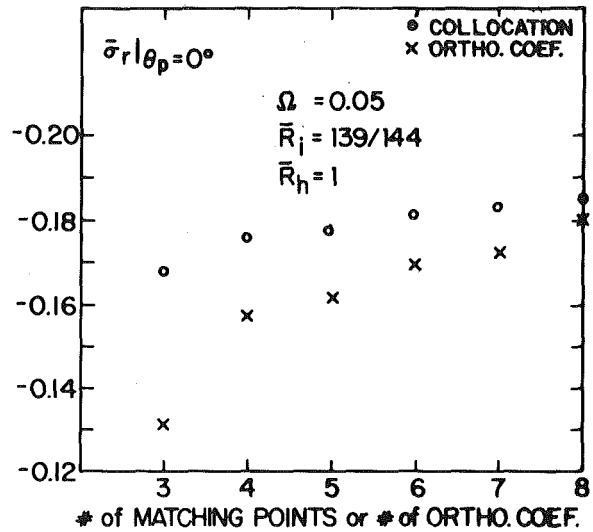


Fig. 2 Comparison between the orthogonal function method and the colocation method of equal number of terms for a problem in reference [5]

Compared with Fig. 2, the collocation method approaches this value quicker and in a steadier manner than the orthogonal coefficient method, and for equal number of terms, the collocation method is closer to the final value than the orthogonal coefficient method.

Some Numerical Results

Fig. 3 shows indentation versus contact angle for a curved beam and a straight beam in the nip. Fig. 4 shows effect of beam curvature and stiffness on contact angle for a given indentation. Fig. 5 shows the shear stresses at the interfaces of a two-layered cylinder with the presence of a curved beam in the nip.

Conclusions

An iterative procedure is suggested to solve the problems of an elastic plate with an initial curvature compressed between two cylinders, one is hard and the other has soft layers with a hard core. The method can be adopted easily to the same problem with the hard roll replaced by another elastic layered soft roll.

References

1 Wu, C. H., and Plunkett, R., "On the Contact Problem of Thin Circular Rings," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, Mar. 1965, pp. 11–20.

2 Callegari, A. J., and Keller, J. B., "Contact of Inflated Membranes With Rigid Surfaces," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, Mar. 1974, pp. 189–191.

3 Weitsman, Y., "On the Unbonded Contact Between Plates and An Elastic Half Space," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, June 1969, 198–202.

4 Hogg, A. H. A., "Equilibrium of a Thin Plate, Symmetrically Loaded, Resting on an Elastic Foundation of Infinite Depth," *Philosophical Magazine*, Vol. 25, 1938, pp. 576–582.

5 Hahn, H. T., and Levinson, M., "Indentation of an Elastic Layer(s) Bonded to a Rigid Cylinder—I. Quasistatic Case Without Friction," *International Journal of Mechanical Sciences*, Pergamon Press, 1974. Vol. 16, pp. 480–502.

6 Hahn, H. T., and Levinson, M., "Indentation of an Elastic Layer(s) Bonded to a Rigid Cylinder—II. Unidirectional Slipping With Coulomb Friction," *International Journal of Mechanical Sciences*, Pergamon Press, 1974. Vol. 16, pp. 503–514.

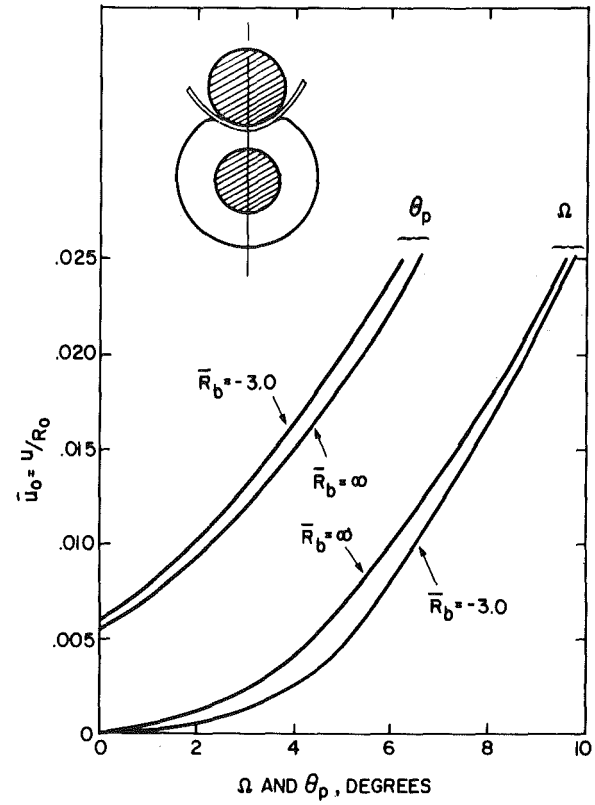


Fig. 3 Indentation versus contact angles for a curved beam and a straight beam; stiffness value $E^* R_0^3/EI = 5670$

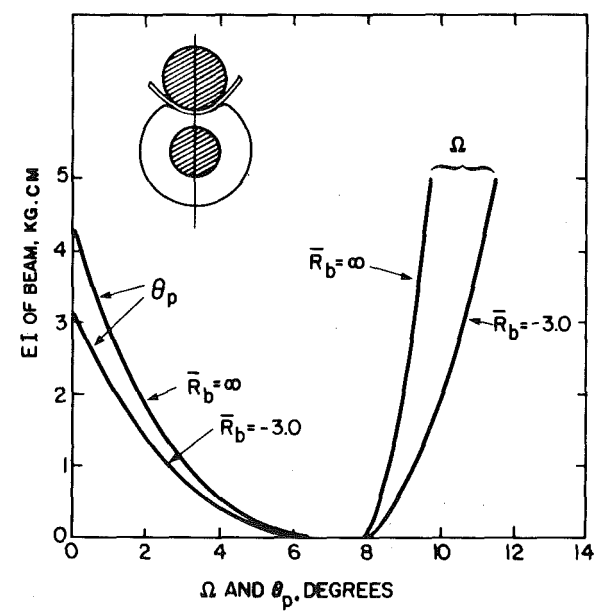


Fig. 4 Beam stiffness effect on contact angles for a fixed indentation $\bar{u}_0 = -0.0185$

APPENDIX *A*

Based on the stress function, equations (13)–(18) yield

$$\bar{\sigma}_r = \bar{a}_0\bar{r}^{-2} + \bar{b}_0 + \cos\theta\,[2\bar{r}\bar{b}_1 - 2\bar{r}^{-3}\bar{c}_1 + \bar{r}^{-1}(\bar{d}_1' + \bar{d}_1)] + \sum_{n=2}^{\infty}\cos n\theta\,[(n-n^2)\bar{r}^{n-2}\bar{a}_n + (2+n-n^2)\bar{r}^n\bar{b}_n - (n+n^2)\bar{r}^{-n-2}\bar{c}_n + (2-n-n^2)\bar{r}^{-n}\bar{d}_n] \quad (30)$$

$$\bar{\tau}_{r\theta} = \sin\theta(2\bar{r}\bar{b}_1 - 2\bar{r}^{-3}\bar{c}_1 + \bar{r}^{-1}\bar{d}_1') + \sum_{n=2}^{\infty}\sin n\theta\,[(n^2-n)r^{n-2}\bar{a}_n + (n^2+n)\bar{r}^n\bar{b}_n - (n^2+n)\bar{r}^{-n-2}\bar{c}_n + (n-n^2)\bar{r}^{-n}\bar{d}_n \quad (31)$$

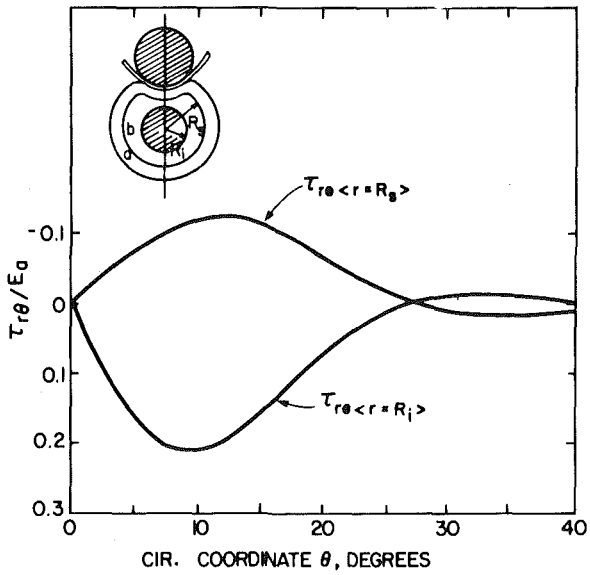


Fig. 5 Shear stress distribution at the two interfaces for a two-layered soft cylinder; indentation $\bar{u}_0 = 0.0185$, beam stiffness $EI = 0.25$ Kg-cm per unit width and curvature $\bar{R}_b = 3.0$

$$\epsilon_\theta = -(\kappa+\beta)\bar{r}^{-2}\bar{a}_0 + (\kappa-\beta)\bar{b}_0 + \cos\theta\,[2(3\kappa-\beta)\bar{r}\bar{b}_1 + 2\bar{r}^{-3}(\kappa+\beta)\bar{c}_1 + \bar{r}^{-1} \times ((\kappa-\beta)\bar{d}_1' - \beta\bar{d}_1)] + \sum_{n=2}^{\infty}\cos n\theta\,[n(n-1)(\kappa+\beta)\bar{r}^{n-2}\bar{a}_n + (\kappa(n^2+3n+2) - \beta(2+n-n^2))\bar{r}^n\bar{b}_n + (\kappa+\beta)(n+n^2)\bar{r}^{-n-2}\bar{c}_n + (\kappa(2-3n+n^2) + \beta(n^2+n-2))\bar{r}^{-n}\bar{d}_n] \quad (32)$$

$$\epsilon_r = (\kappa+\beta)\bar{r}^{-2}\bar{a}_0 + (\kappa-\beta)\bar{b}_0 + \cos\theta\,[2(\kappa-3\beta)\bar{r}\bar{b}_1 - (\kappa+\beta)2\bar{r}^{-3}\bar{c}_1 + \bar{r}^{-1}((\kappa-\beta)\bar{d}_1' + \kappa\bar{d}_1)] + \sum_{n=2}^{\infty}\cos n\theta\,[\bar{a}_n\bar{r}^{n-2}(n-n^2)(\kappa+\beta) + \bar{b}_n\bar{r}^n(\kappa(2+n-n^2) - \beta(2+3n+n^2)) - \bar{c}_n\bar{r}^{-n-2}(n^2+n)(\kappa+\beta) + \bar{d}_n\bar{r}^{-n}(\kappa(2-n-n^2) - \beta(2-3n+n^2))] \quad (33)$$

$$\bar{\sigma}_\theta = -\bar{r}^{-2}\bar{a}_0 + \bar{b}_0 + \cos\theta\,(6\bar{r}\bar{b}_1 + 2\bar{r}^{-3}\bar{c}_1 + \bar{r}^{-1}\bar{d}_1') + \sum_{n=2}^{\infty}\cos n\theta\,[n(n-1)\bar{r}^{n-2}\bar{a}_n + (n^2+3n+2)\bar{r}^n\bar{b}_n + (n^2+n)\bar{r}^{-n-2}\bar{c}_n + (2-3n+n^2)\bar{r}^{-n}\bar{d}_n]. \quad (34)$$

$$\bar{u}(\theta) = -(\kappa+\beta)\bar{r}^{-1}\bar{a}_0 + (\kappa-\beta)\bar{r}\bar{b}_0 + \cos\theta\,[\bar{a}_1 + (\kappa-3\beta)\bar{r}^2\bar{b}_1 + (\kappa+\beta)\bar{r}^{-2}\bar{c}_1 + ((\kappa-\beta)\bar{d}_1' + \kappa\bar{d}_1)\ln\bar{r}] + \sum_{n=2}^{\infty}\cos n\theta\,[-(\kappa+\beta)n\bar{r}^{n-1}\bar{a}_n + (\kappa(2-n) - \beta(2+n))\bar{r}^{n+1}\bar{b}_n + (\kappa+\beta)n\bar{r}^{-n-1}\bar{c}_n + (\kappa(2+n) + \beta(n-2))\,\bar{r}^{-n+1}\bar{d}_n], \quad (35)$$

and

$$\bar{v}(\theta) = \sin\theta\,[-\bar{a}_1 + (5\kappa+\beta)\bar{r}^2\bar{b}_1 + (\kappa+\beta)\bar{r}^{-2}\bar{c}_1 + (\kappa-\beta)\bar{d}_1' - \beta\bar{d}_1 - \ln\bar{r}\,((\kappa-\beta)\bar{d}_1' + \kappa\bar{d}_1)] + \sum_{n=2}^{\infty}\sin n\theta\,[(\kappa+\beta)n\bar{r}^{n-1}\bar{a}_n + (\kappa(n+4) + \beta n)\bar{r}^{n+1}\bar{b}_n + (\kappa+\beta)n\bar{r}^{-n-1}\bar{c}_n + (\kappa(n-4) + \beta n)\bar{r}^{-n+1}\bar{d}_n], \quad (36)$$

where $\bar{a}_1$ is an integration constant, and

$$\bar{d}_1' = \frac{\beta-\kappa}{4\kappa}\bar{d}_1 \quad (37)$$

is obtained by requirements for compatibility.

APPENDIX *B*

Relationships between stress function coefficients and the surface load coefficients of equation (19*c*)

$$\{\bar{a}_0, \bar{b}_0\} = \{(\kappa-\beta)\bar{R}_i, (\kappa+\beta)\bar{R}_i^{-1}\}\,((\kappa+\beta)\bar{R}_i^{-1} + (\kappa-\beta)\bar{R}_i)^{-1}A_0 \quad (38)$$

$$\bar{a}_1 = -\left[(\kappa-3\beta)\bar{R}_i^2 B_{13}^{-1} + (\kappa+\beta)\bar{R}_i^{-2}B_{23}^{-1} + \left(\kappa - \frac{(\kappa-\beta)^2}{4\kappa}\right)\ln\bar{R}_i B_{33}^{-1}\right]A_1 \quad (39)$$

$$\{\bar{b}_1, \bar{c}_1, \bar{d}_1\} = \{B_{13}^{-1}, B_{23}^{-1}, B_{33}^{-1}\}\,A_1, \quad (40)$$

$$\{\bar{a}_n, \bar{b}_n, \bar{c}_n, \bar{d}_n\} = \{c_{n14}^{-1}, c_{n24}^{-1}, c_{n34}^{-1}, c_{n44}^{-1}\}\,A_n, \quad n = 2, 3, \ldots, \infty, \quad (41)$$

where B_{ij}^{-1} and c_{nij}^{-1} denote the (ij)th elements of the matrices $[B]^{-1}$ and $[C_n]^{-1}$, respectively. Matrices $[B]^{-1}$ and $[C_n]^{-1}$ are given as follows:

$$[B]^{-1} = \begin{bmatrix} (6\kappa-2\beta)\bar{R}_i^2 & 2(\kappa+\beta)\bar{R}_i^{-2} & -\beta - \dfrac{(\kappa-\beta)^2}{4\kappa} \\ 2 & -2 & \dfrac{\beta-\kappa}{4\kappa} \\ 2 & -2 & 1 + \dfrac{\beta-\kappa}{4\kappa} \end{bmatrix}^{-1} \quad (42)$$

$$[C_n]^{-1} = \begin{bmatrix} -(\kappa+\beta)n\bar{R}_i^{n-1} & -[\kappa(n-2)+\beta(n+2)]\bar{R}_i^{n+1} & \dfrac{(\kappa+\beta)n}{\bar{R}_i^{n+1}} & \dfrac{\kappa(n+2)+\beta(n-2)}{\bar{R}_i^{n-1}} \\ (\kappa+\beta)n\bar{R}_i^{n-1} & [\kappa(n+4)+\beta n]\bar{R}_i^{n+1} & \dfrac{(\kappa+\beta)n}{\bar{R}_i^{n+1}} & \dfrac{\kappa(n-4)+\beta n}{\bar{R}_i^{n-1}} \\ n(n-1) & n(n+1) & -n(n+1) & -(n-1) \\ -n(n-1) & -(n-2)(n+1) & -n(n+1) & -(n+2)(n-1) \\ & n = 2, 3, \ldots, \infty & & \end{bmatrix}^{-1} \quad (43)$$

APPENDIX *C*

The coefficients of equation (22) are

$$\alpha_0 = \frac{(\kappa^2 - \beta^2)(-\overline{R}_i + \overline{R}_i^{-1})}{(\kappa + \beta)\overline{R}_i^{-1} + (\kappa - \beta)\overline{R}_i} \tag{44}$$

$$\alpha_1 = (\kappa - 3\beta)(1 - \overline{R}_i^2)B_{13}^{-1} + (\kappa + \beta)(1 - \overline{R}_i^{-2})B_{23}^{-1} + \left(\kappa - \frac{(\kappa - \beta)^2}{4\kappa}\right) \ln \overline{R}_i^{-1} B_{33}^{-1} \tag{45}$$

(*Cont.*)

$$\alpha_n = -(\kappa + \beta)nc_{n14}^{-1} - [\kappa(n - 2) + \beta(n + 2)]\, c_{n24}^{-1} + (\kappa + \beta)nc_{n34}^{-1} + [\kappa(n + 2) + \beta(n - 2)]c_{n44}^{-1} \qquad n = 2, 3, \ldots, \infty \tag{46}$$

H. P. Yagoda

Senior Technical Specialist,
AMF Inc.,
Morehead Patterson Center,
689 Hope Street,
Stamford, Conn. 06907
Mem. ASME

Resolution of a Core Problem in Wound Rolls

Previous misunderstandings about the behavior of stresses in the vicinity of the core of wound rolls are resolved through the development of an explicit closed-form analytical solution for the radial and circumferential stresses which are generated during winding; the solution holds in the general case of variable winding tension. Asymptotic series are derived and then employed to compute profiles of stresses within the wound material for several cases of practical interest. A parametric analysis of the influence of core elasticity on structural integrity of the roll underscores the indispensable support provided by the core at the inner boundary of the roll. Results show that the circumferential stress in the vicinity of the core strongly depends on core stiffness. A relation that is derived between interlayer pressure and circumferential stress at the core boundary may be useful as a guide in core design and in preventing roll collapse.

1 Introduction

Proper winding of flexible sheetlike materials into rolls is an area in the field of mechanics which has broad utility in a number of diverse and important industries and which, therefore, has been of considerable interest over the past few decades. For example, the winding process fulfills vital functions in the flexible packaging film industry [3], the magnetic tape industry [4, 5], the paper industry [6–9], and more recently, the tire retreading industry [10, 11].

Thus a central consideration of a number of analytical and experimental investigations has been the determination of stresses within a roll, e.g., [3, 6–9]. Nonetheless, the markedly sharp variations of the radial and circumferential stresses, which are shown in this paper to occur in the vicinity of the core, have been either overlooked or incorrectly calculated in these earlier investigations, e.g., [6, Fig. 1; 3, Fig. 5; 9, Fig. 4]. Experimental measurement has also failed to detect these sharp stress variations at the core, e.g., [7, 12]. As a consequence, apparently, the contribution by the core to structural cohesiveness of the roll has not been fully recognized. Indeed, it is shown in Section 6 herein, through a parametric analysis of core elasticity, that an inadequately designed core, i.e., one that is too soft to resist the compression of the material wound on it, will shift this intended function to the roll; thus high compressive hoop stress will develop in the surrounding material and the roll will be prone to buckling, with the likely formation of defects in the wound material.

A series solution is developed in this paper for the general case when winding tension may vary with roll growth, specifically as arbitrary non-negative real powers of the radius. The series converges rapidly; thus it provides an efficient method of evaluation that is less sensitive than the numerical integration employed in earlier theoretical investigations, e.g., [3, 4, 6], especially in the vicinity of the core.

2 The Altmann Integral Solution

Structural cohesiveness of wound rolls is developed through the building up of stressed layers within the wound material. A general incremental method of analysis of accreted bodies is given by Brown and Goodman [13]. By employing this incremental method, Altmann [6] derived integral expressions for the internal stresses in a center-wound roll[1] in the general case when the circumferential stress, σ_{T_w}, applied to the instantaneous outermost layer (the web) of the roll, may vary during winding. The initial radial pressure applied to the web, σ_{P_w}, is zero for center-wound rolls, e.g., Fig. 1(*a*). Nevertheless, σ_{P_w} is introduced in the following reformulation of Altmann's integrals for the radial pressure σ_P and the circumferential stress σ_T to facilitate generalization of the formulas to other winding configurations:

$$\sigma_P = \sigma_{P_w} + \left(\frac{1 + ar^{-2\gamma}}{r^b}\right) \int_r^R \left(\frac{s^b}{1 + as^{-2\gamma}}\right)\left(\frac{\sigma_{T_w}}{s}\right) ds \tag{1}$$

$$\sigma_T = \sigma_{T_w} - \left(\frac{\alpha - a\beta r^{-2\gamma}}{r^b}\right) \int_r^R \left(\frac{s^b}{1 + as^{-2\gamma}}\right)\left(\frac{\sigma_{T_w}}{s}\right) ds \tag{2}$$

The reader is referred to [6] for the notation used here. Briefly, $\bar{R}$, $\bar{c}$, $\bar{r}$, $\bar{s}$ are the dimensional radii (barred) of the finished roll, of the core, of a point within the finished roll, and of the web during winding ($\bar{r}$

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, December 1979; final revision, May, 1980. Paper No. 80-WA/APM-23.

[1] A center-wound roll is supported and driven entirely by the core against the tension in the web, e.g., Fig. 1(*a*).

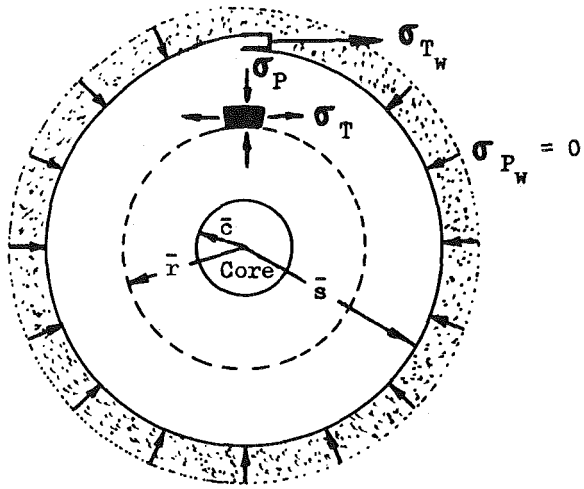


Fig. 1(*a*) Stresses applied to and generated within a center-wound roll during winding

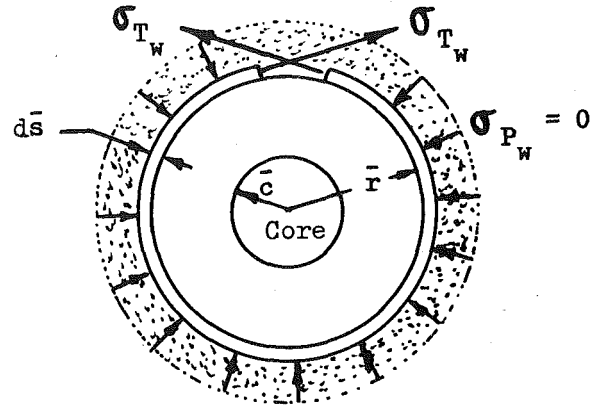


Fig. 1(*b*) Initial stresses at radial location $\bar{r}$ in a center-wound roll

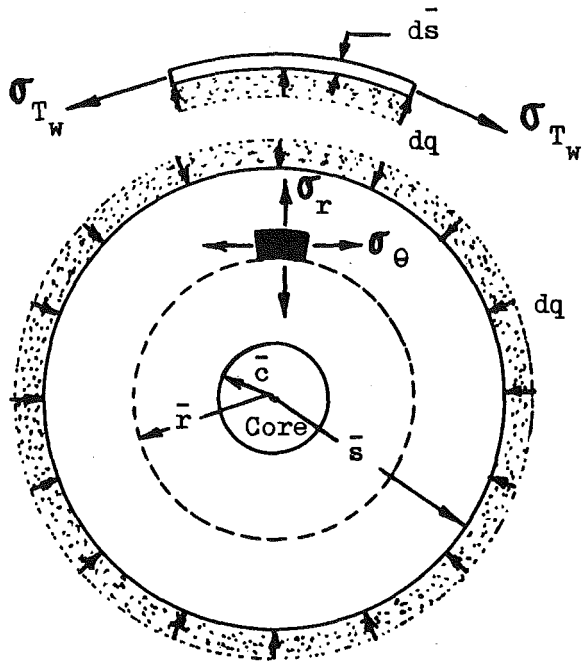


Fig. 1(*c*) Incremental stresses at radial location $\bar{r}$ which are induced when a layer is wound on at radial location $\bar{s}$ of a center-wound roll

$\leq \bar{s} \leq \bar{R}$), respectively. The corresponding nondimensional geometric quantities are defined as

$$R = \bar{R}/\bar{c}, \quad r = \bar{r}/\bar{c}, \quad s = \bar{s}/\bar{c} \tag{3}$$

Anisotropic elasticity parameters of the roll and web, and core elasticity are represented by E_r, E_t, μ_r, μ_t, and E_c. Based on these parameters, secondary dimensionless elasticity parameters are defined, respectively, for plane stress (unprimed) and plane strain (primed) as follows:

$$\begin{aligned}
\mu &= \frac{1}{2}\left(\mu_t + \frac{E_t}{E_r}\mu_r\right),\\
\delta &= \frac{1}{2}\left(\mu_t - \frac{E_t}{E_r}\mu_r\right),\\
\gamma &= \left|\sqrt{\delta^2 + \frac{E_t}{E_r}}\right|,\\
\alpha &= \gamma - \delta,\\
\beta &= \gamma + \delta,\\
b &= -(\alpha - 1)\\
a &= \frac{\gamma - \mu - \dfrac{E_t}{E_c}}{\gamma + \mu + \dfrac{E_t}{E_c}},\\
\mu' &= \mu/(1 - \mu_t)\\
\delta' &= \delta/(1 - \mu_t)\\
\gamma' &= |\sqrt{\delta'^2 + (1 - \mu_r\mu_t)E_t/(1 - \mu_t{}^2)E_r}|\\
\alpha' &= \gamma' - \delta'\\
\beta' &= \gamma' + \delta'\\
b' &= -(\alpha' - 1)\\
a' &= \frac{(1 - \mu_t{}^2)(\gamma' - \mu') - \dfrac{E_t}{E_c}}{(1 - \mu_t{}^2)(\gamma' + \mu') + \dfrac{E_t}{E_c}}
\end{aligned} \tag{4}$$

The derivation of equations (1) and (2) given in [6] is briefly reviewed here expressly to indicate how these equations can be extended to include general winding configurations. As mentioned earlier, the solution for the final state of stress may be formulated as an integral of infinitesimal increments, specifically,

$$\boldsymbol{\sigma}_P = \boldsymbol{\sigma}_{P_w} + \int_{\bar{r}}^{\bar{R}} \frac{d\boldsymbol{\sigma}_P}{d\bar{s}}\, d\bar{s}, \quad \boldsymbol{\sigma}_T = \boldsymbol{\sigma}_{T_w} + \int_{\bar{r}}^{\bar{R}} \frac{d\boldsymbol{\sigma}_T}{d\bar{s}}\, d\bar{s} \tag{5}$$

Fig. 1(*b*) shows the initial stresses, $\boldsymbol{\sigma}_{P_w}$ and $\boldsymbol{\sigma}_{T_w}$, acting on and in the layer at radial location $\bar{r}$, before additional plies are wound on. An incremental pressure, dq, is exerted on the underlying roll surface as each infinitesimal layer is added to the instantaneously outermost layer of the roll, which is shown at radial location $\bar{s}$ in Fig. 1(*c*). The loading dq will incrementally alter the radial and circumferential stresses at radial location $\bar{r}$ within the roll, thus

$$d\boldsymbol{\sigma}_P = -\boldsymbol{\sigma}_r = \left(\frac{1 + ar^{-2\gamma}}{r^b}\right)\left(\frac{s^b}{1 + as^{-2\gamma}}\right) dq,$$

$$d\boldsymbol{\sigma}_T = \boldsymbol{\sigma}_\theta = \left(\frac{a\beta r^{-2\gamma} - \boldsymbol{\alpha}}{r^b}\right)\left(\frac{s^b}{1 + as^{-2\gamma}}\right) dq \tag{6}$$

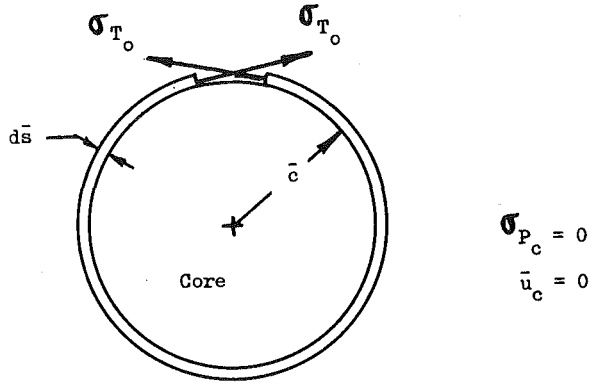


Fig. 1(*d*) Stress in the first layer before additional wraps are wound on

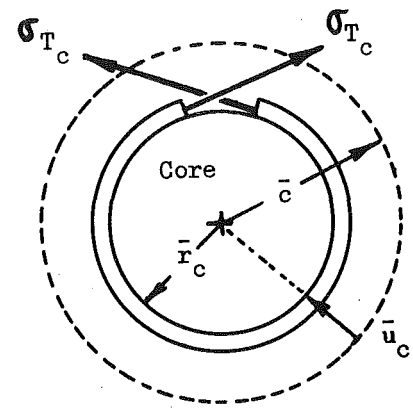


Fig. 1(*e*) Stress and deformation of the first layer after the last wrap is wound on

Equations (6),[2] which give the stresses in the anisotropic cylindrical continuum of Fig. 1(*c*), were derived in [6] through a displacement formulation of the following boundary-value problem:
The differential Equation of Stress Equilibrium

$$r\frac{d\sigma_r}{dr} + \sigma_r - \sigma_\theta = 0 \tag{7a}$$

The Anisotropic Stress-Strain and Strain-Displacement Relations

$$e_r = \frac{du}{dr} = \frac{\sigma_r}{E_r} - \mu_t\frac{\sigma_\theta}{E_t} \tag{7b}$$

$$e_\theta = \frac{u}{r} = \frac{\sigma_\theta}{E_t} - \mu_r\frac{\sigma_r}{E_r} \tag{7c}$$

where $u = \bar{u}/\bar{c}$ is the nondimensional displacement in the radial direction (positive if outward), and the boundary conditions,

$$\begin{aligned} \sigma_r &= E_c u, \quad \text{at } r = 1 \\ \sigma_r &= -dq \quad \text{at } r = s \end{aligned} \tag{7d}$$

If equations (5) are first transformed to nondimensional form and equations (6) then substituted therein, expressions for the stresses within the finished roll are obtained in terms of the incremental loading dq, i.e.,

$$\frac{\sigma_P}{\sigma_{T_0}} = \frac{\sigma_{P_w}}{\sigma_{T_0}} + \left(\frac{1 + ar^{-2\gamma}}{r^b}\right)\int_r^R \left(\frac{s^b}{1 + as^{-2\gamma}}\right)\left(\frac{dq/ds}{\sigma_{T_0}}\right) ds \tag{8}$$

$$\frac{\sigma_T}{\sigma_{T_0}} = \frac{\sigma_{T_w}}{\sigma_{T_0}} - \left(\frac{\alpha - a\beta r^{-2\gamma}}{r^b}\right)\int_r^R \left(\frac{s^b}{1 + as^{-2\gamma}}\right)\left(\frac{dq/ds}{\sigma_{T_0}}\right) ds \tag{9}$$

where σ_{T_0} is introduced here to render the stresses nondimensional and it is arbitrarily defined as the initial circumferential stress in the web at the start of winding, i.e., $\sigma_{T_0} = \sigma_{T_w}$ at $r = 1$. The center-wound rolls, Fig. 1(*c*), equilibrium of the forces acting on the web requires that

$$dq = \frac{\sigma_{T_w}}{\bar{s}}\, d\bar{s} = \frac{\sigma_{T_w}}{s}\, ds \tag{10}$$

If the average tension rate in the dislocating outer layers of the roll is restricted, equation (10) will also hold for other winding configuration, e.g., [9], for a description of two-drum winding. A derivation of the restriction on average tension rate will be given in a later paper. It is noted that the values of σ_{T_w} should now correspond to the location where slip between the outer layers and the roll ceases, and not at the outermost layer as described earlier. Thus a more general form of Altmann's integral solutions for the interlayer pressure σ_P and the circumferential stress σ_T within a wound roll is obtained by substituting equation (10) into equation (8) and equation (9):

$$\frac{\sigma_P}{\sigma_{T_o}} = \frac{\sigma_{P_w}}{\sigma_{T_0}} + \left(\frac{1 + ar^{-2\gamma}}{r^b}\right) I(r, R) \tag{11}$$

$$\frac{\sigma_T}{\sigma_{T_0}} = \frac{\sigma_{T_w}}{\sigma_{T_0}} - \left(\frac{\alpha - a\beta r^{-2\gamma}}{r^b}\right) I(r, R) \tag{12}$$

where

$$I(r, R) = \int_r^R \left(\frac{s^{b-1}}{1 + as^{-2\gamma}}\right)\left(\frac{\sigma_{T_w}}{\sigma_{T_0}}\right) ds \tag{13}$$

The integral (13) is transformed in Section 4 to a closed-form analytical expression in terms of hypergeometric functions. Before proceeding to the work in Section 4 however, a relation is first derived in Section 3 between the interlayer pressure and circumferential stress precisely at the core boundary. It is then demonstrated how this relation may be used to evaluate core design.

3 A Relation Between Pressure and Circumferential Stress at the Core

A condition may be derived, which holds precisely at the core interface, between the radial pressure σ_{P_c} and the circumferential stress σ_{T_c}. It is convenient, however, to first define the nondimensional radial displacement of the outer core surface, that is

$$u_c = \bar{u}_c/\bar{c} \tag{14}$$

where the corresponding dimensional quantity $\bar{u}_c$ is positive if inward. The inward radial displacement $\bar{u}_c$ and in-roll stresses at the core, i.e., for $r = 1$, are shown in Fig. 1(*d*) and Fig. 1(*e*) after the first and last ply have been wound on, respectively. We may express the compression of the outer core surface in terms of the core elasticity, thus

$$u_c = \sigma_{P_c}/E_c \tag{15}$$

The circumferential strain of the core can be expressed by combining the strain displacement equation (7*c*) and equation (15), thus

[2] Plane stress conditions were employed to derive equations (6) in [6]; these conditions are justified in the case of rolls of tape, e.g., [4, 5], or in the case of isotropic cylinders, e.g., [14, p. 70]. Plane strain conditions may be more appropriate in the case of anisotropic cylinders if Poisson's ratios are not small compared to unity. The stresses (6′) are obtained for plane strain conditions by a procedure similar to that used in the case of plane stress:

$$\sigma_r = -\left(\frac{1 + a'r^{-2\gamma'}}{r^{b'}}\right)\left(\frac{s^{b'}\, dq}{1 + a's^{-2\gamma'}}\right);$$

$$\sigma_\theta = \left(\frac{a'\beta' r^{-2\gamma'} - \alpha'}{r^{b'}}\right)\left(\frac{s^{b'}\, dq}{1 + a's^{-2\gamma'}}\right) \tag{6′}$$

Equations (6′) almost are identical to the corresponding stresses (6) for plane stress, except that the secondary elasticity parameters μ', δ', γ', α', β', b', a', defined in (4), replace the corresponding parameters in equation (6). Thus the formulas derived in this work for the radial and circumferential stresses hold either for plane stress or plane strain conditions, provided that the appropriate secondary elasticity parameters (4) are selected. It is noted that the solutions (6) and (6′) become identical as $\mu_r \to 0$ and $\mu_t \to 0$.

$$\Delta e_\theta]_{\text{core}} = -u_c = -\frac{\sigma_{P_c}}{E_c} \tag{16}$$

Employing the stress-strain equation (7c), the relieved strain in the first layer of material may be written as

$$\Delta e_\theta]_{\text{roll}} = \frac{\sigma_{T_c} - \sigma_{T_0}}{E_t} + \mu_r \frac{\sigma_{P_c}}{E_r} \tag{17}$$

Equality of expressions (16) and (17) assures geometric compatibility of core and roll deformations. After some mathematical manipulation, the following relation between internal stresses at the core is obtained:

$$\frac{\sigma_{T_c}}{\sigma_{T_0}} + \left(\mu_r \frac{E_t}{E_r} + \frac{E_t}{E_c}\right) \frac{\sigma_{P_c}}{\sigma_{T_0}} = 1 \tag{18}$$

It is interesting to observe, from equation (18), that relaxation of the initially generated circumferential stress at the core, σ_{T_0}, evolves by interaction of the radial pressure with the circumferential strain through two different elastic effects. One of these effects stems from the decrease in circumferential strain which is directly attributable to core compliance (E_c). The other effect stems from the resisting radial pressure which would be induced by a decrease in circumferential strain, through Poisson's ratio μ_r. Thus some relaxation of circumferential stress at the core should be expected even in the extreme case of an absolutely rigid core, i.e., $E_c \to \infty$, except if $\mu_r = 0$.

Relation (18) may be employed to develop guides for core design. We can, for example, deduce a condition from equation (18) which provides that the circumferential stress will not become compressive (negative tension) at the core. This condition on the core pressure is obtained by setting $\sigma_{T_c} \geqq 0$ in equation (18), i.e.,

$$\frac{\sigma_{P_c}}{\sigma_{T_0}} \leqq \frac{1}{\mu_r \dfrac{E_t}{E_r} + \dfrac{E_t}{E_c}} \tag{19}$$

Condition (19) may be used to assess whether the elastic stiffness of the core is adequate in cases when σ_{P_c} can be estimated by measurement or is approximately known *a priori*, as in the case of Fig. 3(b). Upper limits (19) on the core pressure correspond to the intersections on the abscissa of the lines of constant core elasticity in Fig. 2, which was constructed by employing equation (18) for the specific case of a cellophane film whose properties are described subsequently. Elastic properties of the 140V5 saran-coated cellophane film are given in [3], namely,

$$E_t = 168{,}000 \text{ psi} = 1158 \text{ MPa}$$

$$E_r = \quad 5{,}000 \text{ psi} = 34.47 \text{ MPa}$$

$$\mu_r = \mu_t = 0.10 \tag{20}$$

For convenience, the corresponding secondary nondimensional elasticity parameters defined for plane stress conditions in [6] and (4) are listed in Table 1. Fig. 2 shows the effect of core elasticity E_c on the relation between circumferential stress and interface pressure at the core. A solution is taken apriori from Section 6 herein and plotted in Fig. 2 for the cellophane roll, which has been center-wound at constant web tension to a diameter of 9.5 in. (24.13 cm) on a 3.4 in. (8.64 cm) diameter core. It can be seen that the circumferential stress becomes compressive at the core for values of $E_c < 64{,}000$ psi (441.3 MPa) approximately. The solution to be developed subsequently in Section 4 and Fig. 4(a) shows, in fact, that the minimum roll circumferential stress will shift from within the roll to the core interface for values of $E_c < 40{,}000$ psi (275.8 MPa), approximately. Of course, a core that is too soft will lead to collapse of the roll.

4 An Explicit Closed-Form Solution

The integral $I(r, R)$ defined by equation (13) can be evaluated analytically in closed form in those cases when the applied winding stress σ_{T_w} is, or may be, expressed in powers of the dimensionless radius s, i.e.,

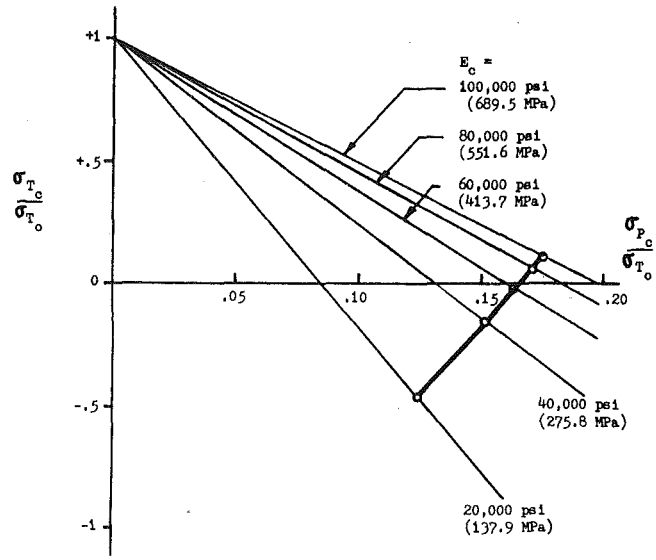


Fig. 2 Core elasticity influence on the relation between stresses at the core for 140 *V* cellophane film, e.g., equation (20)

Table 1 Nondimensional secondary elasticity parameters of equation (4) for the 33-lb newsprint roll of [6, 7], e.g., equation (43), and the 140 *V*5 saran-coated cellophane roll of [3], e.g., equation (20)

Non-Dimensional Secondary Elasticity Parameters (Plane Stress)	33-lb Newsprint Roll	140V5 Cellophane Roll				
		E_c, Core Elasticity				
		100,000 psi* (689.5 MPa)	80,000 psi (551.6 MPa)	60,000 psi (413.7 MPa)	40,000 psi (275.8 MPa)	20,000 psi (137.9 MPa)
μ	4.0050	.....		1.7300	.....	
δ	-3.9950	.....		-1.6300	.....	
γ	28.5650	.....		6.0214	.....	
α	32.560	.....		7.6514	.....	
β	24.570	.....		4.3914	.....	
b	-31.560	.....		-6.6514	.....	
a	.40818	.27688	.22244	.14134	.0076452	-.25438

*Reference [3] assumed this value of core elasticity

$$\frac{\sigma_{T_w}}{\sigma_{T_0}} = \sum_{j=0}^{M} C_j s^{\phi_j} \tag{21}$$

where $\phi_j \geqq 0$ and C_j are arbitrary constants. If equation (21) is substituted into equation (13) and the orders of integration and summation are interchanged, $I(r, R)$ may be written as a summation of $M + 1$ integrals, namely,

$$I(r, R) = \sum_{j=0}^{M} I_j(r, R, \phi_j) \tag{22}$$

where the integrals $I_j(r, R, \phi_j)$ are of the following general form:

$$I_j(r, R, \phi_j) = C_j \int_r^R \frac{s^{-\xi_j}}{1 + as^{-2\gamma}} ds \tag{23}$$

and

$$\xi_j = 1 - b - \phi_j \tag{24}$$

The integrals (23) can be transformed by the change of variable

$$\eta = (s/R)^{2\gamma} \tag{25}$$

to

$$I_j(r, R, \phi_j) = \frac{C_j R^{2\gamma - \xi_j + 1}}{2\gamma a} \int_{(r/R)^{2\gamma}}^{1} \frac{\eta^{-(\xi_j - 1)/2\gamma}}{1 + (R^{2\gamma}/a)\eta} d\eta \tag{26}$$

or

$$I_j(r, R, \phi_j) = \frac{C_j R^{2\gamma-\xi_j+1}}{2\gamma a} \times \left[\int_0^1 \frac{\eta^{-(\xi_j-1)/2\gamma}}{1+(R^{2\gamma}/a)\eta} d\eta - \int_0^{(r/R)^{2\gamma}} \frac{\eta^{-\xi_j-1/2\gamma}}{1+(R^{2\gamma}/a)\eta} d\eta\right] \tag{27}$$

We can now employ the formula given in [15, 3.194(1)] to integrate $I_j(r, R, \phi_j)$ in terms of hypergeometric functions, specifically,

$$I_j(r, R, \phi_j) = \frac{C_j R^{2\gamma-\xi_j+1}}{(2\gamma-\xi_j+1)\, a}\left\{{}_2F_1\left(1,1-\frac{\xi_j-1}{2\gamma}; 2-\frac{\xi_j-1}{2\gamma}; -\frac{R^{2\gamma}}{a}\right) - \left(\frac{r}{R}\right)^{2\gamma-\xi_j+1} {}_2F_1\left(1,1-\frac{\xi_j-1}{2\gamma}; 2-\frac{\xi_j-1}{2\gamma}; -\frac{R^{2\gamma}}{a}\left[\frac{r}{R}\right]^{2\gamma}\right)\right\} \tag{28}$$

which holds for $(\xi_j - 1)/2\gamma < 1$. Equations (11), (12), (22), (23), and (28) comprise the closed-form analytical solution for the interlayer pressure and circumferential stresses within a roll wound with a general applied web tension (21). In the special case of constant web tension, for which $M = 0$, $\phi_0 = 0$, $C_0 = 1$, $\xi_0 = 1 - b$, $\sigma_{T_w} \equiv \sigma_{T_0}$, the internal stresses reduce to the following expressions:

$$\frac{\sigma_P}{\sigma_{T_0}} = \frac{\sigma_{P_w}}{\sigma_{T_0}} + \left(\frac{1+ar^{-2\gamma}}{r^b}\right) I_0(r, R, 0) \tag{29}$$

$$\frac{\sigma_T}{\sigma_{T_0}} = 1 - \left(\frac{\alpha - a\beta r^{-2\gamma}}{r^b}\right) I_0(r, R, 0) \tag{30}$$

where

$$I_0(r, R, 0) = \frac{R^{2\gamma+b}}{(2\gamma+b)\, a}\left\{{}_2F_1\left(1,1-\frac{-b}{2\gamma}; 2-\frac{-b}{2\gamma}; -\frac{R^{2\gamma}}{a}\right) - \left(\frac{r}{R}\right)^{2\gamma+b} {}_2F_1\left(1,1-\frac{-b}{2\gamma}; 2-\frac{-b}{2\gamma}; -\frac{R^{2\gamma}}{a}\left[\frac{r}{R}\right]^{2\gamma}\right)\right\} \tag{31}$$

5 Asymptotic Solutions

Values of the secondary elasticity constants (e.g., Table 1) for the cellophane film and newspaper rolls which are considered in this work typically lead to hypergeometric functions in equation (28) with large negative arguments. The behavior of ${}_2F_1(a,b;c;z)$ for large $|z|$ is described by the transformation formulas for hypergeometric functions, e.g., [16]. One of these formulas [16, 15.3.7] is employed in [1] to obtain a useful series representation for these functions, namely,

$${}_2F_1(1,b;b+1;-z) = \Gamma(b+1)\Gamma(1-b)\frac{1}{z^b} + \left(\frac{b}{b-1}\right)\frac{1}{z}\sum_{N=0}^{\infty}\frac{(-1)^N(1-b)}{(1-b+N)}\frac{1}{z^N} \tag{32}$$

Equation (32) is central to an orderly derivation of an asymptotic expansion of $I_j(r, R, \phi_j)$, equation (28). For the practical cases considered in the next Section 6,

$$|R^{2\gamma}/a| \gg 1 \quad \text{and} \quad \left(\frac{r}{R}\right)^{2\gamma} \leqq 1; \tag{33a}$$

it will be found, however, that

$$\left|\frac{R^{2\gamma}}{a}\left(\frac{r}{R}\right)^{2\gamma}\right| > 1, \tag{33b}$$

even at the core where its minimum value $1/a > 2.44$. Thus we can employ equation (32) to write the following two expressions:

$${}_2F_1\left(1,1-\frac{\xi_j-1}{2\gamma}; 2-\frac{\xi_j-1}{2\gamma}; -\frac{R^{2\gamma}}{a}\right) = \frac{\Gamma\left(2-\frac{\xi_j-1}{2\gamma}\right)\Gamma\left(\frac{\xi_j-1}{2\gamma}\right)}{\left(\frac{R^{2\gamma}}{a}\right)^{[1-(\xi_j-1)/2\gamma]}} + \left(\frac{1-\frac{\xi_j-1}{2\gamma}}{-\frac{\xi_j-1}{2\gamma}}\right)\frac{1}{\frac{R^{2\gamma}}{a}}\sum_{N=0}^{\infty}\frac{(-1)^N\left(\frac{\xi_j-1}{2\gamma}\right)}{\left(N+\frac{\xi_j-1}{2\gamma}\right)}\frac{1}{\left(\frac{R^{2\gamma}}{a}\right)^N} \tag{34a}$$

$$\left(\frac{r}{R}\right)^{2\gamma-\xi_j+1} {}_2F_1\left(1,1-\frac{\xi_j-1}{2\gamma}; 2-\frac{\xi_j-1}{2\gamma}; -\frac{R^{2\gamma}}{a}\left[\frac{r}{R}\right]^{2\gamma}\right) = \frac{\Gamma\left(2-\frac{\xi_j-1}{2\gamma}\right)\Gamma\left(\frac{\xi_j-1}{2\gamma}\right)}{\left(\frac{R^{2\gamma}}{a}\right)^{[1-(\xi_j-1)/2\gamma]}} + \left(\frac{1-\frac{\xi_j-1}{2}}{-\frac{\xi_j-1}{2\gamma}}\right)\frac{\left(\frac{r}{R}\right)^{1-\xi_j}}{\frac{R^{2\gamma}}{a}} \times \sum_{N=0}^{\infty}\frac{(-1)^N\left(\frac{\xi_j-1}{2\gamma}\right)}{\left(N+\frac{\xi_j-1}{2\gamma}\right)}\frac{1}{\left[\frac{R^{2\gamma}}{a}\left(\frac{r}{R}\right)^{2\gamma}\right]^N} \tag{34b}$$

The leading order terms will cancel on taking the difference of equations (34). Therefore, if the first terms are extracted from both summations in equations (34) in expectation of the cancellation, the series for $I_j(r, R, \phi_j)$ is obtained by substituting equation (34) into equation (28), specifically,

$$r^{\xi_j-1} I_j(r, ifR, \phi_j) = C_j\left\{\frac{1}{\xi_j-1}\left[1-\left(\frac{r}{R}\right)^{\xi_j-1}\right] - \frac{r^{2\gamma}}{2\gamma a}\sum_{n=2}^{\infty}\frac{(-1)^n}{\left(n-1+\frac{\xi_j-1}{2\gamma}\right)}\frac{1}{\left[\frac{R^{2\gamma}}{a}\left(\frac{r}{R}\right)^{2\gamma}\right]^n} + \frac{R^{2\gamma-\xi_j+1}}{2\gamma a r^{-(\xi_j-1)}}\sum_{n-2}^{\infty}\frac{(-1)^n}{\left(n-1+\frac{\xi_j-1}{2\gamma}\right)}\frac{1}{\left(\frac{R^{2\gamma}}{a}\right)^n}\right\} \tag{35}$$

where the summation variable has been changed to $n = N + 1$ and $\xi_j \neq 1$. The last summation in equation (35) typically will be of higher order for the practical cases considered in Section 6 and thus can be neglected. The abbreviated notation ξ_j, introduced earlier for the sake of convenience, is now eliminated from equation (35) by equation (24), thus

$$\frac{1}{r^b} I_j(r,R,\phi_j) = C_j r^{\phi_j}\left\{\frac{1}{-(b+\phi_j)}\left[1-\left(\frac{r}{R}\right)^{-(b+\phi_j)}\right] - \frac{r^{2\gamma}}{2\gamma a}\sum_{n=2}\frac{(-1)^n}{\left(n-1-\frac{b+\phi_j}{2\gamma}\right)}\frac{1}{\left[\frac{R^{2\gamma}}{a}\left(\frac{r}{R}\right)^{2\gamma}\right]^n}\right\} \tag{36}$$

where higher-order terms are omitted. By combining equations (11), (12), (22), (23), (24), and (36), the following asymptotic series solution is obtained, which provides a convenient and accurate method to compute the internal stresses generated in rolls wound under a general web tension (21):

$$\frac{\sigma_P}{\sigma_{T_0}} = \frac{\sigma_{P_w}}{\sigma_{T_0}} + (1+ar^{-2\gamma})\sum_{j=0}^{M} S(r, R, \phi_j)\, C_j\, r^{\phi_j} \tag{37}$$

$$\frac{\sigma_T}{\sigma_{T_0}} = \frac{\sigma_{T_w}}{\sigma_{T_0}} - (\alpha - a\beta r^{-2\gamma})\sum_{j=0}^{M} S(r, R, \phi_j)\, C_j\, r^{\phi_j} \tag{38}$$

where

$$S(r, R, \phi_j) = \frac{1}{-(b+\phi_j)}\left[1-\left(\frac{r}{R}\right)^{-(b+\phi_j)}\right] \tag{39}$$

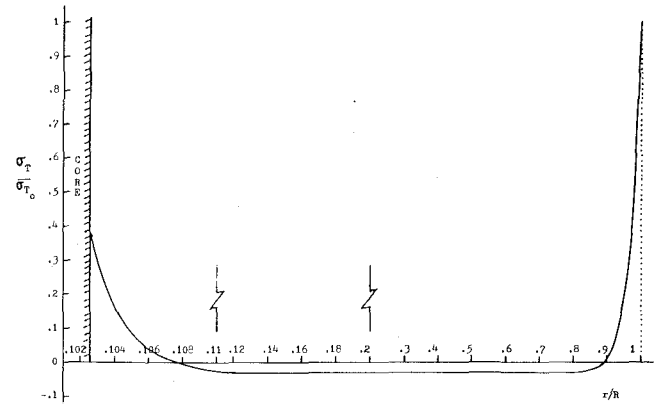


Fig. 3(*a*) Circumferential stress profile for the 33-lb newsprint roll, e.g., equation (43)

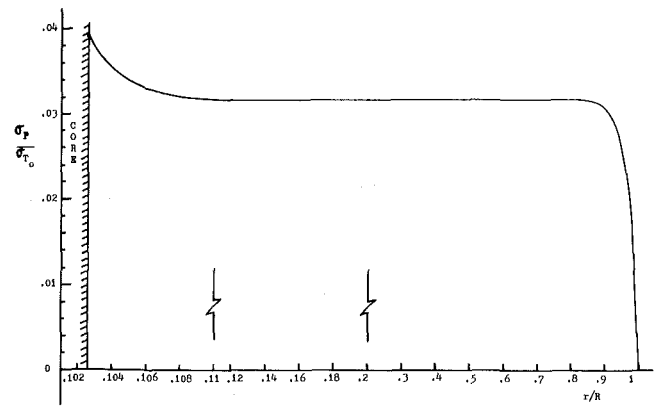


Fig. 3(*b*) Pressure profile for the 33-lb newsprint roll, e.g., equation (43)

$$-\frac{r^{2\gamma}}{2\gamma a}\sum_{n=2}\frac{(-1)^n}{\left(n-1-\dfrac{b+\phi_j}{2\gamma}\right)}\frac{1}{\left[\dfrac{R^{2\gamma}}{a}\left(\dfrac{r}{R}\right)^2\right]^n} \tag{39}$$

(*Cont.*)

and $b + \phi_j \neq 0$. It may be appropriate to indicate here that the major effect of core elasticity stems from the summation indicated in equation (39). The first term of (39) is independent of core elasticity and typically is the leading order term outside the vicinity of the core region; it may be of interest to note that this first term may be obtained by elementary integration [1] if the denominator of integral (23) is assumed to be unity.[3] The form of the foregoing asymptotic solution may be simplified in two special cases of particular interest. In the case of constant web tension, i.e., $M = 0$, $\phi_0 = 0$, $C_0 = 1$, $\sigma_{T_w} = \sigma_{T_0}$, the summations in equations (37) and (38) reduce to a single term, i.e.,

$$\frac{\sigma_P}{\sigma_{T_0}} = \frac{\sigma_{P_w}}{\sigma_{T_0}} + (1 + ar^{-2\gamma})\, S(r, R, 0),$$

$$\frac{\sigma_T}{\sigma_{T_0}} = 1 - (\alpha - a\beta r^{-2\gamma})\, S(r, R, 0) \tag{40}$$

where

$$S(r, R, 0) = \frac{1}{-b}\left[1 - \left(\frac{r}{R}\right)^{-b}\right] - \frac{r^{2\gamma}}{2\gamma a}\sum_{n=2}\frac{(-1)^n}{\left(n-1+\dfrac{-b}{2\gamma}\right)}\frac{1}{\left[\dfrac{R^{2\gamma}}{a}\left(\dfrac{r}{R}\right)^{2\gamma}\right]^n} \tag{41}$$

The asymptotic solution again simplifies if $|\phi_j/b| \ll 1$; under these conditions, the function $S(r, R, \phi_j)$ will not significantly depend on ϕ_j, which may be relevant in cases with the typical values of $|b|$ given in Table 1. In this case, $S(r, R, \phi_j)$ may be extracted from the summations in equations (37) and (38), which then sum to the applied web stress (21); consequently, the form of the solution reduces, namely, to[4]

$$\frac{\sigma_P}{\sigma_{T_w}} \approx \frac{\sigma_{P_w}}{\sigma_{T_w}} + (1 + ar^{-2\gamma})\, S(r, R, 0),$$

$$\frac{\sigma_T}{\sigma_{T_w}} \approx 1 - (\alpha - a\beta r^{-2\gamma})\, S(r, R, 0) \tag{42}$$

Results for several cases of cellophane film and newsprint rolls wound at constant web tension are discussed in the next section.

6 Internal Stress Profiles

Employing the asymptotic series derived in Section 5, circumferential and radial stress profiles are constructed[5] under plane stress conditions in Figs. 3(*a*, *b*), respectively, for the 33-lb newsprint roll of 0.003 in. (0.00762 cm) caliper (layer thickness) in [6, 7], which was center-wound at constant web tension to a diameter of 39 in. (99.06 cm) on a 4 in. (10.16 cm) diameter core. Elastic properties of the roll and core are given in [6] and listed here for convenience:

$$E_t = 800{,}000 \text{ psi} = 5516 \text{ MPa} \qquad E_r = 1000 \text{ psi} = 6.895 \text{ MPa}$$
$$E_c = 100{,}000 \text{ psi} = 689.5 \text{ MPa}, \qquad \mu_r = \mu_t = 0.01 \tag{43}$$

Note that the abscissae of Figs. 3 have been discontinuously scaled to illustrate the extremely narrow region of tension bands which surround the core in the particular case of this newspaper roll. Despite the diminutive size of this region, the core provides, notwithstanding, indispensable support at the inner boundary of the roll. It resists the inward crush and simultaneously prevents buckling of the innermost plies. Fig. 3(*a*) shows, in fact, that the newsprint core is sufficiently stiff to retain nearly 40 percent of the initial tension of the innermost ply. Indeed, a core that is too soft would burden the roll with this function of the core, leading to high circumferential compression and, thereby, to instability of the inner layers of material.

The preceding remarks are effectively demonstrated through a parametric analysis of the effect of core elasticity on stresses within the roll. The dependence on core elasticity of stresses in the vicinity of the core is constructed under plane stress conditions in Fig. 4 for the roll of 140V5 saran-coated cellophane film, e.g., equation (20) and [3]. In particular, Figs. 4(*a*, *b*) show that as core elasticity decreases, a relatively modest decrease in core pressure is compensated by the development of significant circumferential compression of the inner plies. For example, the negative tension reaches nearly half of the

[3] This assumption does not hold at the core, i.e., when $s = 1$ and $1 + as^{-2\gamma} = 1 + a$.

[4] The equations for the interlayer pressure and circumferential stress given in [9] may be derived from case (42) if the effect of core elasticity is arbitrarily neglected (by omitting the summation) in equation (41) and $ar^{-2\gamma}$ and $a\beta r^{-2\gamma}$ in equation (42) are assumed small compared to unity and α, respectively; thus

$$\frac{\sigma_P}{\sigma_{T_w}} = \frac{1}{-b}\left[1 - \left(\frac{r}{R}\right)^{-b}\right]; \quad \frac{\sigma_T}{\sigma_{T_w}} = 1 - \frac{\alpha}{-b}\left[1 - \left(\frac{r}{R}\right)^{-b}\right] \approx \frac{1}{-b}\left[(1-b)\left(\frac{r}{R}\right)^{-b} - 1\right]$$

Note that the nomenclature used here is different from that of [9] and that the foregoing equations become identical to [9, equations (4) and (5)] if the symbols are properly reconciled and if it is further assumed that $1 - b \approx \sqrt{E_t/E_r}$. Note however that this solution does not hold in the vicinity of the core.

[5] Approximately 11 terms of the summation in equation (41) are required at the core ($r/R = 0.10256$) to secure four significant figures; three significant figures are obtained after 8 terms.

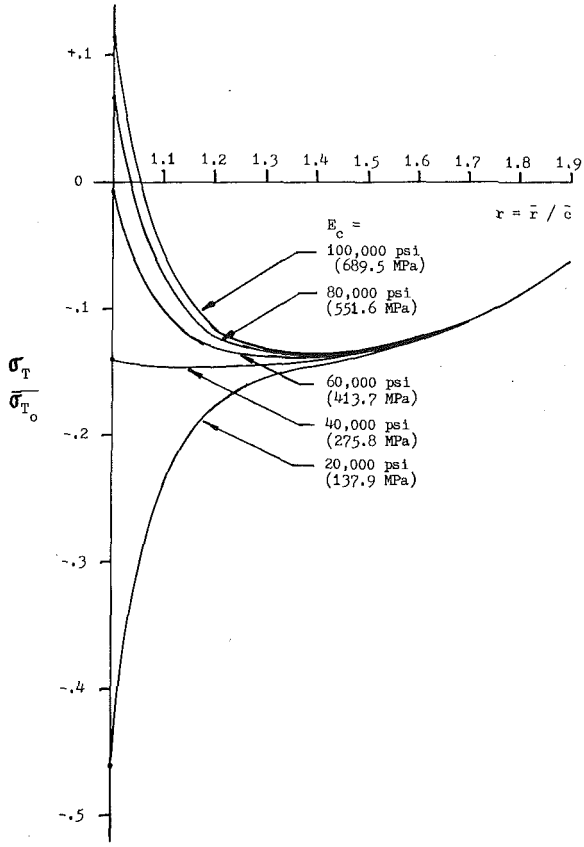


Fig. 4(*a*) Core elasticity influence on circumferential stress in the vicinity of the core for 140 *V* cellophane film, e.g., equation (20)

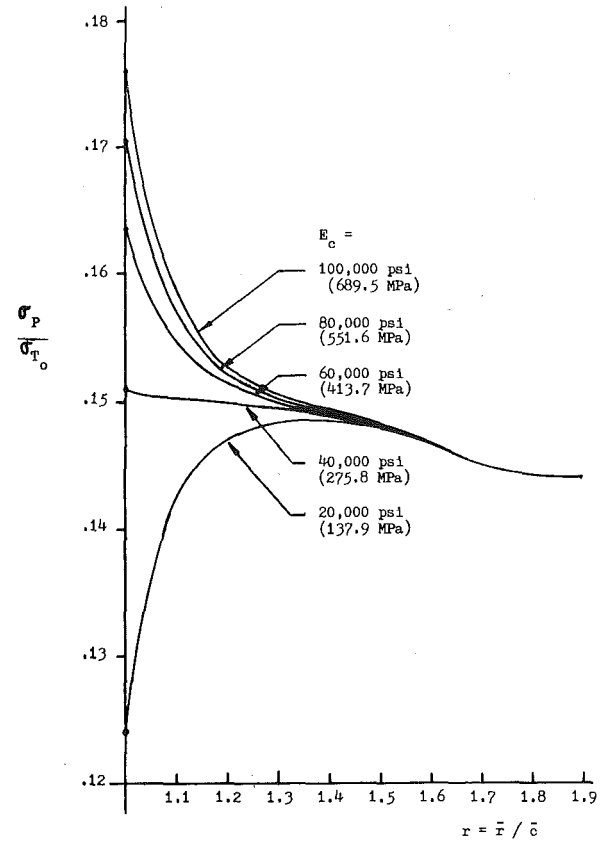


Fig. 4(*b*) Core elasticity influence on pressure in the vicinity of the core for 140 *V* cellophane film, e.g., equation (20)

initial wound-on tension in the case when E_c = 20,000 psi (137.9 MPa). The comparative insensitivity of the core pressure and, conversely, the pronounced dependence of the circumferential stress on core elasticity may be observed by inspection of Fig. 4(*c*). As mentioned previously, the behavior described here, in conjunction with relation (18) and Fig. 2, provide a guide for core design; specifically, an appropriate value of E_c may be estimated from Fig. 2 if core pressure is known *a priori,* or can be estimated either by measurement or an approximate analysis, i.e., the leading order term of equation (36).

Before concluding this section, it is noted that the stress profiles computed in the parametric study of Fig. 4 progressively converge with increasing radial distance from the core. No appreciable differences in the curves are discernible at $r = \bar{r}/\bar{c} = 1.7$, which corresponds to a location approximately 39 percent within the thickness of wound material, i.e., $0.39(R\text{-}c)$.

7 Summary

An explicit closed-form analytical solution is derived in terms of hypergeometric functions for the radial and circumferential stresses within rolls which may be wound under variable winding tension. With appropriately selected parameters, the solution holds for plane stress or plane strain conditions and for general winding configurations, i.e., center or surface-wound rolls. Employing well-known relations and series for the hypergeometric function, an asymptotic series solution is developed for the general case when winding tension may vary with roll growth, specifically as arbitrary non-negative real powers of the radial location. The series converges rapidly and thus facilitates computation of internal stress profiles that previously were obtained by numerical integration.

A parametric analysis of the influence of core elasticity on structural integrity of the roll underscores the indispensable support provided

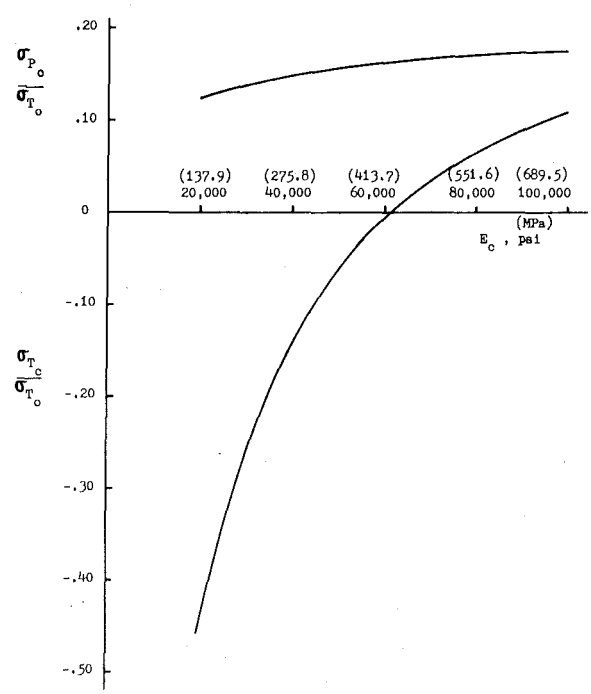


Fig. 4(*c*) Dependence on core elasticity of stresses at the core for 140 *V* cellophane film, e.g., equation (20)

by the core at the inner boundary of the roll. It is seen that the circumferential stress in the vicinity of the core markedly depends on core elasticity; for example, high compressive hoop stresses are generated in the material surrounding an excessively compliant core. A relation is derived between interlayer pressure and circumferential stress precisely at the core, which may be used as a guide to core design.

Acknowledgment

The author appreciates some helpful discussions with John Dundurs of Northwestern University.

References

1 Yagoda, H. P., "Integral Formulas for Wound Rolls," *Mechanics Research Communications,* Vol. 7, No. 2, 1980, pp. 103–113.

2 Yagoda, H. P., "Centrifugally Induced Stresses Within Center-Wound Roll-Part I," *Mechanics Research Communications,* Vol. 7, No. 3, 1980, pp. 181–193.

3 Monk, D. W., Lautner, W. K., and McMullen, J. F., "Internal Stresses Within Rolls of Cellophane," *The Journal of the Technical Association of the Pulp and Paper Industry,* Vol. 58, No. 8, Aug. 1975, pp. 152–155.

4 Tramposch, H., "Anisotropic Relaxation of Internal Forces in a Wound Reel of Magnetic Tape," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, Dec. 1967, pp. 888–894.

5 Tramposch, H., "Relaxation of Internal Forces in a Wound Reel of Magnetic Tape," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, Dec. 1965, pp. 865–873.

6 Altmann, H. C., "Formulas for Computing the Stresses in Center-Wound Rolls," *The Journal of the Technical Association of the Pulp and Paper Industry,* Vol. 51, No. 4, Apr. 1968, pp. 176–179.

7 Pfeiffer, J. D., "Internal Pressures in a Wound Roll of Paper," *The Journal of the Technical Association of the Pulp and Paper Industry,* Vol. 49, No. 8, Aug. 1966, pp. 342–347.

8 Frye, K. G., "Winding Variables and Their Effect on Roll Hardness and Roll Quality," *The Journal of the Technical Association of the Pulp and Paper Industry,* Vol. 50, No. 7, July 1967, pp. 81A–86A.

9 Rand, T., and Eriksson, L. G., "Physical Properties of Newsprint Rolls During Winding," *The Journal of the Techincal Association of the Pulp and Paper Industry,* Vol. 56, No. 6, June 1973, pp. 153–156.

10 Wireman, J., "Strip Winding of Tread Rubber," *Rubber World,* Vol. 169, No. 2, Nov. 1973, pp. 33–40.

11 Markey, W. F., "Retread Production Methods and Equipment," Akron Rubber Group Technical Symposium, Fall 1975, available from the Rubber Division Library of the University of Akron.

12 Hussain, S. M., and Farrell, W. R., "Roll Winding—Causes, Effects and Cures of Loose Cores in Newsprint Rolls," *The Journal of the Technical Association of the Pulp and Paper Industry,* Vol. 60, No. 5, May 1977, pp. 112–114.

13 Brown, C. B., and Goodman, L. E., "Gravitational Stresses in Accreted Bodies," *Proceedings of the Royal Society,* London, Series A, Vol. 276, 1963, pp. 571–576.

14 Timoshenko, S. P., and Goodier, J. N., *Theory of Elasticity,* 3rd ed., McGraw-Hill, New York, 1970.

15 Gradshteyn, I. S., and Ryzhik, I. M., *Table of Integrals, Series and Products,* 4th ed., Academic Press, New York, 1965.

16 Abramowitz, M., and Stegun, I. A., ed., *Handbook of Mathematical Functions,* National Bureau of Standards, Applied Mathematics Series, Vol. 55, May 1968.

E. Reissner

Department of Applied Mechanics
and Engineering Sciences,
University of California, San Diego,
La Jolla, Calif. 92093
Fellow ASME

On Torsion and Transverse Flexure of Orthotropic Elastic Plates[1]

The equations of transverse bending of shear-deformable plates are used for the derivation of a system of one-dimensional equations for beams with unsymmetrical cross section, with account for warping stiffness, in addition to bending, shearing, and twisting stiffness. Significant results of the analysis include the observation that the rate of change of differential bending moment is given by the difference between torque contribution due to plate twisting moments and torque contribution due to plate shear stress resultants; a formula for shear center location which generalizes a result by Griffith and Taylor so as to account for transverse shear deformability and end-section warping restraint; a second-order compatibility equation for the differential bending moment; a contracted boundary condition of support for unsymmetrical cross-section beam theory in place of an explicit consideration of the warping deformation boundary layer; and construction of a problem where the effect of the conditions of support of the beam is such as to give noncoincident shear center and twist center locations.

Introduction

The following considerations are intended as a contribution to the understanding of the foundations of the theory of beams of unsymmetrical cross section and of the extent to which the concept of warping stiffness, in addition to the concepts of bending, shearing, and twisting stiffness, is an essential element of such an understanding. The starting point of our considerations are the basic equations of the linear theory of thin elastic plates including the effect of transverse shear deformation. The principal result of the analysis is the establishment of a system of one-dimensional beam equations with contents going in a simple manner beyond the contents of "elementary" beam theory, with this extension of elementary theory being both necessary and sufficient for the solution of the problem of torsion and flexure of narrow cross-section beams with a widthwise axis of symmetry.

Plate Equations

We take the differential equations of linear plate theory, including the effect of transverse shear deformation, in the form of three equilibrium equations,

$$\begin{aligned} M_{xx,x} + M_{xy,y} - Q_x + m_x &= 0, \\ M_{xy,x} + M_{yy,y} - Q_y + m_y &= 0, \\ Q_{x,x} + Q_{y,y} + q_z &= 0, \end{aligned} \tag{1}$$

and of five constitutive equations,

$$\begin{aligned} M_{xx} = D_x\, \phi_{x,x}, \quad M_{xy} &= D_y(\phi_{x,y} + \phi_{y,x}), \\ Q_x = C(\phi_x + w_{,x}), \quad \phi_{y,y} &= \phi_y + w_{,y} = 0. \end{aligned} \tag{2}$$

The foregoing constitutive equations incorporate a limiting-type orthotropy assumption, which causes M_{yy} and Q_y to be reactive quantities, with this making a relatively simple transition from plate theory to beam theory, in a physically reasonable way, possible.

We assume that the plate is rectangular, with spanwise edges $y = y_1, y = y_2$, and widthwise edges $x = 0$ and $x = L$, and we assume that the spanwise edges are free of load. The boundary conditions for these edges are then,

$$y = y_1, y_2; \quad M_{yx} = M_{yy} = Q_y = 0. \tag{3}$$

The formulation of conditions for the edges $x = 0, L$ will follow the step from two-dimensional plate theory to one-dimensional beam theory.

Derivation of Beam Equations

The constitutive rigidity conditions in (2) imply the representations

$$\phi_y = \theta(x), \quad w = v(x) - y\theta(x), \tag{4a}$$

with θ and v being twisting and bending deflections.

---

[1] A report on work supported by the Office of Naval Research.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, November, 1979; final revision, April, 1980. Paper No. 80-WA/APM-16.

Considering the form of the expression for Q_x in (2) it is consistent with (4a) to stipulate the supplementary representation

$$\phi_x = \phi(x) + y\psi(x), \tag{4b}$$

with ϕ and ψ being bending and warping rotational displacements, respectively.

The equilibrium equations (1) and (3) imply the four integrated relations

$$\begin{aligned} (\int M_{xx}dy)' - \int Q_x dy + \int m_x dy = 0, \\ (\int M_{xy}dy)' - \int Q_y dy + \int m_y dy = 0, \\ (\int Q_x dy)' + \int q_z dy = 0, \\ (\int yQ_x dy)' - \int Q_y dy + \int yq_z dy = 0. \end{aligned} \tag{5}$$

In these, primes indicate differentiation with respect to x, and the integrals extend from y_1 to y_2.

The form of equations (5) suggests introduction of the definitions

$$\begin{aligned} \int M_{xx}dy = M, \quad \int M_{xy}dy = T_M, \\ \int Q_x dy = Q, \quad \int yQ_x dy = -T_Q, \end{aligned} \tag{6}$$

with corresponding definitions of load terms m, q, t_M, and t_Q, and the elimination of $\int Q_y dy$ in (5), so as to have the three beam equilibrium equations

$$\begin{aligned} M' - Q + m = 0, \quad Q' + q = 0, \\ T_M' + T_Q' + t_M + t_Q = 0. \end{aligned} \tag{7}$$

Equations (7) are associated with constitutive equations which follow from (2), (4) and (6) in the form

$$\begin{aligned} M = D_0\phi' + D_1\psi', \qquad Q = C_0(v' + \phi) + C_1(\psi - \theta'), \\ T_M = D_T(\psi + \theta'), \quad -T_Q = C_1(v' + \phi) + C_2(\psi - \theta'), \end{aligned} \tag{8}$$

where

$$D_m = \int y^m D_x dy, \quad D_T = \int D_y dy, \quad C_m = \int y^m C\, dy. \tag{9}$$

Equations (8) involve the four displacement variables ϕ, ψ, θ, and v and, evidently, the three equilibrium equations (7) are too few in number for their determination.

The way out of this difficulty is suggested by the appearance of the system (8) which asks for a companion to the expression for M, in the same way that the expression for $-T_Q$ is a companion to the expression for Q. Accordingly, we introduce an additional beam stress measure R, defined by the relation

$$R = \int yM_{xx}dy. \tag{6'}$$

We note the equivalence of R, which we shall call differential bending moment, to what, in a more general context, has been designated as bimoment by Vlasov [5].

With R as in (6′) we now have the additional constitutive relation

$$R = D_1\phi' + D_2\psi', \tag{8'}$$

and it remains to obtain a fourth equilibrium equation involving R, so as to have as many equilibrium equations as displacement variables. We obtain this additional equation from the first relation in (1), with (6′), (6), and (3), in the form

$$R' - T_M + T_Q + r = 0. \tag{7'}$$

Equations (7) and (7′), in conjunction with (8) and (8′) are now in fact equivalent to a system of four differential equations for the four displacement variables ϕ, ψ, θ, and v.

Beam Equations Without Transverse Shear Deformation

The assumption of absent transverse shear deformation consists in stipulating the limiting relations

$$C_m = \infty, \tag{10}$$

in conjunction with the constraint conditions

$$v' + \phi = 0, \quad \theta' - \psi = 0, \tag{11}$$

which result in the one-dimensional stress measures Q and T_Q now being reactive quantities.

The remaining constitutive relations are then

$$M = D_0\phi' + D_1\psi', \quad T_M = 2D_T\psi, \quad R = D_1\phi' + D_2\psi', \tag{12}$$

with the four equilibrium equations (7) and (7′) now serving to determine the four dependent variables ϕ, ψ, Q, and T_Q and with v and θ then following from (11).

Sequential Determination of Statical and Geometrical Quantities

Given that the four equilibrium equations (7) contain statical quantities only (for the case that the load intensity quantities m, q, t, and r are prescribed rather than displacement-dependent) it will often be useful to use constitutive equations giving displacements in terms of statical quantities, rather than statical quantities in terms of displacements, as in (8). The inverted equations (8) may be written in the form

$$\begin{aligned} \phi' = \frac{M}{D_M} - \frac{R}{D_{MR}}, \quad \psi' = \frac{R}{D_R} - \frac{M}{D_{MR}}, \quad \theta' + \psi = \frac{T_M}{D_T}, \\ v' + \phi = \frac{Q}{C_Q} + \frac{T_Q}{C_{QT}}, \quad \theta' - \psi = \frac{T_Q}{C_T} + \frac{Q}{C_{QT}}, \end{aligned} \tag{13}$$ [2]

where

$$\begin{aligned} \left(\frac{1}{D_M}, \frac{1}{D_{MR}}, \frac{1}{D_R}\right) = \frac{(D_2, D_1, D_0)}{D_2D_0 - D_1{}^2}, \\ \left(\frac{1}{C_Q}, \frac{1}{C_{QT}}, \frac{1}{C_T}\right) = \frac{(C_2, C_1, C_0)}{C_2C_0 - C_1{}^2}. \end{aligned} \tag{14}$$

Equations (13) contain four relations for the determination of the four quantities ϕ', ψ, θ', and v', in terms of the five quantities M, R, T_M, Q, and T_Q. Since there are only four equilibrium equations for these five statical quantities it is necessary to derive a fifth equation, which will have to be of the nature of a compatibility equation. Such a compatibility equation follows from a consideration of the three relations in (13) which involve the warping displacement ψ, in the form

$$\frac{R}{D_R} - \frac{M}{D_{MR}} = \frac{1}{2}\left(\frac{T_M}{D_T} - \frac{T_Q}{C_T} - \frac{Q}{C_{QT}}\right). \tag{15}$$

For the case that the constitutive coefficients are independent of x equation (15) may be transformed usefully, with the help of (7), so as to read

$$\begin{aligned} \frac{R}{D_R} - \left(\frac{1}{D_T} + \frac{1}{C_t}\right)\frac{R''}{4} = \frac{M}{D_{MR}} + \frac{1}{C_{QT}}\frac{q}{2} \\ - \left(\frac{1}{D_T} - \frac{1}{C_T}\right)\frac{t_M + t_Q}{4} + \left(\frac{1}{D_T} + \frac{1}{C_T}\right)\frac{r'}{4}. \end{aligned} \tag{16}$$

The form of this relation makes evident the possibility of localized portions of R as a quantitative measure of the contents of Saint-Venant's principle.

Cantilever Torsion and Flexure

The classical problem of torsion and flexure of end-loaded cantilever beams of span L is given upon assuming that the load terms m, q, t_M, t_Q, and r are absent, and upon stipulating as boundary conditions the loading conditions

[2] We may note the resemblance of equations (13) and (7) to the contents of equations (12), (5), and (3) for a related problem on the subject of box beams [4].

$$x = 0; \quad Q = Q_0, \quad T_M + T_Q = T_0, \quad M = R = 0, \tag{17}$$

and the support conditions

$$x = L; \quad \upsilon = \theta = \psi = \phi = 0. \tag{18}$$

The equilibrium differential equations (7) in conjunction with the first three boundary conditions in (17) now give as expressions for forces and moments

$$Q = Q_0, \quad M = Q_0 x, \quad 2T_M = T_0 + R', \quad 2T_Q = T_0 - R', \tag{19}$$

and the constitutive equations (13) may be written in the form

$$\phi' = \frac{Q_0 x}{D_M} - \frac{R}{D_{MR}}, \quad \upsilon' = -\phi + \frac{Q_0}{C_Q} + \frac{T_0 - R'}{2C_{QT}}, \tag{20}$$

$$\psi = \frac{T_0 + R'}{4D_T} - \frac{T_0 - R'}{4C_T} - \frac{Q_0}{2C_{QT}},$$
$$\theta' = \frac{T_0 + R'}{4D_T} + \frac{T_0 - R'}{4C_T} + \frac{Q_0}{2C_{QT}}. \tag{21}$$

In this R is determined by the differential equation

$$\frac{R}{D_R} - \left(\frac{1}{D_T} + \frac{1}{C_T}\right)\frac{R''}{4} = \frac{Q_0 x}{D_{MR}}, \tag{22}$$

in conjunction with the boundary conditions

$$R(0) = 0, \quad \psi(L) = 0. \tag{23}$$

Equations (22) and (23) give as expression for R,

$$\frac{R}{L} = Q_0\left[\frac{D_R}{D_{RM}}\left(\frac{x}{L} - \frac{\sinh \lambda x}{\lambda L \cosh \lambda L}\right) + \frac{2D_T C_T}{C_{QT}(D_T + C_T)}\frac{\sinh \lambda x}{\lambda L \cosh \lambda L}\right]$$
$$+ T_0 \frac{D_T - C_T}{D_T + C_T}\frac{\sinh \lambda x}{\lambda L \cosh \lambda L}, \quad \lambda^2 = \frac{4}{D_R}\frac{D_T C_T}{D_T + C_T}. \tag{24}$$

Having R, as in (24), we obtain from the second relation in (21), in conjunction with the condition $\theta(L) = 0$.

$$\frac{\theta}{L} = \left[\frac{T_0}{4}\frac{C_T + D_T}{D_T C_T} + \frac{Q_0}{2C_{QT}}\right]\left(\frac{x}{L} - 1\right) + \frac{C_T - D_T}{D_T C_T}\frac{R(x) - R(L)}{4L}. \tag{25}$$

Expressions for ϕ and υ may be determined in an analogous way (where it is convenient to use (22) in order to write R in the first relation in (20) in terms of Q_0 and R'').

With equation (25) we may further write

$$\theta_0 = C_{\theta Q} Q_0 + C_{\theta T} T_0, \tag{26}$$

where θ_0 is the value of θ for $x = 0$, and where the flexibility coefficients $C_{\theta Q}$ and $C_{\theta T}$ come out to be

$$-\frac{C_{\theta Q}}{L} = \frac{D_R}{D_{RM}}\frac{C_T - D_T}{4C_T D_T}\left(1 - \frac{\tanh \lambda L}{\lambda L}\right)$$
$$+ \frac{1}{2C_{QT}}\left(1 + \frac{C_T - D_T}{C_T + D_T}\frac{\tanh \lambda L}{\lambda L}\right), \tag{27}$$

$$-\frac{C_{\theta T}}{L} = \frac{C_T + D_T}{4C_T D_T}\left[1 - \left(\frac{C_T - D_T}{C_T + D_T}\right)^2 \frac{\tanh \lambda L}{\lambda L}\right]. \tag{28}$$

Equations (26)–(28) imply as expression for the coordinate y_S of the center of shear, defined by the stipulation that $T_0 = -Q_0 y_S$ for $\theta_0 = 0$,

$$y_S = \frac{C_{\theta Q}}{C_{\theta T}} = \frac{\dfrac{D_R}{D_{RM}}\dfrac{C_T - D_T}{C_T + D_T}\left(1 - \dfrac{\tanh \lambda L}{\lambda L}\right) + \dfrac{C_T}{C_{TQ}}\dfrac{2D_T}{C_T + D_T}\left(1 + \dfrac{C_T - D_T}{C_T + D_T}\dfrac{\tanh \lambda L}{\lambda L}\right)}{1 - \left(\dfrac{C_T - D_T}{C_T + D_T}\right)^2 \dfrac{\tanh \lambda L}{\lambda L}} \tag{29}$$

In this result the values of λL determine the effect of warping restraint at the fixed end of the beam, and the values of D_T/C_T and D_T/C_{TQ} determine the effect of transverse shear deformation. We note that when $D_T/C_T = D_T/C_{TQ} = 0$ the effects of end section warping restraint cancel out and equation (29) reduces to the simple form

$$y_S = \frac{D_R}{D_{RM}} = \frac{D_1}{D_0}. \tag{29a}$$

It does not seem well known that this result has first been given by Griffith and Taylor [1]. It has been rederived in [2], as a special case of a general shear center formula for open and closed cross-section (simply connected) cylindrical shell beams.

We also note that when λL is large enough for end section warping restraint to be negligible, independent of the values of D_T/C_T and D_T/C_{TQ}, then equation (29) reduces to the form

$$y_S = \frac{D_R}{D_{RM}}\frac{C_T - D_T}{C_T + D_T} + \frac{C_T}{C_{TQ}}\frac{2D_T}{C_T + D_T}, \tag{29b}$$

with (29b) containing (29a) as a special case.

In the range of significant transverse shear deformation effects (29b) implies the simple particular results

$$\frac{D_T}{C_T} = 1; \quad y_S = \frac{C_T}{C_{TQ}} = \frac{C_1}{C_0}, \tag{29c}$$

and

$$1 \ll \frac{D_T}{C_T}; \quad y_S = 2\frac{C_T}{C_{TQ}} - \frac{D_R}{D_{RM}} = 2\frac{C_1}{C_0} - \frac{D_1}{D_0}. \tag{29d}$$

References

1 Griffith, A. A., and Taylor, G. I., "The Problem of Flexure and Its Solution by the Soap-Film Method," Reports and Memoranda of the Advisory Committee for Aeronautics, No. 399, 1917.

2 Reissner, E., and Tsai, W. T., "On the Determination of the Centers of Twist and of Shear for Cylindrical Shell Beams," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, 1972, pp. 1098–1102.

3 Reissner, E., "Some Considerations on the Problem of Torsion and Flexure of Primatical Beams," *International Journal of Solids and Structures*, Vol. 15, 1979, pp. 41–53.

4 Reissner, E., "Note on a Nontrivial Example of Higher-Order One-Dimensional Beam Theory, ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 337–340.

5 Vlasov, V. Z., "Thin-Walled Elastic Beams," Office of Technical Services, U.S. Department of Commerce, Washington, D.C., 1961.

APPENDIX *A*

Asymptotic Solution for Cantilever Torsion and Flexure

We are led to an asymptotic solution of the problem, for $1 \ll \lambda L$, upon writing the solution of the differential bending moment equation (22) in the form

$$R = R_i + R_e, \tag{30}$$

where

$$R_i = \frac{D_R}{D_{RM}} Q_0 x, \tag{31}$$

and

$$R_e'' - \lambda^2 R_e = 0, \tag{32}$$

with λ as in (24).

Having (30)–(32), we rewrite equations (20) and (21) in the form

$$\phi' = \phi_i' - \frac{R_e''}{D_{RM}\lambda^2}, \quad \upsilon' = \upsilon_i' + \frac{R_e'}{D_{RM}\lambda^2} - \frac{R_e'}{2C_{QT}} \tag{33}$$

$$\psi = \psi_i + \left(\frac{1}{D_T} + \frac{1}{C_T}\right)\frac{R_e'}{4}, \quad \theta' = \theta_i' + \left(\frac{1}{D_T} - \frac{1}{C_T}\right)\frac{R_e'}{4}, \tag{34}$$

where the solution of (32)–(34) is to be determined in such a way that the five boundary conditions in (18) and (23) are satisfied. As they must, equations of (32)–(34) imply just the right number of constants of integration to make possible the satisfaction of these five conditions.

In order to see the possibility of an asymptotic solution of the problem, with separate determination of interior and edge-zone solution contributions, for the case of sufficiently large values of λL, we reintroduce in (33) the defining relation for λ^2 from equation (24) and then integrate in (33) and (34) with respect to x, so as to obtain in addition to the expression for ψ in (34)

$$\phi = \phi_i - \frac{D_R}{D_{RM}}\left(\frac{1}{D_T} + \frac{1}{C_T}\right)\frac{R_e'}{4},$$
$$\upsilon = \upsilon_i + \left[\frac{D_R}{D_{RM}}\left(\frac{1}{D_T} + \frac{1}{C_T}\right) - \frac{2}{C_{QT}}\right]\frac{R_e}{4} \tag{35}$$

$$\theta = \theta_i + \left(\frac{1}{D_T} - \frac{1}{C_T}\right)\frac{R_e}{4}. \tag{36}$$

Having (34)–(36) we now consider the boundary conditions (18), in conjunction with the boundary condition $R_e(0) = 0$ which is implied by equation (30), (31), and (23). We will be using these conditions in an asymptotic sense, for large λL, if we take account of the fact that for large λL the terms with R_e in (34)–(36) will be small of relative order $1/\lambda L$ compared to the terms with R_e', in such a way that the two boundary conditions in (18) which involve R_e itself rather than R_e' may be replaced by the simpler conditions

$$\upsilon_i(L) = 0, \quad \theta_i(L) = 0. \tag{37}$$

There now remain the two conditions for ψ and ϕ, for the determination of ϕ_i and R_e in terms of the known expression for ψ_i. The possibility of a determination of ϕ_i without reference to the determination of R_e is given upon observing the existence of a *contracted* condition involving ϕ_i and ψ_i in the form

$$D_{RM}\phi_i(L) + D_R\psi_i(L) = 0. \tag{38}$$

At the same time we retain as boundary condition for the separate determination of the edge zone contribution R_e the one condition

$$R_e'(L) = -\frac{4D_TC_T}{D_T + C_T}\psi_i(L). \tag{39}$$

Having equation (38), in conjunction with (37), for the determination of the interior state, we now find from (33) and (34), in conjunction with (20) and (21), as expressions for θ_i, ϕ_i, and υ_i,

$$\theta_i = \left\{\left(\frac{1}{D_T} + \frac{1}{C_T}\right)T_0 + \left[\frac{D_R}{D_{RM}}\left(\frac{1}{D_T} - \frac{1}{C_T}\right) + \frac{2}{C_{QT}}\right]Q_0\right\}\frac{x - L}{4}, \tag{40}$$

$$\phi_i = \left(1 - \frac{D_MD_R}{D_{MR}^2}\right)\frac{Q_0}{D_M}\frac{x^2 - L^2}{2} - \frac{D_R}{D_{RM}}\left\{\left(\frac{1}{D_T} - \frac{1}{C_T}\right)\frac{T_0}{4} + \left[\frac{D_R}{D_{RM}}\left(\frac{1}{D_T} + \frac{1}{C_T}\right) - \frac{2}{C_{QT}}\right]\frac{Q_0}{4}\right\}, \tag{41}$$

$$\upsilon_i = -\left(1 - \frac{D_MD_R}{D_{MR}^2}\right)\frac{Q_0}{D_M}\frac{x^3 - 3L^2x + 2L^3}{6} + \left\{\frac{4}{C_Q} + \frac{D_R}{D_{RM}}\left[\frac{D_R}{D_{RM}}\left(\frac{1}{D_T} + \frac{1}{C_T}\right) - \frac{4}{C_{QT}}\right]\right\}Q_0\frac{x - L}{4} + \left[\frac{D_R}{D_{RM}}\left(\frac{1}{D_T} - \frac{1}{C_T}\right) + \frac{2}{C_{QT}}\right]T_0\frac{x - L}{4}. \tag{42}$$

We note that equations (40)–(42) assume a somewhat simpler form with $(1 - D_MD_R/D_{MR}^2)/D_M = 1/D_0$ and $D_R/D_{RM} = D_1/D_0$, which relations are implied by equations (14). A comparison of (40) with (25) shows agreement insofar as the expression for θ_i is concerned. As ϕ and υ had not earlier been determined we will not here carry out a complementary verification of the expressions for ϕ_i and υ_i.

It remains to determine R_e and the edge zone displacement contributions θ_e and υ_e. We find from (32), (30), (31), (23), and (39) as expression for R_e

$$R_e = -\left[\frac{C_T - D_T}{C_T + D_T}T_0 + \left(\frac{D_R}{D_{RM}} - \frac{2}{C_{QT}}\frac{D_TC_T}{C_T + D_T}\right)Q_0\right]\frac{\sinh \lambda x}{\lambda \cosh \lambda L} \tag{43}$$

while, in accordance with (36) and (37),

$$\theta_e = \frac{C_T - D_T}{4C_TD_T}R_e, \quad \upsilon_e = \left(\frac{D_R}{D_{RM}}\frac{C_T + D_T}{4C_TD_T} - \frac{1}{2C_{QT}}\right)R_e, \tag{44}$$

with (43) and (44) making it apparent that $\theta \approx \theta_i$ and $\upsilon \approx \upsilon_i$ except for terms of relative order $1/\lambda L$.

Having now an expression for $\theta = \theta_i + \theta_e$, we may once again determine the values of flexibility coefficients $C_{\theta T}$ and $C_{\theta Q}$ including the effect of warping restraint. The results obtained in this way are in complete agreement with the results in (27) and (28), which means that the simple asymptotic analysis in this section leads not just to an asymptotic solution but to the actual exact solution, by the simple device of retaining the factor $\tanh \lambda L$ in the analysis, instead of replacing it by a factor "1," as would be asymptotically consistent.

We conclude this section by a consideration of those aspects of the force and moment distribution which are of a nonelementary nature. These are the differential bending moment distribution $R = R_i + R_e$ with R_i and R_e as in (31) and (43), and expressions for the torque components T_M and T_Q which follow from (19) in conjunction with (31) and (43) in the form

$$\begin{Bmatrix} T_M \\ T_Q \end{Bmatrix} = \left\{\frac{T_0}{2} \pm \frac{Q_0}{2}\frac{D_R}{D_{RM}} \pm \frac{\cosh \lambda x}{\cosh \lambda L}\left[\frac{T_0}{2}\frac{C_T - D_T}{C_T + D_T} + \frac{Q_0}{2}\left(\frac{D_R}{D_{RM}} - \frac{C_T}{C_{TQ}}\frac{2D_T}{C_T + D_T}\right)\right]\right\}. \tag{45}$$

APPENDIX *B*

Approximate Determination of Flexibility Coefficients by a Variational Procedure

It follows from general principles that an appropriate statement of the theorem of minimum complementary energy for the cantilever torsion-bending problem, as considered in the foregoing, consists in the variational equation

$$\delta\left\{T_0\theta_0 + Q_0\upsilon_0 + \frac{1}{2}\int_0^L\left[\frac{M^2}{D_M} - 2\frac{MR}{D_{MR}} + \frac{R^2}{D_R} + \frac{T_M^2}{D_T} + \frac{T_Q^2}{C_T} + 2\frac{QT_Q}{C_{TQ}} + \frac{Q^2}{C_Q}\right]dx\right\} = 0. \tag{46}$$

In this M, R, T_M, T_Q, and Q must satisfy the four homogeneous equilibrium equations (7) and in addition the stress boundary conditions $M(0) = R(0) = 0$, with the other two stress boundary conditions in (17) replaced by displacement boundary conditions $\theta(0) = \theta_0$ and $\upsilon(0) = \upsilon_0$.

While earlier considerations [3] had indicated that for the purpose of obtaining bound-relations, equation (46) should be used for the determination of stiffness coefficient approximations we will here once again, as in [2], consider the use of (48) for the determination of flexibility coefficient approximations.

The first two equilibrium relations in (7), in conjunction with the stipulation $M(0) = 0$, make it apparent that we have no choice other than to use $Q = Q_0$ and $M = Q_0x$ as expressions for Q and M. As regards R, T_M, and T_Q we are free to choose approximative expressions restricted in no other way than by the relations

$$T_M = \frac{1}{2}(T_0 + R') \quad \text{and} \quad T_Q = \frac{1}{2}(T_0 - R'),$$

in conjunction with the boundary condition $R(0) = 0$.

It remains then to make a suitable choice of a function which ap-

proximates the function R. We will consider two specific possibilities, both of them leading to results of interest.

As a first possibility we chose R in such a way that the state of stress is in accordance with Saint-Venant's theory of torsion and flexure, in which there is no reference to the condition of warping restraint as a part of the conditions of support. We obtain a Saint-Venant stress distribution by assuming that the warping displacement measure ψ does not vary along the span, that is, by setting $\psi' = 0$ in the constitutive equations (13). Setting $\psi' = 0$ we find immediately that now $R = D_R Q_0 x/D_{MR}$ and therewith $2T_M = T_0 + D_R Q_0/D_{MR}$ and $2T_Q = T_0 - D_R Q_0/D_{MR}$, with equation (46) now reducing to the form

$$\delta\{T_0\theta_0 + Q_0 v_0 + f(T_0, Q_0)\} = 0, \tag{47}$$

where f is a simple quadratic expression.

The evaluation of (47) gives

$$-\frac{\theta_0}{L} = \left(\frac{1}{D_T} + \frac{1}{C_T}\right)\frac{T_0}{4} + \left[\frac{D_R}{D_{RM}}\left(\frac{1}{D_T} - \frac{1}{C_T}\right) + \frac{2}{C_{QT}}\right]\frac{Q_0}{4}, \tag{48}$$

$$-\frac{v_0}{L} = \left\{\frac{4}{3}\frac{L^2}{D_M}\left(1 - \frac{D_M D_R}{D_{MR}{}^2}\right) + \frac{4}{C_Q} + \frac{D_R}{D_{RM}}\left[\frac{D_R}{D_{RM}}\left(\frac{1}{D_T} + \frac{1}{C_T}\right) - \frac{4}{C_{QT}}\right]\right\}\frac{Q_0}{4} + \left[\frac{D_R}{D_{RM}}\left(\frac{1}{D_T} - \frac{1}{C_T}\right) + \frac{2}{C_{QT}}\right]\frac{T_0}{4}, \tag{49}$$

with these formulas being in agreement with the results of the determination of θ_0 and v_0 through solution of the boundary value problem, for the limiting case $1/\lambda L = 0$.

A second possibility to choose R, of comparable simplicity, is to assume that $R = R_0 x$, with R_0 being determined from a variational equation

$$\delta\{T_0\theta_0 + Q_0 v_0 + f(T_0, Q_0, R_0)\} = 0, \tag{50}$$

with this procedure being equivalent to what has previously been proposed in a more general context in [3]. Remarkably, equation (50) gives an expression for R_0 of the form $R_0 = D_R Q_0/D_{MR} + O(1/\lambda^2 L^2)$, and the substitution of this in the associated equations for θ_0 and v_0 gives expressions for θ_0 and v_0 which differ from those in (48) and (49) by no more than terms of relative order of magnitude $1/\lambda^2 L^2$, consistent with a more general statement in [3]. While this result and the statement in [3] are meaningful insofar as upper and lower bound calculations are concerned, the differential equation solution described earlier indicates a warping effect on flexibility coefficients of relative order $1/\lambda L$, rather than of order $1/\lambda^2 L^2$, and, evidently, to obtain the correct order of magnitude by the variational procedure, an approximation for R must account for the boundary layer aspects of this function in the parameter value range $1 \ll \lambda L$.

APPENDIX *C*

On Saint-Venant Torsion and Flexure

The results for Saint-Venant *torsion* are obtained by setting, in a *semi-inverse sense*, $M = R = Q = 0$ and therewith $\theta' = \psi' = 0$. Equations (13) then give as expressions for θ' and ψ

$$2\theta' = \frac{T_M}{D_T} + \frac{T_Q}{C_T}, \quad 2\psi = \frac{T_M}{D_T} - \frac{T_Q}{C_T}, \tag{51}$$

with $T_M + T_Q = T_0$. In order to express θ' in terms of T_0 it is essential to make use of the fourth equilibrium equation, (7′). This equation reduces here to the form $T_M - T_Q = 0$, and with this we now have that $T_M = T_Q = \frac{1}{2}T_0$ and therewith $4\theta' = T_0(1/D_T + 1/C_T)$.

The results for Saint-Venant *flexure* are obtained by now setting *a priori*, in a semi-inverse sense, $\psi' = 0$. With this we have then $M = D_0\phi'$ and $R = D_1\phi'$ and this in turn implies the crucial relation $R = D_1 M/D_0 = D_1 Q_0 x/D_0$. In order to evaluate the constitutive relation

$$2\theta' = \frac{T_M}{D_T} + \frac{T_Q}{C_T} + \frac{Q}{C_{QT}}, \tag{52}$$

it is again essential to utilize the equilibrium equation $R' = T_M - T_Q$. Therewith and with $T_M + T_Q = T_0$, it is then possible to express θ' in terms of T_0 and Q_0, in such a way that the result coincides with the expression for θ' in equation (34). It follows from this and from the boundary condition $\theta(L) = 0$ that the Saint-Venant rotation θ is the same as the θ_i in equation (40).

It remains to determine ϕ and v with the help of the relations $\phi' = Q_0 x/D_0$ and $v' = -\phi + Q_0/C_Q + \frac{1}{2}(T_0 - D_1 Q_0/D_0)/C_{QT}$, in conjunction with suitable boundary conditions. The important point here is that Saint-Venant's theory makes no reference to the subject of boundary conditions of support. In order to bridge this gap one may use a variational condition, as in the preceding section, in place of explicitly stated conditions of support. Alternately, our earlier considerations show how the derivation of appropriate boundary conditions, involving the *contraction* of two conditions for θ and ψ into one condition, as in (38), is in fact a problem of asymptotic analysis.

APPENDIX *D*

On the Relation Between Shear Center and Twist Center

The present formulation of the problems of torsion and flexure of a class of narrow-cross-section beams makes possible an explicit demonstration of the significance of earlier formulas [2, 3] for the location y_T of the center of twist and the location y_S of the center of shear in terms of flexibility coefficients. One aspect of these formulas was the observation that symmetry of the flexibility coefficient matrix is a necessary and sufficient condition for the coincidence of the two centers. As far as is known, no cases have yet been considered for which this symmetry does not hold. In what follows we construct an example of such a case by modifying the boundary conditions (17) and (18) for the problem of the end-loaded cantilever in such a way that we replace the two conditions $\theta(L) = \phi(L) = 0$ in (18) by conditions of *flexible* support of the form

$$-\theta(L) = c_{\theta T}T(L) + c_{\theta M}M(L), \quad -\phi(L) = c_{\phi T}T(L) + c_{\phi M}M(L). \tag{53}$$

In order for this modified beam problem to be such as to allow a variational formulation, the coefficients in (53) must satisfy the symmetry condition $c_{\theta M} = c_{\phi T}$. We will show that the same condition is necessary and sufficient to ensure validity of the relation $y_T = y_S$.

In connection with the foregoing problem we assume here, for simplicity's sake, that transverse shear deformation is negligible, by setting $C_n = \infty$, and that the beam cross section is *doubly* symmetric, by setting $D_1 = 0$. Evidently, for the original problem, with the boundary conditions (18), we will now, by symmetry, have that $y_T = y_S = 0$. For the present more general case, equations (19) remain unchanged, equations (20) and (21) assume the form

$$\phi' = \frac{Q_0 x}{D_0}, \quad v' = -\phi, \quad \psi = \frac{T_0 + R'}{4D_T}, \quad \theta' = \frac{T_0 + R'}{4D_T}, \tag{54}$$

and the differential equation (22) becomes

$$D_2 R'' - 4D_T R = 0, \tag{55}$$

with the remaining boundary conditions being equations (53), together with the conditions $R(0) = 0$ and $v(L) = \psi(L) = 0$.

Equation (55), together with the boundary conditions for R and ψ now gives as expression for R,

$$R = -\frac{T_0}{\lambda}\frac{\sinh \lambda x}{\cosh \lambda L}, \tag{56}$$

where $\lambda = 2\sqrt{D_T/D_2}$.

In order to obtain the values of θ_0 and v_0 which are needed for the determination of y_T and y_S, we find, from (54),

$$\theta = \frac{T_0 x + R}{4D_T} + c_1, \quad \phi = \frac{Q_0 x^2}{2D_0} + c_2 \tag{57}$$

where, in accordance with (53), c_1 and c_2 must be such that

$$\frac{T_0 L}{4D_T}\left(1 - \frac{\tanh \lambda L}{\lambda L}\right) + c_1 = -c_{\theta T} T_0 - c_{\theta M} Q_0 L, \tag{58}$$

and

$$\frac{Q_0 L^2}{2D_0} + c_2 = -c_{\phi T} T_0 - c_{\phi M} Q_0 L. \tag{59}$$

Having ϕ, the second relation in (54) gives

$$v = -\left(\frac{Q_0 x^3}{6D_0} + c_2 x\right) + c_3, \tag{60}$$

and the condition $v(L) = 0$ gives as expression for c_3

$$c_3 = \frac{Q_0 L^3}{6D_0} + c_2 L. \tag{61}$$

Therewith, and with the abbreviation $C_{\theta T} = L/4D_T$, we have now

$$\theta_0 = c_1 = -\left[c_{\theta T} + C_{\theta T}\left(1 - \frac{\tanh \lambda L}{\lambda L}\right)\right] T_0 - c_{\theta M} L Q_0 \tag{62}$$

and, with $L/3D_0 = C_{\phi M}$,

$$\frac{v_0}{L} = \frac{c_3}{L} = -c_{\phi T} T_0 - [c_{\phi M} + C_{\phi M}] L Q_0. \tag{63}$$

Equations (62) and (63) give, with $T_0 = -y_S Q_0$ for $\theta_0 = 0$, and with $v_0 = \theta_0 y_T$ for $Q_0 = 0$,

$$\frac{(y_S, y_T)}{L} = \frac{(c_{\theta M}, c_{\phi T})}{c_{\theta T} + C_{\theta T}[1 - (\lambda L)^{-1} \tanh \lambda L]}, \tag{64}$$

so that, in fact, $y_S \neq y_T$ in the event that $c_{\theta M} \neq c_{\phi T}$.

G. V. Ranjan
Director,
Shelltech Associates,
Stanford, Calif. 94304

C. R. Steele
Professor,
Division of Applied Mechanics,
Department of Mechanical Engineering,
Stanford University,
Stanford, Calif. 94305
Mem. ASME

Nonlinear Corrections for Edge Bending of Shells[1]

Asymptotic expansions for self-equilibrating edge loading are derived in terms of exponential functions, from which formulas for the stiffness and flexibility edge influence coefficients are obtained, which include the quadratic nonlinear terms. The flexibility coefficients agree with those previously obtained by Van Dyke for the pressurized spherical shell and provide the generalization to general geometry and loading. In addition, the axial displacement is obtained. The nonlinear terms in the differential equations can be identified as "prestress" and "quadratic rotation." To assess the importance of the latter, the problem of a pressurized spherical cap with roller supported edges is considered. Results show that whether the rotation at the edge is constrained or not, the quadratic rotation terms do not have a large effect on the axial displacement. The effect will be large for problems with small membrane stresses.

Introduction

Reissner [1] formulated the nonlinear equations for the finite axisymmetric deformations of thin elastic shells of revolution, which are acted upon by surface loads as well as edge forces and moments. He showed that the change in the meridional slope during deformation is an important geometric nonlinearity. This nonlinearity alters the moment equilibrium equation and is frequently referred to as the nonlinear prestress effect. The other nonlinearity introduced by moderately large rotation alters the compatibility equation. In the analyses of many problems, it is enough to include the nonlinear prestress effect alone. However in certain problems, particularly stability, the nonlinearity due to moderately large rotation is significant. So, the objective of the present investigation is to develop asymptotic expansion solutions to Reissner's nonlinear equations. In particular, nonlinear corrections to the results of linear theory for edge force, moment, and axial displacement in terms of edge radial displacement and rotation are desired.

Reissner [2] developed an asymptotic expansion for the nonlinear "edge effect" solution of a steep, smooth shell acted upon only by self-equilibrating edge forces and moments. Van Dyke [3] included the membrane prestress term and obtained the correct expansion for the spherical shell. In the present investigation, we extend his approach to the general shell of revolution. The nonlinear expressions for meridional edge moment and radial edge force in terms of edge rotation and radial edge displacement are useful in solving shells with, for example, clamped edge conditions. These are inverted to facilitate analysis of shells with either meridional slope discontinuity or roller-supported or simply supported edges. As illustrations, linear and nonlinear solutions for shells used in pressure gauges and for domes with roller-supported edges are compared to show the effects of the two nonlinearities. A further application of the present results is in the analysis by Ranjan [4], reported in Ranjan and Steele [5], of the large displacement dimpling of a spherical shell.

Nonlinear Equations for Moderate Rotation

The linear part of the field equations formulated by Reissner [1] for shells with prescribed surface and edge loadings are written in matrix form in [4,6]. The dependent variable vector is

$$\mathbf{y} = \begin{bmatrix} M_\varphi/(Etc)_e \\ H\lambda \sin\varphi_e/(Et)_e \\ \chi/\lambda \\ h/r_e \end{bmatrix} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \\ y^{(4)} \end{bmatrix} \tag{1}$$

in which M_φ, H, χ, and h are the physical quantities which can be prescribed on the shell edges: meridional bending moment resultant, radial force resultant, rotation, and radial displacement, respectively. Young's modulus is E and the thickness is t, with

$$c = t/[12\,(1-\nu^2)]^{1/2}$$

in which ν is Poisson's ratio. The radius is r and the angle between axis and normal to the shell midsurface is φ. The subscript e denotes a reference point at which the quantities are evaluated. The large parameter of the problem is

$$\lambda = [r/(c\sin\varphi)]_e^{1/2}$$

All the elements of $\mathbf{y}$ are the same order of magnitude for the linear bending solution. The linear equations [1] are used in [4,6] in the form

[1] This study was supported by a grant from the National Science Foundation to Stanford University.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, May, 1980. Paper No. 80-WA/APM-15.

$$-\frac{d\mathbf{y}}{dx} + \lambda \mathbf{A}\mathbf{y} = \mathbf{a} \tag{2}$$

in which the independent variable is the dimensionless arclength

$$x = (s \sin \varphi_e)/r_e$$

The matrix $\mathbf{A}$ and the vector $\mathbf{a}$ have the forms

$$\mathbf{A} = \mathbf{A}_0 + \frac{1}{\lambda}\mathbf{A}_1 + \frac{1}{\lambda^2}\mathbf{A}_2$$

$$\mathbf{a} = \mathbf{a}_0 + \frac{1}{\lambda}\mathbf{a}_1 + \frac{1}{\lambda^2}\mathbf{a}_2 \tag{3}$$

where $\mathbf{A}_i$ and $\mathbf{a}_i$ are given in [4, 6]. The leading terms are

$$\mathbf{A}_0 = \begin{bmatrix} 0 & (\sin\varphi)^* & 0 & 0 \\ 0 & 0 & 0 & (Et/r^2)^* \\ 1/(Etc^2)^* & 0 & 0 & 0 \\ 0 & 0 & -(\sin\varphi)^* & 0 \end{bmatrix}$$

$$\mathbf{a}_0 = \hat{y}\frac{\cos\phi}{\sin\phi_e}[1, 0, 0, 0]^T$$

in which the star denotes that the quantity is divided by its value at the reference point φ_e, and the load magnitude is assumed to be such that

$$\hat{y} = V\lambda^2 \sin\phi_e/(Et)_e = O(1)$$

The equations for moderate rotation [1] are obtained by adding to the meridional strain ϵ_ϕ the nonlinear term

$$\epsilon_\phi \rightarrow \epsilon_\phi + \chi^2/2,$$

where χ is the rotation. The matrix equation becomes

$$-\frac{d\mathbf{y}}{dx} + \lambda \mathbf{A}\mathbf{y} + \mathbf{N} = \mathbf{a} \tag{4}$$

in which the vector on nonlinear terms is

$$\mathbf{N} = \begin{bmatrix} \dfrac{\lambda^2\cos\varphi}{\sin\varphi_e} y^{(2)}y^{(3)} + \dfrac{V\lambda^3\sin\varphi}{(Et)_e} y^{(3)} \\ 0 \\ 0 \\ -\dfrac{\lambda^2\cos\varphi}{2\sin\varphi_e}[y^{(3)}]^2 \end{bmatrix}$$

In the majority of problems, the axial resultant V is statically determinant. Furthermore, for many problems the membrane approximation $H \simeq V \cot\phi$ is valid, since edge bending effects provide a correction to H of higher order ($O(\lambda^{-1})$). In this situation the first element of $\mathbf{N}$ is a linear term, with a known coefficient multiplying $y^{(3)}$. This is, of course, the well-known "prestress" term which provides classical stability limits and a significant correction to edge bending effects [2,4–6]. In this communication, we demonstrate a formal expansion procedure which is valid when the bending contribution to H is as large or larger than the membrane contribution.

Asymptotic Expansion Solution

A formal solution to the nonlinear matrix formulation (4) can be obtained in the form of the asymptotic expansion

$$\begin{aligned} \mathbf{y} = {} & \frac{1}{\lambda}\boldsymbol{\Psi}_0 + \frac{1}{\lambda^2}\boldsymbol{\Psi}_1 + \frac{1}{\lambda^3}\boldsymbol{\Psi}_2 + \dots \\ & + \mathcal{R}e\left\{\exp(\lambda\xi)\left[\frac{1}{\lambda}\boldsymbol{\alpha}_0 + \frac{1}{\lambda^2}\boldsymbol{\alpha}_1 + \frac{1}{\lambda^3}\boldsymbol{\alpha}_2 + \dots\right]\right\} \\ & + \mathcal{R}e\left\{\exp(2\lambda\xi)\left[\frac{1}{\lambda^2}\boldsymbol{\beta}_0 + \frac{1}{\lambda^3}\boldsymbol{\beta}_1 + \frac{1}{\lambda^4}\boldsymbol{\beta}_2 + \dots\right]\right\} \\ & + \exp(2\,\mathcal{R}e\{\lambda\xi\})\left[\frac{1}{\lambda^2}\boldsymbol{\gamma}_0 + \frac{1}{\lambda^3}\boldsymbol{\gamma}_1 + \frac{1}{\lambda^4}\boldsymbol{\gamma}_2 + \dots\right] + \dots \end{aligned} \tag{5}$$

in which the $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i$, $\boldsymbol{\gamma}_i$, and $\boldsymbol{\psi}_i$ are vector functions and ξ is a scalar function of x, all independent of the large parameter λ. Substitution of (5) into (4) and equating the coefficient of each power of λ to zero in the usual procedure gives a system of equations for the unknown functions. The leading equation for the $\boldsymbol{\psi}_i$ is

$$\mathbf{A}_0\boldsymbol{\Psi}_0 + \begin{bmatrix} \dfrac{\Psi_0^{(3)}}{\sin\varphi_e}[\Psi_0^{(2)}\cos\varphi + \hat{y}\sin\varphi] \\ 0 \\ 0 \\ -\dfrac{\cos\varphi}{2\sin\varphi_e}[\Psi_0^{(3)}]^2 \end{bmatrix} = \boldsymbol{\alpha}_0 \tag{6}$$

Because of the simple structure of $\mathbf{A}_0$, the solution of (6) is just

$$\boldsymbol{\Psi}_0 = \hat{y}\cot\phi[0\ 1\ 0\ 0]^T \tag{7}$$

which gives the membrane result

$$H = V\cot\varphi \tag{8}$$

where V can be determined from axial equilibrium equations when the surface loads are known. Using (7), we find that

$$\frac{1}{\sin\varphi_e}[\Psi_0^{(2)}\cos\varphi + \hat{y}\sin\varphi] = 2\rho\left(\frac{Etc}{r_2}\right)^* \tag{9}$$

where

$$\rho = Vr/(2\,Etc\sin^2\varphi) \tag{10}$$

is a quantity $O(1)$. Thus the nonlinear surface load or prestress effect introduces an additional term

$$A_0^{(1,3)} = 2\rho\,(Etc/r_2)^* \tag{11}$$

into the matrix $\mathbf{A}_0$. Vectors $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$ are determined from the equations

$$\mathbf{A}_0\boldsymbol{\Psi}_1 = \mathbf{a}_1 + \frac{d\boldsymbol{\Psi}_0}{dx} - \mathbf{A}_1\boldsymbol{\Psi}_0 \tag{12}$$

$$\mathbf{A}_0\boldsymbol{\Psi}_2 = \mathbf{a}_2 + \frac{d\boldsymbol{\Psi}_1}{dx} - \mathbf{A}_1\boldsymbol{\Psi}_1 - \mathbf{A}_2\boldsymbol{\Psi}_0 \tag{13}$$

It can be observed that the nonlinear prestress term $A_0^{(1,3)}$ does not affect $\boldsymbol{\Psi}_1$, but modifies $\boldsymbol{\Psi}_2$. Therefore, the solution to (12) is

$$\boldsymbol{\Psi}_1 = \left(\frac{r}{r_e}\right)\frac{1}{Et}\left\{r_2P_z - \frac{V}{\sin\varphi}\left(\frac{r_2}{r_1} + \nu\right)\right\}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{14}$$

in which P_z is the normal component of surface loading.

Now the coefficients of the exponential terms in (5) will be considered. The vector $\boldsymbol{\alpha}_0$ is obtained from the equation

$$\left[-\frac{d\xi}{dx}\mathbf{I} + \mathbf{A}_0\right]\boldsymbol{\alpha}_0 = \mathbf{0} \tag{15}$$

For a nontrivial solution of (15), $d\xi/dx$ must be an eigenvalue and $\boldsymbol{\alpha}_0$ the corresponding eigenvector of $\mathbf{A}_0$. For $|\rho| \leq 1$, an angle of η is introduced by the relation

$$\rho = \cos\eta \tag{16}$$

It is seen that for the solution which decreases when $s < 0$, and thus is significant at the lower edge of a smooth shell, one obtains the eigenvalue

$$\left(\frac{d\xi}{dx}\right)\left[\left(\frac{cr}{\sin\varphi}\right)^*\right]^{1/2} = \exp(i\eta/2) \tag{17}$$

which establishes the relation

$$\lambda\xi = \int_0^s \frac{\exp(i\eta/2)}{(cr_2)^{1/2}}\,ds \tag{18}$$

The corresponding eigenvector is

$$\alpha_0 = d_0(x)\begin{bmatrix} -\left(\frac{d\xi}{dx}\right)^2 (Etc^2/\sin\varphi)^* \\ \left(\frac{d\xi}{dx}\right)^{-1} (Et/r^2)^* \\ -\left(\frac{d\xi}{dx}\right)\Big/(\sin\varphi)^* \\ 1 \end{bmatrix} \tag{19}$$

The function $d_0(x)$ in (19) is assumed to be a smooth function. The proper eigenvalue and eigenvector should be considered for the solution that decreases exponentially for $s > 0$ and hence is significant at the upper edge of a smooth shell. For simplicity, only the lower edge of the shell will be considered. On the basis of known linear results, it is assumed that significant changes of stress and deformation due to the self-equilibrating edge forces and moments are limited to a relatively narrow region adjacent to the edge $s = 0$ of the shell. In this region the variation of geometric and material properties may be neglected. Hence, throughout this narrow edge zone, the function $d_0(x)$ in (19) would assume the value $d_0(0) = D$, a complex constant.

Equations to determine the vectors β_0 and γ_0 are now provided.

$$\left[-2\frac{d\xi}{dx}\mathbf{I} + \mathbf{A}_0\right]\beta_0 = \begin{bmatrix} (\cot\varphi_e D^2)/2 \\ 0 \\ 0 \\ [\cot\varphi_e D^2 \exp(i\eta)]/4 \end{bmatrix} \tag{20}$$

$$\left[-2\cos\frac{\eta}{2}\mathbf{I} + \mathbf{A}_0\right]\gamma_0 = \begin{bmatrix} (\cot\varphi_e |D|^2 \cos\eta)/2 \\ 0 \\ 0 \\ (\cot\varphi_e |D|^2)/4 \end{bmatrix} \tag{21}$$

The solutions of (20) and (21) are

$$\beta_0 = \frac{\cot\varphi_e D^2}{4(5+4\rho)(5-4\rho)} \times \begin{bmatrix} 6[\exp(i3\eta/2) - 4\exp(-i\eta/2)] \\ [\exp(i2\eta) - 5 + 4\exp(-i2\eta)] \\ 3[\exp(i\eta) - 4\exp(-i\eta)] \\ 2[\exp(i5\eta/2) - 5\exp(i\eta/2) + 4\exp(-i3\eta/2)] \end{bmatrix} \tag{22}$$

$$\gamma_0 = -\frac{\cot\varphi_e |D|^2}{4(5+4\rho)} \begin{bmatrix} (4\rho^2+4\rho+1)[2(1+\rho)]^{1/2} \\ 2(1-\rho) \\ (4\rho^2+4\rho+1) \\ 2(1-\rho)[2(1+\rho)]^{1/2} \end{bmatrix} \tag{23}$$

Higher-order terms need to be evaluated only when the edge of the shell is not too steep. It can be observed that the vectors Ψ_i, given by (7), (13), and (14), provide the solution due to surface loads while the vectors α_i, β_i, and γ_i in (5) determine the solution due to edge forces and moments.

Edge Influence Coefficients

From the results (19), (22), (23) the first nonlinear corrections to the relations for edge forces $y^{(1)}$ and $y^{(2)}$ in terms of edge displacements $y^{(3)}$ and $y^{(4)}$ can be obtained

$$y^{(1)} = (2+2\rho)^{1/2} y^{(3)} + y^{(4)} + \frac{\lambda\cot\varphi_e}{2(5+4\rho)} \times \{3(2+2\rho)^{1/2}[y^{(3)}]^2 + 2(1+2\rho)y^{(3)}y^{(4)}\} \tag{24}$$

$$y^{(2)} = y^{(3)} + (2+2\rho)^{1/2} y^{(4)} + \frac{\lambda\cot\varphi_e}{2(5+4\rho)} \times \{(1+2\rho)[y^{(3)}]^2 + 2(1+2\rho)[y^{(4)}]^2\} \tag{25}$$

where ρ, defined by (10), introduces the prestress or surface load effect. The terms multiplied by $\lambda\cot\varphi_e$ in (24) and (25) provide the nonlinear correction to the known linear result.

An increment in the nondimensional work performed by the edge loads is given by

$$dU = y^{(1)}dy^{(3)} + y^{(2)}dy^{(4)} \tag{26}$$

To be independent of load path, this must be an exact derivative, so that

$$y^{(1)} = \frac{\partial U}{\partial y^{(3)}}; \quad y^{(2)} = \frac{\partial U}{\partial y^{(4)}} \tag{27}$$

from which the following condition is obtained:

$$\frac{\partial y^{(1)}}{\partial y^{(4)}} = \frac{\partial y^{(2)}}{\partial y^{(3)}} \tag{28}$$

The expressions (24) and (25) do satisfy this criterion.

Equations (24) and (25) are useful in analyzing shells with edges where the displacements are prescribed, as in the case of clamped edges. To facilitate the analyses of shells with either roller-supported edges where $y^{(1)}$ and $y^{(2)}$ are prescribed or simply supported edges where $y^{(1)}$ and $y^{(4)}$ are specified or edges at slope discontinuity where $y^{(2)}$ and $y^{(3)}$ are prescribed, (24) and (25) are inverted to obtain the following relations. The expressions for edge displacements $y^{(3)}$ and $y^{(4)}$ in terms of edge forces $y^{(1)}$ and $y^{(2)}$ are

$$y^{(3)} = \frac{1}{(1+2\rho)}\Bigg[(2+2\rho)^{1/2}y^{(1)} - y^{(2)} - \frac{\lambda\cot\varphi_e}{2(5+4\rho)(1+2\rho)^2} \times \{(4+2\rho)[y^{(1)}]^2 - (3+20\rho+16\rho^2)[y^{(2)}]^2 + (16\rho+8\rho^2)(2+2\rho)^{1/2}y^{(1)}y^{(2)}\}\Bigg] \tag{29}$$

$$y^{(4)} = \frac{1}{1+2\rho}\Bigg[-y^{(1)} + (2+2\rho)^{1/2}y^{(2)} - \frac{\lambda\cot\varphi_e}{2(5+4\rho)(1+2\rho)^2} \times \{(8\rho+4\rho^2)(2+2\rho)^{1/2}[y(1)]^2 + (4+18\rho+8\rho^2)(2+2\rho)^{1/2}[y^{(2)}]^2 - (6+40\rho+32\rho^2)y^{(1)}y^{(2)}\}\Bigg] \tag{30}$$

Equations (29) and (30) satisfy the criterion

$$\frac{\partial y^{(3)}}{\partial y^{(2)}} = \frac{\partial y^{(4)}}{\partial y^{(1)}} \tag{31}$$

which is the condition for the complementary strain energy to be independent of path. The results (29) and (30) for the sphere reduce to those of Van Dyke [3]. The relations for $y^{(2)}$ and $y^{(3)}$ in terms of $y^{(1)}$ and $y^{(4)}$ are

$$y^{(2)} = (2+2\rho)^{-1/2}y^{(1)} + (1+2\rho)(2+2\rho)^{-1/2}y^{(4)} + \frac{\lambda\cot\varphi_e}{4(5+4\rho)(1+\rho)} \times \{(2\rho-2)[y^{(1)}]^2 + (4+18\rho+8\rho^2)[y^{(4)}]^2 + (2-8\rho)y^{(1)}y^{(4)}\} \tag{32}$$

$$y^{(3)} = (2+2\rho)^{-1/2}y^{(1)} - (2+2\rho)^{-1/2}y^{(4)} - \frac{\lambda\cot\varphi_e}{4(5+4\rho)(1+\rho)} \times \{3[y^{(1)}]^2 + (1-4\rho)[y^{(4)}]^2 - (4-4\rho)y^{(1)y(4)}\} \tag{33}$$

and the expressions for $y^{(1)}$ and $y^{(4)}$ in terms of $y^{(2)}$ and $y^{(3)}$ are

$$y^{(1)} = (2+2\rho)^{-1/2}\Bigg[y^{(2)} + (1+2\rho)y^{(3)} + \frac{\lambda\cot\varphi_e}{2(5+4\rho)(1+\rho)} \times \{-(1+2\rho)[y^{(2)}]^2 + (2+\rho)[y^{(3)}]^2 + 2(2+\rho)(1+2\rho)y^{(2)}y^{(3)}\}\Bigg] \tag{34}$$

$$y^{(4)} = (2+2\rho)^{-1/2}\Big[y^{(2)} - y^{(3)} + \frac{\lambda \cot\varphi_e}{2(5+4\rho)(1+\rho)} \times \{-(1+2\rho)[y^{(2)}]^2 - (1+2\rho)(2+\rho)[y^{(3)}]^2 + 2(1+2\rho)y^{(2)}y^{(3)}\}\Big] \quad (35)$$

Axial Edge Displacement

We now proceed to develop an expression for the axial displacement when the shell undergoes finite edge deformation due to self-equilibrating edge loads. The effect of surface load, which introduces a nonzero value for the term $A_0^{(1,3)}$ in $\mathbf{A}_0$, is included.

The compatibility condition which is written in terms of nonlinear meridional midsurface strain measure gives the following expression for axial displacement:

$$v = -h\cot\varphi + \int \frac{\epsilon_\varphi - \frac{r_2}{r_1}\epsilon_\theta}{\sin\varphi}\,ds - \int \frac{\chi^2}{2\sin\varphi}\,ds \quad (36)$$

For the edge bending, the first approximation is

$$\epsilon_\varphi \approx -\nu\epsilon_\theta \quad (37)$$

Substituting this into (36) and assuming that E, ν, r, r_1, t, and ϕ remain essentially constant within the narrow boundary layer, we obtain

$$v = -h\cot\varphi_e - \frac{\left(\frac{r_2}{r_1}+\nu\right)_e}{\sin\varphi_e}\int y^{(4)}ds - \frac{\lambda^2}{2\sin\varphi_e}\int [y^{(3)}]^2 ds \quad (38)$$

As the edge effect solution decays rapidly as we go away from the edge $s = 0$, the axial displacement at the edge is

$$v_e = -(h\cot\varphi)_e - \frac{\left(\frac{r_2}{r_1}+\nu\right)_e (cr_2)_e^{1/2}}{\sin\varphi_e}\int_{-\infty}^{0} y^{(4)}d\omega - \frac{\lambda^2 (cr_2)_e^{1/2}}{2\sin\varphi_e}\int_{-\infty}^{0}[y^{(3)}]^2 d\omega \quad (39)$$

where

$$\omega = s/(cr_2)_e^{1/2} \quad (40)$$

In (39), the first term on the right side is the main contribution from the linear solution for a steep shell; the second term gives the shallow shell correction, whereas the third provides the nonlinear effect.

$$\int_{-\infty}^{0} y^{(4)}d\omega = \Big[y^{(3)} + [2(1+\rho)]^{1/2}y^{(4)} + \frac{\lambda\cot\varphi(1+2\rho)}{2(5+4\rho)}\{[y^{(3)}]^2 + 2[y^{(4)}]^2\}\Big]_e \quad (41)$$

$$\int_{-\infty}^{0}[y^{(3)}]^2 d\omega = \{[y^{(3)}]^2 + [y^{(4)}]^2\}_e/[2(2+2\rho)^{1/2}]_e \quad (42)$$

Substituting these in (39) gives the nondimensional axial displacement at the edge $s = 0$

$$\left(\frac{v}{c}\right)_e = -\lambda^2\cos\varphi_e\left\{y^{(4)} + \frac{2\left(\frac{r_2}{r_1}+\nu\right)}{\lambda\sin 2\varphi}\Big[y^{(3)} + (2+2\rho)^{1/2}y^{(4)} + \frac{\lambda\cot\varphi(1+2\rho)}{2(5+4\rho)}\{[y^{(3)}]^2 + 2[y^{(4)}]^2\}\Big] + \frac{\lambda}{2\sin 2\varphi}\frac{\{[y^{(3)}]^2 + [y^{(4)}]^2\}}{(2+2\rho)^{1/2}}\right\}_e \quad (43)$$

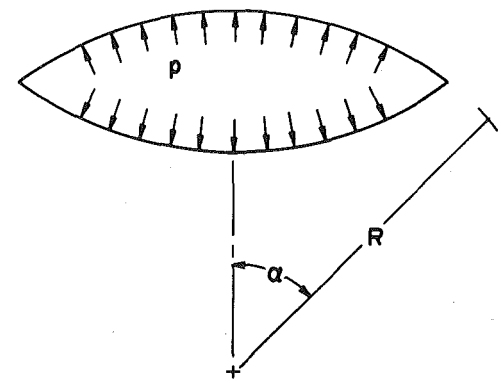


Fig. 1 Spherical caps joined with slope discontinuity as in pressure gauges

For a steep shell edge, the second term in square bracket in (43) is negligible. Hence, the ratio between the nonlinear and linear results for axial displacement at the shell edge is

$$\left(\frac{v_{\text{nonlinear}}}{v_{\text{linear}}}\right)_e = \frac{\left[y_e^{(4)} + \frac{\lambda}{2\sin 2\varphi_e}(2+2\rho)_e^{-1/2}\{[y^{(3)}]^2 + [y^{(4)}]^2\}\right]_{\text{nonlinear}}}{[y_e^{(4)}]_{\text{linear}}} \quad (44)$$

This relation will now be used to show the effect of nonlinearities in two specific problems.

Examples

A shell often used in pressure gauges is shown in Fig. 1. The two spherical caps intersect with a slope discontinuity. Then, for internal pressure the prestress term ρ given by (10) becomes,

$$\rho = PR^2/(4\,Etc) \quad (45)$$

At the edge with slope discontinuity, the membrane solution provides the following boundary conditions for the edge effect solution.

$$y^{(2)} = -(\rho\sin 2\alpha)/\lambda \qquad y^{(3)} = 0 \quad (46)$$

Using (35), $y^{(4)}$ becomes

$$[y^{(4)}]_{\text{linear}} = -(\rho\sin 2\alpha)/(\lambda(2)^{1/2}) \quad (47)$$

$$[y^{(4)}]_{\text{nonlinear}} = \frac{[y^{(4)}]_{\text{linear}}}{(1+\rho)^{1/2}}\left[1 + \frac{\rho(1+2\rho)\cos^2\alpha}{(5+4\rho)(1+\rho)}\right] \quad (48)$$

Substituting these in (44), we obtain

$$\frac{v_{\text{nonlinear}}}{v_{\text{linear}}} = (1+\rho)^{-1/2} - \frac{\rho(1+\rho)^{-3/2}}{4}\left\{1 - \frac{4(1+2\rho)\cos^2\alpha}{(5+4\rho)} + \frac{2\rho(1+2\rho)\cos^2\alpha}{(5+4\rho)(1+\rho)}\right\} \quad (49)$$

The first term in (49) provides the nonlinear prestress effect while the second term within the set of braces gives the effect of the quadratic rotation term. These effects are computed for specific situations and are tabulated in Table 1. For this case the edge rotation is constrained so the dominant nonlinear effect is due to the prestress term, even though the total radial stress resultant is zero at the boundary.

A second example of a spherical dome resting on roller supports with hinged edges is shown in Fig. 2. As in the previous example, the radial stress resultants due to bending and membrane solutions are of the same magnitude and opposite sign. The prestress parameter ρ remains the same as (45). At the edge with roller supports, the membrane solution provides the following boundary conditions for the bending solution:

$$y^{(1)} = 0 \qquad y^{(2)} = -(\rho\sin 2\alpha)/\lambda \quad (50)$$

Table 1 Nonlinear effects in pressure gauge

	Only prestress term included		Both nonlinearities considered					
			$\cos^2\alpha \approx 1$		$\alpha = 30°$		$\alpha = 60°$	
ρ	−0.25	1.0	−0.25	1.0	−0.25	1.0	−0.25	1.0
$\frac{v_{\text{nonlinear}}}{v_{\text{linear}}}$	1.1547	0.7071	1.1948	0.7071	1.2088	0.6850	1.2369	0.6408

Table 2 Nonlinear effects on the pressurized dome

	only prestress term included		Both nonlinearities considered					
			$\cos^2\alpha \approx 1$		$\alpha = 30°$		$\alpha = 60°$	
ρ	−0.25	1.0	−0.25	1.0	−0.25	1.0	−0.25	1.0
$\frac{v_{\text{nonlinear}}}{v_{\text{linear}}}$	1.732	0.4714	2.598	0.4707	2.562	0.4463	2.490	0.3976

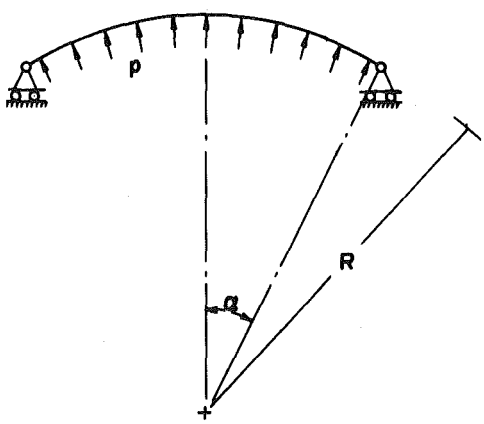


Fig. 2 Spherical dome on roller supports

Using (29) and (30) gives the linear solution

$$[y^{(4)}]_{\text{linear}} = (2)^{1/2} y^{(2)} \tag{51}$$

and the nonlinear solution becomes

$$[y^{(3)}]_{\text{nonlinear}} = -\frac{y^{(2)}}{(1+2\rho)} \times \left[1 - \frac{\lambda \cot\alpha\,(3 + 20\rho + 16\rho^2)}{2(5+4\rho)(1+2\rho)^2}\, y^{(2)}\right] \tag{52}$$

$$[y^{(4)}]_{\text{nonlinear}} = \frac{(2+2\rho)^{1/2} y^{(2)}}{(1+2\rho)} \times \left[1 - \frac{\lambda \cot\alpha(4 + 18\rho + 8\rho^2)}{2(5+4\rho)(1+2\rho)^2}\, y^{(2)}\right] \tag{53}$$

Substituting these in (44), we obtain

$$\frac{v_{\text{nonlinear}}}{v_{\text{linear}}} = \frac{(1+\rho)^{1/2}}{(1+2\rho)} - \frac{\rho(3+2\rho)}{4(1+\rho)^{1/2}(1+2\rho)^2} \times \left[1 - \frac{8(1+\rho)(2+\rho)(1+4\rho)\cos^2\alpha}{(3+2\rho)(5+4\rho)(1+2\rho)} + \frac{2\rho(11 + 64\rho + 68\rho^2 + 16\rho^3)\cos^2\alpha}{(3+2\rho)(5+4\rho)(1+2\rho)^2}\right] \tag{54}$$

The first term in (54) provides the effect of the first geometric nonlinearity, namely, the prestress effect, while the second term within the set of brackets gives the effect of quadratic rotation terms. These effects are tabulated for specific cases in Table 2.

It is somewhat surprising that even in this case, when the dome is internally pressurized so that ρ is positive, the prestress term provides the main portion of the nonlinear effect. When the dome is externally pressurized so that ρ is negative, the effects of both the nonlinearities are significant. Hence, it is important to include both the nonlinearities in stability problems. In [4, 5] for the spherical shell dimpling under a point load, it is found that the prestress term is virtually negligible but the quadratic rotation terms have a significant effect.

References

1 Reissner, E., "On Axisymmetrical Deformations of Thin Shells of Revolution," *Proceedings of Symposia in Applied Mathematics*, Vol. 3, McGraw-Hill, New York, 1950, pp. 27–52.

2 Reissner, E., "On Influence Coefficients and Nonlinearity for Thin Shells of Revolution," ASME JOURNAL OF APPLIED MECHANICS, Vol. 26, Mar. 1959, pp. 69–72.

3 Van Dyke, P., "Nonlinear Influence Coefficients for a Spherical Shell and Pressure Loading," *AIAA Journal*, Vol. 4, No. 11, Nov. 1966, pp. 2045–2047.

4 Ranjan, G. V., "Edge Zone Expansion for Thin Shells With Application to (1) Torispherical Pressure Vessel and (2) Large Deflection of Sphere With Point Load," PhD Dissertation, Stanford University, Nov. 1976.

5 Ranjan, G. V., and Steele, C. R., "Large Deflection of Deep Spherical Shells Under a Concentrated Load," *Proceedings, 18th Structures, Structural Dynamics, and Materials Conference*, San Diego, Mar. 1979.

6 Steele, C. R., and Skogh, J., "Slope Discontinuities in Pressure Vessels," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, Sept. 1970, pp. 587–595.

T. Yabuta

Engineer,
Outside Plant Development Division,
Ibaraki Electrical Communication Laboratory,
Nippon Telegraph and Telephone Public Corp.,
Tokai, Ibaraki-ken, 319-11, Japan

Effects of Elastic Supports on the Buckling of Circular Cylindrical Shells Under Bending

This paper presents the effects of elastic supports on the buckling of circular cylindrical shells under bending. Stability was investigated using Donnell's equation and the Galerkin method, including the spring constant of the elastic support. The results of this investigation indicate that the effects are similar in the cases of bending and axial compression.

1 Introduction

A submarine coaxial cable (cross-sectional view is shown in Fig. 1) used for international communications has bending stress applied when it is laid by ship. When a submarine coaxial cable is bent, buckling occurs in the outer conductor (shown as the cylindrical shell in Fig. 1). Cable bending characteristics become worse, if the Young's modulus for both sheath and dielectric become low [1].

This paper reports results of an investigation on the buckling of the elastically supported cylindrical shell, under bending, as the first step in clarifying the buckling of the outer conductor, assuming that both sheath and dielectric act as elastic supports for the outer conductor.

This analysis is an extension of the work done by Seide [2, 3]. Batdorf's modified Donnell's equation for buckling is used, in conjunction with the Galerkin method [4], to obtain the critical bending stress for a cylinder from the small deflection theory, introducing the spring constant of the elastic support obtained by Seide [2]. Critical bending stresses are calculated for a cylinder with both inner and outer elastic support.

Experimental results are shown and compared with numerical results. On the whole, the experimental results are qualitatively in agreement with the theoretical results.

2 Theory

Notations for the problem under consideration are shown in Fig. 1. An equation for the buckling stress is provided by Batdorf's modification of Donnell's equation [4], which may be written as

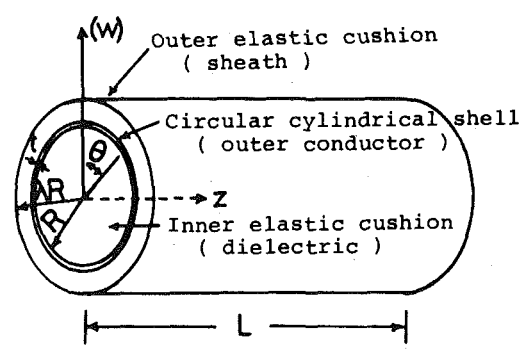


Fig. 1 Notation for problem under consideration

$$Q(w) = D\nabla^4 w + \frac{Et}{R^2}\nabla^{-4}\left(\frac{\partial^4 w}{\partial z^4}\right) - \left(t\sigma_z\frac{\partial^2 w}{\partial z^2} + 2t\tau_{z\theta}\frac{\partial^2 w}{R\partial z\partial\theta} + t\sigma_\theta\frac{\partial^2 w}{R^2\partial\theta^2}\right) + P(z,\theta) = 0 \quad (1)$$

Stresses prior to buckling are given by the following equation, using both σ_b (maximum stress due to bending) and σ_c (stress due to compression), when considering a cylinder under combined bending and compression.

$$\sigma_z = -(\sigma_c + \sigma_b \cos\theta)$$
$$\sigma_\theta = \tau_{z\theta} = 0 \quad (2)$$

Radial deflection is assumed to be in the form

$$w = \sin\frac{m\pi z}{L}\sum_{n=0}^{\infty} a_n \cos(n\theta) \quad (3)$$

which satisfies the end conditions of simple support.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, July, 1979; final revision, April, 1980.

Pressure by the elastic support is assumed to be in the form

$$P(z, \theta) = \sum_{n=0}^{\infty} k_n^*(\lambda, n, \overline{R})\, a_n \sin \frac{m\pi z}{L} \cos (n\theta) \tag{4}$$

where $k_n^*(\lambda, n, \overline{R})$ is the equivalent spring constant, discussed in Section 3.

If the Galerkin method is employed to satisfy equation (1) in order to derive the stability criterion, the following equation is obtained:

$$\int_0^{2\pi} \int_0^L Q(w) \sin \frac{p\pi z}{L} \cos (q\theta) R dz d\theta = 0;$$

$$p = 1, 2, 3 \ldots, \quad q = 0, 1, 2 \ldots \tag{5}$$

The following system of homogenous equation is obtained after integrating equation (5):

$$a_n \left[\frac{(\overline{R} + n^2/\overline{R})^2}{S} + \frac{S}{(\overline{R} + n^2/\overline{R})^2} - 2h_c + 2K_n^*/\overline{R}^2 \right]$$

$$-h_b[(1 + \delta_{1n} - \delta_{0n})a_{n-1} + a_{n+1}] = 0; \quad n = 0, 1, 2 \ldots \tag{6}$$

where

$$\delta_{jn} = \begin{cases} 1 & \text{if} \quad n = j \\ 0 & \text{if} \quad n \neq j \end{cases}$$

The stability criterion is determined by the condition that the coefficient determinant of equation (6) vanishes. The critical bending stress is obtained by minimizing the value h_b with respect to wavelength parameter $\overline{R}$.

3 Spring Constant

The spring constant was obtained by Seide [2] for the elastic support, when shear stress between the elastic support and cylinder is neglected. According to Seide [2], the spring constant is expressed as a function of the longitudinal wavelength parameter $\overline{R}$, circumferential wave number n, and thickness parameter λ, as follows:

$$k_n(\lambda, n, \overline{R}) = \frac{E_c}{1 + \nu_c} \frac{1}{R} f_n(\lambda, n, \overline{R}) \tag{7}$$

where $f_n(\lambda, n, \overline{R})$ is a nondimensional spring constant described in the Appendix.

When a cylindrical shell has different spring constants between inner and outer elastic cushion, and is fastened to the inner and outer elastic cushions, the equivalent spring constant is given by

$$k_n^*(\lambda, n, \overline{R}) = k_1(\lambda_1, n, \overline{R}) + k_2(\lambda_2, n, \overline{R}) \tag{8}$$

where λ_1 and λ_2 are the ratios of the inner and outer cushion radii to the cylinder radius.

Consider the problem in Fig. 1. As the inner elastic cushion has no hole, λ_1 is equal to zero and only λ_2 varies. Therefore equation (8) is written as follows:

$$k_n^*(\lambda, n, \overline{R}) = k_1(0, n, \overline{R}) + k_2(\lambda, n, \overline{R}) \tag{9}$$

By introducing equation (9) into equation (4), effects of elastic cushion on the buckling of the cylindrical shell can be calculated.

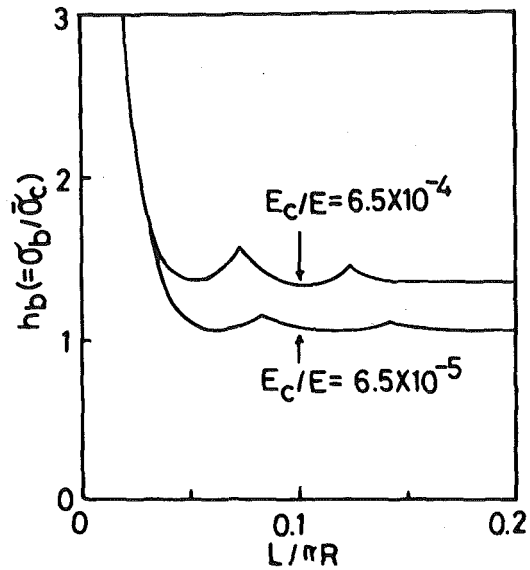


Fig. 2 Relation between buckling stress and cylinder length

4 Numerical Results and Discussion

For given values of radius-thickness parameter S, spring constant parameter K_n^* and axial compression parameter h_c, critical bending stresses are obtained as h_b values minimized with respect to wavelength parameter $\overline{R}$ from a suitable truncated coefficient determinant in equation (6). The size of the determinant is increased until three significant critical stress figures are assured. For example, 14 terms are used in deflection function for $R/t = 100$, $E_c/E = 6.5 \times 10^{-4}$.

Numerical results are shown in Figs. 2–8. Results shown in Figs. 2–7 are calculated for the cylinder with only an inner elastic cushion core in order to simplify the problem. Therefore the equivalent spring constant is $k_n^*(\lambda, n, \overline{R}) = k_1(0, n, \overline{R})$. Fig. 8 shows the results for a cylinder with both inner and outer elastic supports having the same Young's modulus and with Poisson's ratios ν and ν_c taken as 0.3.

The relation between critical bending stress and cylinder length is shown in Fig. 2. Numerical results are shown for $E_c/E = 6.5 \times 10^{-4}$ and $E_c/E = 6.5 \times 10^{-5}$. It can be seen that critical stress increases extremely as the length becomes short. As the length increases, the critical bending stress becomes almost constant and is affected by the elastic support. Minimized h_b values are used in the following discussion.

Fig. 3 represents the relation between critical bending stress and Young's modulus for the elastic support. The broken line in Fig. 3 indicates the critical compressive stress for a cylinder with an inner elastic support obtained as follows. The equation to determine the critical compressive stress is obtained from equation (6) introducing $h_b = 0$.

Nomenclature

a_n = coefficient in deflection function
D = flexural stiffness of cylinder wall, $Et^3/12(1 - \nu^2)$
E, E_c = Young's modulus of cylinder and elastic cushion, respectively
$f_n(\lambda, n, \overline{R})$ = nondimensional spring constant
h_b = buckling stress ratio for bending, $\sigma_b/\overline{\sigma}_c$
h_c = buckling stress ratio for compression, $\sigma_c/\overline{\sigma}_c$
$k_n(\lambda, n, \overline{R})$ = spring constant for n circumferential wave
k_1 = outer elastic spring constant
k_2 = inner elastic spring constant
k_n^* = equivalent spring constant for n circumferential wave
K_n^* = spring constant parameter, $k_n^*[3(1 - \nu^2)]^{1/2}R^3/Et^2$
L = cylinder length
m, n = integers
R = cylinder radius
$\overline{R}$ = wavelength parameter, $m\pi R/L$
S = radius-thickness parameter, $[12(1 - \nu^2)]^{1/2}R/t$
w = radial deflection
σ_b = maximum stress due to bending
σ_c = stress due to compressive load
$\overline{\sigma}_c$ = theoretical compressive buckling stress for a long cylinder, $Et/[3(1 - \nu^2)]^{1/2}R$
σ_θ, σ_z = direct stresses
$\tau_{z\theta}$ = shear stress
λ = thickness parameter
ν, ν_c = Poisson's ratio of cylinder and elastic cushion, respectively
∇^4 = operator, $(\partial^4/\partial z^4 + 2\partial^4/R^2\partial z^2\partial\theta^2 + \partial^4/R^4\partial\theta^4)$

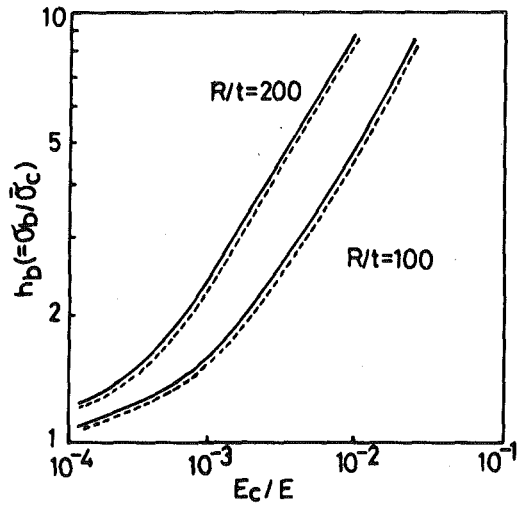


Fig. 3 Relation between buckling stress and elastic cushion stiffness

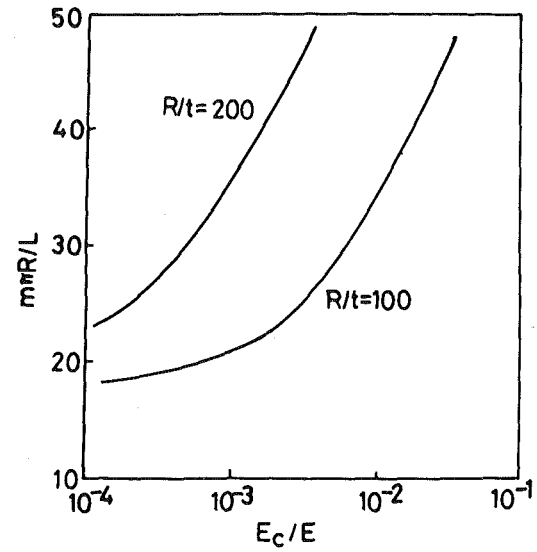


Fig. 5 Relation between wavelength parameter and elastic cushion stiffness

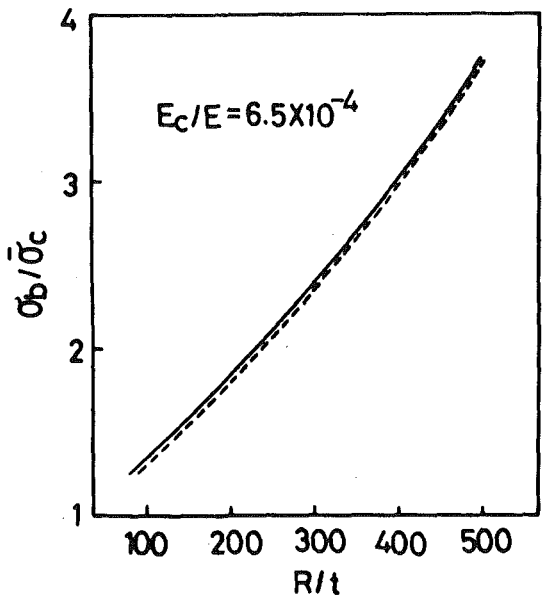


Fig. 4 Relation between buckling stress and radius-thickness ratio

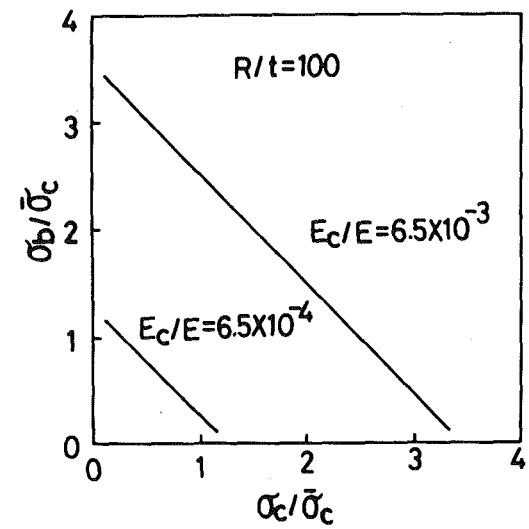


Fig. 6 Interaction between axial load and bending stress

$$h_c = \frac{1}{2}\left[\frac{(\bar{R}+n^2/\bar{R})^2}{S} + \frac{S}{(\bar{R}+n^2/\bar{R})^2}\right] + \frac{K_n{}^*}{\bar{R}^2} \tag{10}$$

The critical compressive stresses are obtained as h_c values minimized with respect to wavelength parameter $\bar{R}$, because the axismetric mode ($n = 0$) yields lower compressive buckling stress than the other mode [2]. Comparison between bending and compression results indicates that effects of the elastic supports are nearly same.

Fig. 4 indicates the relation between critical bending stress and cylinder thickness. The effect of the elastic support increases as R/t increases. For a cylinder without elastic support, h_b is approximately 1 for all values of R/t [3]. The compression results from equation (9) are shown for comparison and are seen to coincide with those for pure bending.

Table 1 indicates the comparison of critical stresses between bending and compression. The approximation of critical compressive stress is also indicated from equation (11), which is accurate for values of h_c less than 1.5 [2].

$$h_c{}^* = 1 + \frac{\sqrt[4]{12(1-\nu^2)}}{4(1-\nu_c{}^2)}\frac{E_c}{E}\left(\frac{R}{t}\right)^{3/2} \tag{11}$$

Table 1 Comparison between bending and compression

E_c/E	h_b	h_c	h_c^*
6.50×10^{-4}	1.33	1.31	1.32
3.25×10^{-3}	2.43	2.37	–
6.50×10^{-3}	3.56	3.48	–

h_c^* is obtained from Eq. (11)

The critical compressive stress h_c is a few percent less than the critical bending stress h_b.

The results between wavelength parameter $\bar{R}$ and Young's modulus for the elastic support are shown in Fig. 5. It can be seen that wavelength parameter $\bar{R}$ increases as Young's modulus for the elastic support increases. This effect becomes especially significant when E_c/E is greater than 10^{-3} for $R/t = 100$ and when E_c/E is greater than 5×10^{-4} for $R/t = 200$.

The effect of additional compression on a cylinder under bending

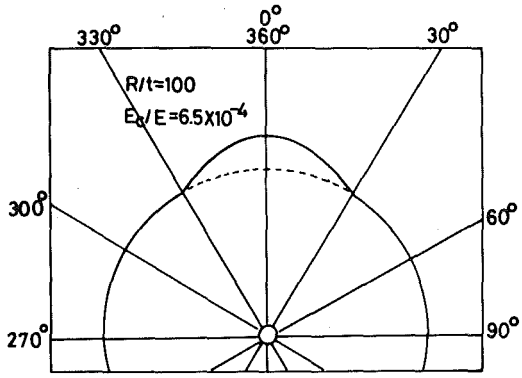


Fig. 7 Circumferential deflection shape

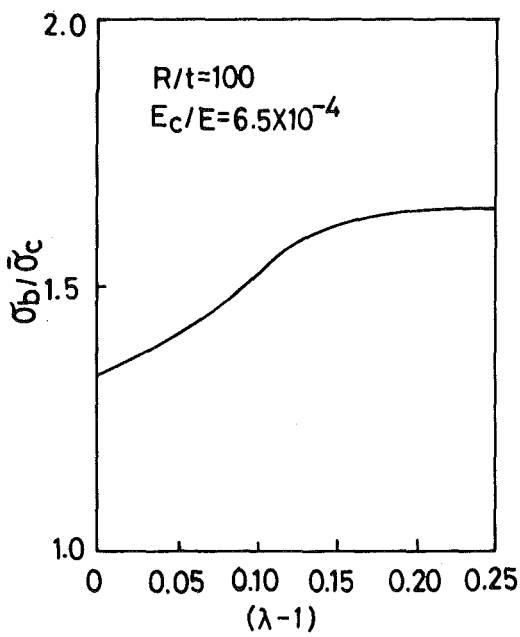


Fig. 8 Relation between buckling stress and outer elastic cushion thickness

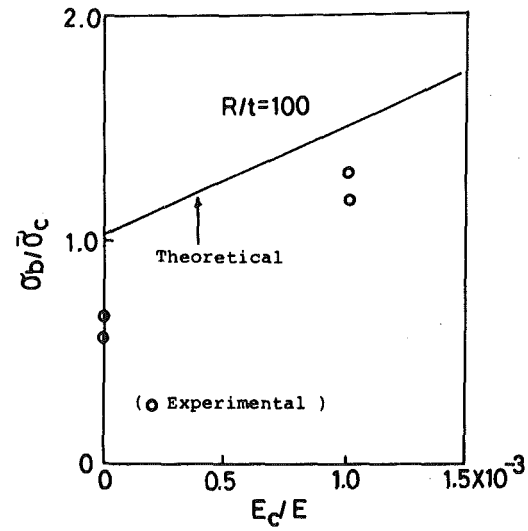


Fig. 9 Theory and experimental results comparison

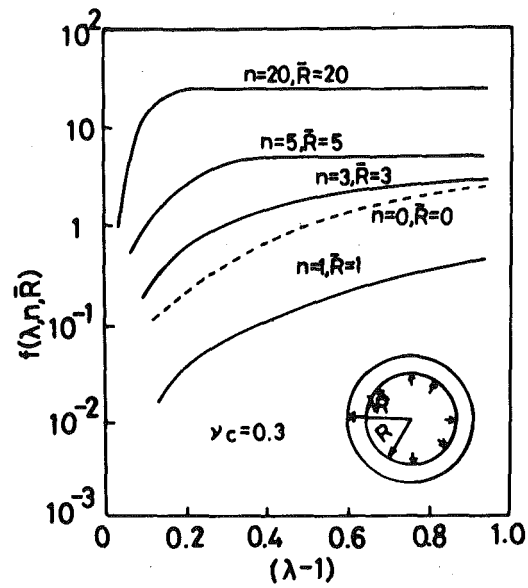


Fig. 10 Relation between nondimensional spring constant and elastic cushion thickness parameters

is shown in Fig. 6. Fig. 6 indicates that the critical bending stress σ_b decreases linearly as the axial compression stress σ_c increases. This variation does not appear to be affected by the presence of the elastic support.

The circumferential deflection is computed with equation (3) and the eigenvector $\{a_n\}$ associated with the minimum eigenvalues. The result for a cylinder whose $R/t = 100$ and $E_c/E = 6.5 \times 10^{-4}$ is shown in Fig. 7.

Calculated results, for a cylinder with both inner and outer elastic cushions, whose Young's moduli are the same, $E_c/E = 6.5 \times 10^{-4}$ (shown in Fig. 1), are shown in Fig. 8. Calculated h_b values for $R/t = 100$ are calculated with respect to the outer thickness parameter λ. The thickness parameter λ indicates that the thickness of the elastic cushion vanishes, when λ is equal to 1, and the thickness increases as λ increases. The effect of the outer elastic cushion is pronounced when the thickness parameter $(\lambda - 1)$ is less than 0.15. However, when the parameter $(\lambda - 1)$ is more than 0.15, the effect becomes small and critical bending stress σ_b converges. This is why the spring constant depends upon the thickness parameter λ, which is discussed in the Appendix.

5 Experimental Investigation

An experimental study was conducted to compare experimental results with the numerical results, using cylinders with inner elastic core and cylinders without an elastic cushion. The cylindrical walls of the test specimens were made of Mylar (polyester film) and the elastic cushion was a silicon rubber rod. Two of the test specimens contained no silicon rubber and the other two contained a silicon rubber rod as the inner elastic cushion. Details of the specimens are described in Table 2.

Table 2 Test specimen details

Specimen No	R(mm)	t(mm)	L(mm)	R/t	L/R	E(N/m²)	E_c(N/m²)	Note
1	20	0.2	70	100	3.5	4x10⁹	0	without elastic cushion
2	20	0.2	68	100	3.4	4x10⁹	0	
3	20	0.2	78	100	3.9	4x10⁹	4x10⁶	with elastic cushion
4	20	0.2	80	100	4.0	4x10⁹	4x10⁶	

Test specimens were bent, little by little, by a bending instrument

made for cable bending tests. Critical bending data were obtained from an automatic graphic record of strain gage readings as a function of varying bending radius, since it had been found that there were sharp discontinuities in these records when buckling occurred [2].

Comparison between experimental results and theory is shown in Fig. 10. The experimental results are less than the theoretical values and the difference is larger for a cylinder without an elastic core than for a cylinder with an elastic core. This tendency agrees with the compression results reported by Seide [2] and Almroth [5]. As the small deflection theory indicates the upper bound for the critical compressive buckling stress in reference [5], this analysis is seen as the upper bound for the critical bending stress.

6 Conclusion

The stability of a circular cylindrical shell, supported by elastic cushions and subjected to bending, was investigated using small deflection theory. The results indicate that the effect of the cushions on the buckling of circular cylindrical shells under bending is similar as the effect under compression. By comparison between theoretical and experimental results, this analysis is expected to indicate the upper bound for the critical bending stress.

Acknowledgment

The authors thanks N. Kojima and Y. Negishi for their encouragement and their effective suggestions.

Reference

1 Yabuta, T, and Negishi, Y., "Submarine Coaxial Cable Structure Designing," *Transaction of the Institute of Electrical Engineers of Japan,* Vol. 98-B, No. 5, 1978, pp. 441–448.

2 Seide, P., "The Stability Under Axial Compression and Lateral Pressure of Circular Cylindrical Shells With a Soft Elastic Core," *Journal of the Aerospace Sciences,* Vol. 29, No. 7, 1962, pp. 851–862.

3 Seide, P., and Weingarten, V. I., "On the Buckling of Circular Cylindrical Shells Under Pure Bending," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, No. 1, 1961, pp. 112–116.

4 Batdorf, S. B., "A Simplified Method of Elastic-Stability Analysis for Thin Cylindrical Shells, II—Modified Equilibrium Equation," NACA TN 1342, 1947

5 Almroth, B. O., and Brush, D. O., "Postbuckling Behavior of Pressure or Core-Stabilized Cylinders Under Axial Compression," *American Institute of Aeronautics and Astronautics Journal,* Vol. 1, No. 10, 1963, pp. 2338–2341

APPENDIX

Nondimensional Spring Constant

The determination of the spring constant is reported by Seide in reference [2]. Although the procedure of reference [2] applies for a cushion with a central hole, the result is only given for an elastic cushion without a central hole. The same procedure for the calculating of the spring constant of an elastic cushion with central hole is used here in order to determine the effect of the thickness. The procedure is omitted here for brevity. Only the numerical results are shown in Fig. 10. Fig. 10 represents the relation between the nondimensional spring constant for the outer elastic cushion in Fig. 1 and the thickness. When $\overline{R}$ and n are large, the nondimensional spring constant f rapidly converges to a constant value in the small λ range. However, when $\overline{R}$ and n are small, f increases gradually as λ increases.

J. R. Barber
Department of Mechanical Engineering,
University of Newcastle upon Tyne,
Newcastle upon Tyne NE1 7RU,
England

J. Dundurs
Professor,
Department of Civil Engineering,
Northwestern University,
Evanston, Ill. 60201
Fellow ASME

M. Comninou
Associate Professor,
Department of Civil Engineering,
The University of Michigan,
Ann Arbor, Mich. 48109
Mem. ASME

Stability Considerations in Thermoelastic Contact

A simple one-dimensional model is described in which thermoelastic contact conditions give rise to nonuniqueness of solution. The stability of the various steady-state solutions discovered is investigated using a perturbation method. The results can be expressed in terms of the minimization of a certain energy function, but the authors have so far been unable to justify the use of such a function from first principles in view of the nonconservative nature of the system.

Introduction

It is known that mathematical difficulties arise in problems of thermoelastic contact between bodies with geometrically smooth surfaces. In particular the conventional boundary conditions of perfect thermal contact (no resistance to heat flow leading to continuity of temperature) in regions of mechanical contact and complete insulation (no heat flux) in regions of separation lead to ill-posed boundary-value problems whenever the hotter solid has the higher thermal distortivity defined by

$$\delta = \alpha(1 + \nu)/k \tag{1}$$

where α, k, ν are, respectively, the coefficient of linear thermal expansion, thermal conductivity, and Poisson's ratio for the material [1–3]. An asymptotic analysis of the transitions between the various boundary regions [3] suggests that the conventional boundary conditions can be safely applied when heat flows in the opposite direction, but there is evidence that, in this case, the solution obtained is not necessarily unique [4]. This immediately raises the question of stability which forms the subject of this paper.

Stability questions can be probed by energy arguments or by an analysis of small perturbations about the steady state. The former would have the advantage of analytic simplicity, since a perturbation analysis involves a consideration of transient heat conduction, but it is far from clear what energy formulation would be appropriate for thermoelastic contact, since the system is inherently nonconservative, i.e., it is possible to devise loading cycles which would cause the contacting solids to act as a heat engine. In order to elucidate this question and to investigate the fundamental characteristics of thermoelastic contact, we give here an exhaustive treatment of the simplest contact system which exhibits thermoelastic nonuniqueness—a one-dimensional rod conducting heat between rigid walls. A perturbation analysis is used to determine the conditions for stability of the various steady-state solutions, but it transpires that these conditions can be stated in terms of the minimization of an energy function with a straightforward physical interpretation.

The Model: Steady-State Solutions

The system to be analyzed is illustrated in Fig. 1. Two perfectly conducting and rigid walls, separated by a distance l, are maintained at temperatures T_0 and zero, respectively, and a uniform elastic rod of cross-sectional area A is built into the cold wall as shown. The length of the rod is such as to leave a gap $g = g_0$ when the temperature is everywhere zero. The same model, with the temperature difference reversed, has been used to investigate the nonexistence of solutions in thermoelastic contact [2].

For sufficiently high values of the hot wall temperature T_0, we should expect two steady-state solutions to the problem: One involving contact between the rod and the wall, the gap being closed by thermal expansion of the rod; the other with separation between the rod and the wall, the gap being sufficient to prevent significant heat flow into the rod. To investigate the matter further, we postulate the existence of a thermal resistance R between the hot wall and the rod, which will be a function of contact force P when contact occurs and of gap size g when it does not. No assumptions will be made about the nature of this function though, on physical grounds, we should expect

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, February, 1980. Paper No. 80-WA/APM-25.

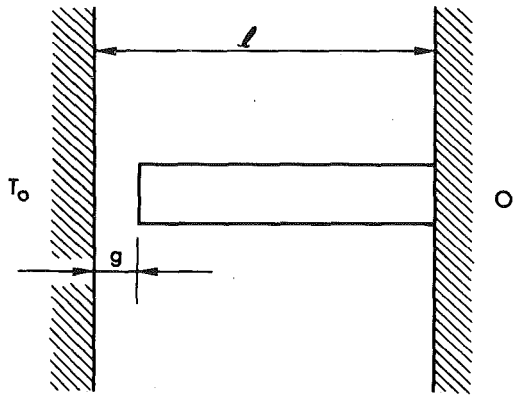


Fig. 1 The rod transferring heat from the hot to the cold wall

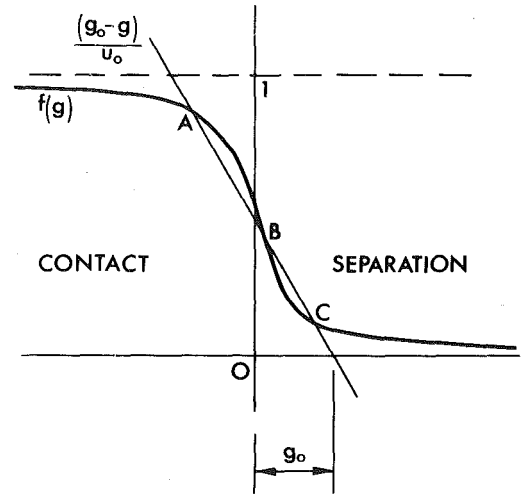


Fig. 2 Graphical interpretation of the stability criterion

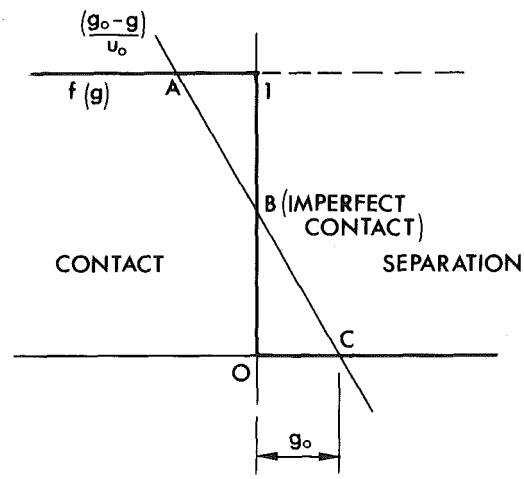


Fig. 3 Graphical interpretation for the idealized boundary conditions

it to fall monotonically as the gap is reduced or the pressure increased.

If the contact resistance R is known, the temperature T' at the hot end of the rod can be determined from continuity of heat flux Q in the steady state. Thus

$$Q = (T_0 - T')/R = T'Ak/l \tag{2}$$

Solving for T', we have

$$T' = \frac{T_0}{1 + AkR/l} \tag{3}$$

and hence the unrestrained thermal expansion of the rod is

$$u_{th} = \frac{1}{2}\alpha T' l = \frac{\alpha l^2 T_0}{2(l + AkR)} = u_0 f \tag{4}$$

where $u_0 = \frac{1}{2}\alpha l T_0$ is the thermal expansion which would be developed with $T' = T_0$, i.e., perfect thermal contact between the rod and the wall, and

$$f = \frac{l}{l + AkR} \tag{5}$$

The function f ranges from zero for complete thermal insulation ($g \to \infty$) to unity for perfect thermal contact ($P \to \infty$).

We can now write down the equations determining the steady-state solutions. For separation,

$$g = g_0 - u_0 f(g), \quad g \geq 0 \tag{6a}$$

whereas for contact,

$$0 = g_0 - u_0 f(P) + Pl/AE, \quad P \geq 0 \tag{6b}$$

The variables g and P apply to separate regimes which intersect only in the point $g = P = 0$. We can therefore define a continuation of g into negative values by the relation

$$g = -Pl/AE, \quad P > 0 \quad \text{and} \quad g < 0 \tag{7}$$

With this definition, the two equations (6a, b) reduce to the same form which is conveniently written

$$f(g) = (g_0 - g)/u_0 \tag{8}$$

A graphical solution to equation (8) could be envisaged as shown in Fig. 2. The two sides of the equation are plotted as separate functions of g, and steady-state solutions are represented by intersections between these functions. In general, there will be either one or three solutions, depending on the values of g_0, u_0 and the nature of the function $f(g)$.

It is instructive to examine the limiting case where the contact resistance passes from zero to infinity over an infinitesimal range of values about $g = 0$. The corresponding limit for $f(g)$ is the step function $f(g) = H(-g)$ shown in Fig. 3. When three steady-state solutions occur, two of them lie on the horizontal branches of the step functions at A and C, corresponding to perfect thermal contact and separation, respectively, while the third lies on the vertical step at B and corresponds to the state defined by a similar limiting process as "imperfect contact" in reference [2].

Stability Analysis

In order to investigate the stability of the various steady states described by equation (8), we examine the conditions under which a small perturbation can grow exponentially with time. Such a perturbation will only be possible for certain eigenvalues of the exponential growth rate, and the condition for stability is that there should be no positive eigenvalues. If complex eigenvalues are possible—corresponding to exponentially growing oscillatory perturbations—there must be none with a positive real part.

Temperature Distribution in the Bar. The perturbation in temperature and heat flux in the bar from the steady-state value must satisfy the transient heat conduction equation

$$\frac{\partial^2 T}{\partial x^2} = \frac{1}{\kappa}\frac{\partial T}{\partial t} \tag{9}$$

where κ is the thermal diffusivity of the bar material and x is measured along the bar from the cold end.

Assuming a perturbation of the form $T = \phi(x)e^{at}$, we have

$$\frac{d^2\phi}{dx^2} - \frac{a}{\kappa}\phi = 0 \tag{10}$$

the solution of which is

$$\phi = A \cosh \lambda x + B \sinh \lambda x \tag{11}$$

with $\lambda = (a/\kappa)^{1/2}$. Applying the condition $T = 0$ at $x = 0$, the constant A must be zero, and hence the perturbation in temperature is described by

$$T = Be^{at} \sinh \lambda x \tag{12}$$

while the heat flux is

$$Q_x = -Ak\frac{\partial T}{\partial x} = -BAk\lambda e^{at} \cosh \lambda x \tag{13}$$

The change in the gap between the rod and the wall due to this temperature perturbation is

$$\Delta g = -\int_0^l \alpha T dx = -B\frac{\alpha}{\lambda}e^{at}(\cosh \lambda l - 1) \tag{14}$$

and the perturbations in heat flux Q and temperature T at the free end $x = l$ are

$$\Delta Q = -Q_x(l) = BAk\lambda e^{at} \cosh \lambda l \tag{15}$$

$$\Delta T = T(l) = Be^{at} \sinh \lambda l \tag{16}$$

Contact Resistance Equation. To complete the analysis, we linearize the relation (2) between heat flux and temperature in the vicinity of the steady-state solution, obtaining

$$\Delta Q = -\frac{(T_0 - T')\Delta R}{R^2} - \frac{\Delta T}{R} \tag{17}$$

where ΔR is the corresponding perturbation in contact resistance and R, T' here describe the steady-state values.

Equation (17) can be cast in terms of the function $f(g)$ by using equation (5). Thus

$$R = \frac{l}{Ak}\left(\frac{1}{f} - 1\right) \tag{18}$$

and hence

$$\Delta R = \frac{\partial R}{\partial g}\Delta g = -\frac{lf'\Delta g}{Akf^2} \tag{19}$$

Substituting for T', R, ΔR from equations (3), (18), and (19), respectively, into equation (17), we obtain

$$\frac{l}{Ak}\left(\frac{1}{f} - 1\right)\Delta Q = \frac{T_0 f'\Delta g}{f} - \Delta T \tag{20}$$

Characteristic Equation. Finally, we substitute for the perturbations Δg, ΔQ, ΔT from equations (14)–(16) into (20) to obtain the characteristic equation for a which is

$$B\lambda l\left(\frac{1}{f} - 1\right)e^{at} \cosh \lambda l = -B\alpha T_0 \frac{f'}{\lambda f}e^{at}(\cosh \lambda l - 1) - Be^{at} \sinh \lambda l \tag{21}$$

and which can be simplified to the form

$$(1 - f)y^2 \cosh y + 2u_0 f'(\cosh y - 1) + fy \sinh y = 0 \tag{22}$$

where

$$y^2 = al^2/\kappa \tag{23}$$

Restricting attention initially to the case of real roots, we expand equation (22) in the form

$$\sum_{j=1}^{\infty} \frac{y^{2j}}{(2j)!}[2j(2j - 1)(1 - f) + 2jf + 2u_0 f'] = 0 \tag{24}$$

The functions f, $(1 - f)$ are both positive for positive values of R (see the section, "The Model: Steady-State Solutions") and hence the series will be positive for large values of y. However, it will be negative at small values of y, giving a zero somewhere on the real axis if

$$2(1 - f) + 2f + 2u_0 f' < 0 \tag{25}$$

or

$$-f' > 1/u_0 \tag{26}$$

(Note that f' is generally negative—see Fig. 2). Furthermore, it is clear that if this condition is *not* satisfied, all the terms of equation (24) will be positive and there will be no real root for y (except the trivial solution $y = 0$).

Of course, equation (22) may have complex roots, describing oscillatory perturbations. This possibility is investigated in the Appendix, where it is shown that (26) also describes the condition for a complex root for the exponential growth rate a to have a positive real part. Hence, we conclude that the system is unstable *if and only if* condition (26) is satisfied.

Graphical Interpretation and History Dependence. By referring to equation (8) it is clear that the criterion (26) for instability describes those intersections in Fig. 2 at which the function $f(g)$ crosses $(g_0 - g)/u_0$ from above with increasing g. Thus solution B is unstable, while A and C are stable.

The steady-state solution C is possible only if the imposed temperature T_0 is smaller than a certain temperature, say, T_C which can be determined graphically for given g_0 from the curve $f(g)$ in Fig. 2. Similarly, solution A can occur only if T_0 is above a certain temperature, say, T_A. All three solutions A, B, and C are possible in the intermediate range $T_C < T_0 < T_A$. Which of the stable steady states are reached depends on the history of the thermal process. Suppose that the rod has a certain initial temperature distribution and that, during the early stages, T_0 depends on time, but that later in the process T_0 is kept constant so that a steady-state distribution of temperature in the rod is eventually achieved. It is then clear that the final state depends on the previous manipulations of the process. Consider for instance the rod being initially at zero temperature with T_0 slowly raised from zero to some finite value. In such case, the steady state reached will correspond to solution C, provided the final value of T_0 is smaller than T_C. If, on the other hand, the rod has initially a temperature distribution such that it is in contact with the hot wall, and the temperature T_0 is not suddenly dropped below a level to break contact, the steady state A will establish itself for long time values of $T_0 > T_A$. The unstable steady state B could conceivably be reached by carefully steering the process during its early stages. However, any temperature disturbance that corresponds to thermal elongation of the rod will eventually make the system settle down in state A. Conversely, disturbances that make the rod contract slightly will make it go into state C.

More generally, we conclude that, if the contact resistance is a continuous function of g, there will be an odd number of steady-state solutions which are alternately stable and unstable. The stable solutions might be thought of as separated from each other by "higher energy" unstable barriers. In the particular case of the step change in resistance shown in Fig. 3, the imperfect contact solution acts as such a barrier and is unstable.

Energy Considerations

If we define an "energy function"

$$U(g) = \frac{AE}{l}\left[\frac{1}{2}(g_0 - g)^2 + \int u_0 f dg\right] \tag{27}$$

with E denoting Young's modulus, the condition (8) for a steady-state solution can be expressed as

$$\partial U/\partial g = 0 \tag{28}$$

while the condition for instability (26) is

$$\partial^2 U/\partial g^2 < 0 \tag{29}$$

In other words, the function U is stationary at all steady-state solutions, being a maximum if the solution is unstable and a minimum if

it is stable. Thus it behaves in all respects as the total energy of a conservative mechanical system.

Furthermore, we can give a physical interpretation to the two terms in equation (27). The first term, $\frac{1}{2}(g_0 - g)^2 AE/l$, is the elastic strain energy involved in extending or compressing the rod isothermally at temperature zero, while the second term can be expressed as

$$- \int u_0 f dP = - \int u_{th} dP \tag{30}$$

(see equations (4) and (7)) where the compressive force P has a continuation $(-gAE/l)$ into negative values.

As long as conditions are changed slowly enough for the temperature distribution in the rod to be in a quasi-steady state, the rod with pressure/gap dependent contact resistance will exhibit a unique relation between load (gap) and extension. Now, if a mechanical system could be constructed with the *same* load extension relation, the normal energy theorems could be applied to it, since we should not now have a continuous flow of energy across a boundary containing the system. However, such a mechanical system would only be conservative if the load was always varied incrementally, i.e., the sudden application or removal of a finite load may lead to the system doing extra work on the surroundings. Of course, this quasi-static, incremental behavior cannot be guaranteed in the thermal system, but the energy function obtained in the foregoing from perturbation arguments is closely related to that which would be obtained by imposing the requirement of minimum complementary energy on such a system.

The authors have as yet been unable to justify use of such an energy argument other than through the perturbation analysis set out in the section, "Stability Analysis," but the result summarized in equations (27)–(29) is extremely suggestive. If such a justification could be produced, it would be capable of extension to more difficult thermoelastic contact problems involving nonuniqueness, such as those concerned with the half space [4], for which a perturbation analysis would be of formidable complexity. The reader's attention is drawn to this unsolved problem.

References

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

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

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

4 Comninou, M., and Dundurs, J. "On Lack of Uniqueness in Heat Conduction Through a Solid to Solid Interface," ASME *Journal of Heat Transfer*, Vol. 102, 1980, pp. 319–323.

APPENDIX

Criterion for Complex Roots of Equation (22)

We are concerned to find the conditions under which the function

$$F(z) = z^2 \cosh z + \alpha z \sinh z + \beta(\cosh z - 1) \tag{31}$$

has complex roots corresponding to values of z^2 with positive real part. In the z-plane, the corresponding zeros of $F(z)$ must lie in the two sectors shaded in Fig. 4 and bounded by the lines $z = \pm(1 \pm i)\omega$. The origin is also excluded by two small quarter circles, as a zero there is introduced by the multiplication by y in the derivation of equation (22) and is of no physical significance.

$F(z)$ is a continuous function of the coefficients $\alpha = f/(1-f)$ and $\beta = 2u_0 f'/(1-f)$ and has no zeros in the domains of Fig. 4 if $\alpha = \beta = 0$. Hence, if we start from this condition and change α, β continuously, the zeros will move continuously about the complex plane and will only be able to enter the domain by crossing one of its boundaries. The stability boundary is therefore equivalent to the condition that $F(z)$

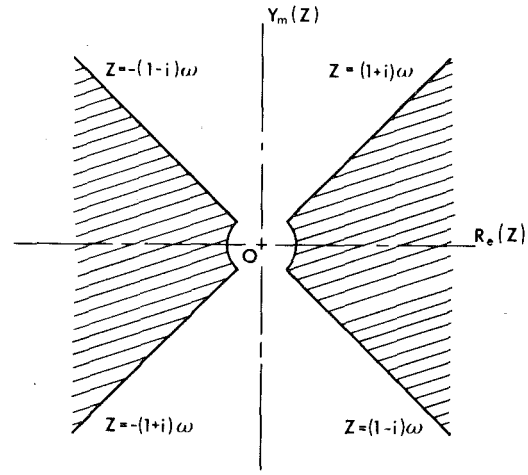


Fig. 4 Regions in the complex plane

has a zero on the domain boundary. We consider these boundaries in turn:

(*i*) $z = (1+i)\omega$. If we treat (31) as an equation for α, we have

$$\alpha = -\frac{z^2 \cosh z + \beta(\cosh z - 1)}{z \sinh z} \tag{32}$$

which can be decomposed in the form

$$\alpha = \frac{2\omega^2 sh \cdot s - \beta(ch \cdot c - 1) - i(2\omega^2 ch \cdot c + \beta sh \cdot s)}{\omega[sh \cdot c - s \cdot ch + i(sh \cdot c + ch \cdot s)]} \tag{33}$$

where $s = \sin\omega$, $c = \cos\omega$, $sh = \sinh\omega$, $ch = \cosh\omega$.

Now α is a *real* constant and hence the imaginary part of equation (33) is zero, i.e.,

$$[2\omega^2 sh \cdot s - \beta(ch \cdot c - 1)](sh \cdot c + ch \cdot s) - [2\omega^2 ch \cdot c + \beta sh \cdot s](sh \cdot c - ch \cdot s) = 0 \tag{34}$$

from which, after some manipulation we have

$$\beta = \frac{2\omega^2(ch \cdot sh - c \cdot s)}{(sh - s)(ch - c)} \tag{35}$$

Back substitution into equation (33) then gives

$$\alpha = -\frac{\omega(ch \cdot sh^2 - c \cdot s^2)}{2(sh^2 \cdot c^2 + ch^2 \cdot s^2)(sh - s)} \tag{36}$$

which is negative for all ω, whereas $\alpha = f/(1-f)$ must be positive for all possible resistance functions f.

We conclude that no zero can enter the domain across $z = (1+i)\omega$. A similar argument can be used to prove that zeros cannot enter across the other diagonal boundaries.

(*ii*) $z = re^{i\theta}$, r *small*, $-\pi/4 < \theta < \pi/4$. For this boundary, $F(z)$ can be expanded as in equation (24) and condition (25) is immediately obtained on dropping all except the lowest order terms in r.

(*iii*) *The Point at Infinity.* As $z \to \infty$, the function $F(z)$ becomes dominated by the term $z^2 \cosh z$ which is insensitive to variation of α, β. Thus no zeros can enter the domain through the point at infinity.

We therefore conclude that condition (25) is the correct stability criterion for real or complex roots. In fact, a further analysis along the same lines, but excluding the real axis by the two lines $z = \pm i\delta$, shows that zeros of $F(z)$ in the domain of Fig. 4 can only occur on the real axis. In other words, there are no oscillatory perturbation solutions to the problem.

Carl Panek

Ebasco Services, Inc.,
2 Rector Street,
New York, N. Y. 10006

A Thermomechanical Example of Auto-Oscillation

A system which progresses repeatedly through a cyclic sequence of physical configurations (continuous or discrete), in the absence of any periodic excitation, is said to be in the state of auto-oscillation. In this paper the author examines a thin thermoelastic rod projecting from the warmer of two parallel rigid walls, their separation being slightly wider than the length of the rod. Assuming heat transfer by conduction only, it is found that for certain combinations of physical parameters the rod does auto-oscillate (repeatedly makes and breaks contact with the cooler wall at finite frequency) provided that part of the rod near the cooler wall has a nonpositive thermal coefficient of expansion. This last condition is absolutely necessary since otherwise one encounters certain thermomechanical pathologies. The only other way that these pathologies could be eliminated would be to employ the concept of imperfect contact as suggested by Barber, which may also be interpreted as allowing heat transfer by radiation.

Introduction

Auto-oscillation describes the ability of a system to progress from one configuration through a number of subsequent configurations, eventually duplicating the initial configuration. Fundamental to the concept of auto-oscillation is the requirement that the system in question present its cyclic behavior in the total absence of periodic excitation. Briefly stated, auto-oscillation is the existence of periodic behavior in the absence of periodic stimulus. Some examples of auto-oscillation are as follows:

1 A pendulum with escapement mechanism.
2 Certain chemical reactions [1].
3 Oscillating configurations in cellular automata theory [2].

The current interest in thermoelastic contact, and Barber's concept of heat conduction through zones of imperfect contact [3], have led the author to examine one simple thermomechanical system for auto-oscillation. While such a study does not yield results which are directly applicable to specific questions, it does cause ont to think carefully about the physical meaning of thermomechanical boundary conditions; and in any case it is difficult to ignore a system which presents the possibility of exhibiting auto-oscillation.

Description of the Problem

Consider a thin uniform rod of initial length π, projecting from a warm rigid wall, as shown in Fig. 1. At a distance $\pi + g$ from the warm wall let there be a cooler wall. Without loss of generality it may be assumed that the warm wall is maintained at temperature $T = 1$, and that the cooler wall is maintained at temperature $T = 0$. Materially the rod is divided into two collinear parts having thermal coefficient of expansion α_1 and α_2, thermal conductivity k_1 and k_2, and thermal diffusivity D_1 and D_2, respectively. Finally, it is assumed that heat may flow by means of conduction only, and only where there is intimate contact between two surfaces.

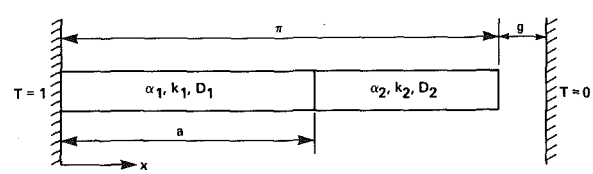


Fig. 1 Geometry of the problem

If g is sufficiently small or sufficiently large one steady state or another develops, but for a certain range of g between these two extremes it is impossible to develop a steady-state solution consistent with the assumed mechanism of heat transfer. On the other hand, any oscillatory solution has an intrinsic pathology since the heated rod cannot remain in contact with the cooler wall for any finite period of time. This behavior is due to the fact that the rate of contraction of the rod is proportional to the heat flux out of the rod times the thermal coefficient of expansion of the rod, and at the instant of contact with the wall the net heat flux out of the rod is infinite.

One solution to this dilemma is to introduce the concept of imperfect contact. In this case the rod expands by an amount g, develops a zone of imperfect contact at the wall, and then remains in imperfect contact with the wall as all transient effects die out. Alternatively, it is possible to find an *oscillatory* solution to this problem consistent with both mechanisms of heat transfer *provided that the part of the*

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, May, 1980. Paper No. 80-WA/APM-21.

rod near the cooler wall has a nonpositive thermal coefficient of expansion.

Thermal Requirements for Auto-Oscillation

Regardless of the feasibility of satisfying the combined thermomechanical boundary conditions, one can solve for a periodic temperature field consistent with a regularly alternating thermal boundary condition at the cooler end of the rod. After such a solution has been found, it will be examined to determine whether or not it satisfies the mechanical boundary conditions necessary for auto-oscillation.

To find a periodic temperature field, first express the temperature in the rod as two different infinite series, one for each part of the cycle, whose individual terms each satisfy the one-dimensional heat-conduction equation

$$D\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} \tag{1}$$

Thus one has

$$T(x, t) = 1 - \frac{x}{\pi} + \sum_{n=1}^{\infty} C_n \exp(-n^2 Dt) \sin nx, \quad 0 < t < \tau \tag{2}$$

and

$$T^*(x, t^*) = 1 + \sum_{n=1}^{\infty} C_n^* \exp\{-(n - \tfrac{1}{2})^2 Dt^*\} \sin (n - \tfrac{1}{2})x, \quad 0 < t^* < \tau^* \tag{3}$$

where T denotes the temperature field when the rod is in contact with the wall, and T^* denotes the temperature when the rod is out of contact with the wall. τ and τ^* denote the length of time the rod is in or out of contact with the wall, respectively. Note that the form of each series was chosen so as to satisfy the boundary conditions at $x = \pi$,

$$T(\pi, t) = 0, \quad 0 < t < \infty \tag{4}$$

and

$$\frac{\partial}{\partial x} T^*(x, t^*) \big|_{x=\pi} = 0, \quad 0 < t^* < \infty \tag{5}$$

as well as to yield the correct steady-state solutions as t or t^* goes to infinity.

In the interest of mathematical simplicity, it has been assumed that $D_1 = D_2 = D$ and $k_1 = k_2 = k$. Doing this results in the continuity of both $T(x, t)$ and $\partial T(x, t)/\partial x$ at the junction of the two rods. Actually the second condition is not necessary since it is always possible to transpose the length a to an equivalent length a' according to $a' = k_2 a\pi/[k_1\pi + (k_2 - k_1)a]$.

By requiring continuity of the two temperature fields at the instant the rod breaks contact with the wall, one finds that

$$T(x, \tau) = T^*(x, 0), \quad 0 < x < \pi \tag{6}$$

Because the behavior is expected to be cyclic, one must also require continuity of the two temperature fields at the instant the rod again makes contact with the wall. Thus one has

$$T^*(x, \tau^*) = T(x, 0), \quad 0 < x < \pi \tag{7}$$

It is proposed to satisfy the two continuity equations in a least-squares sense. Proceeding, one defines an error integral $I(\{C\}, \{C^*\})$ according to

$$I(\{C\}, \{C^*\}) = \int_0^{\pi} \{[T(x, \tau) - T^*(x, 0)]^2 + [T^*(x, \tau^*) - T(x, 0)]^2\} dx \tag{8}$$

Differentiating $I(\{C\}, \{C^*\})$ with respect to the individual coefficients C_i and C_i^* in turn, and carrying out all the resulting integrations, one finds two infinite sets of simultaneous equations:

$$-\frac{2}{\pi}[1 + \exp(-m^2 D\tau)] = (-1)^m m[1 + \exp(-2m^2 D\tau)] C_m \tag{9}$$

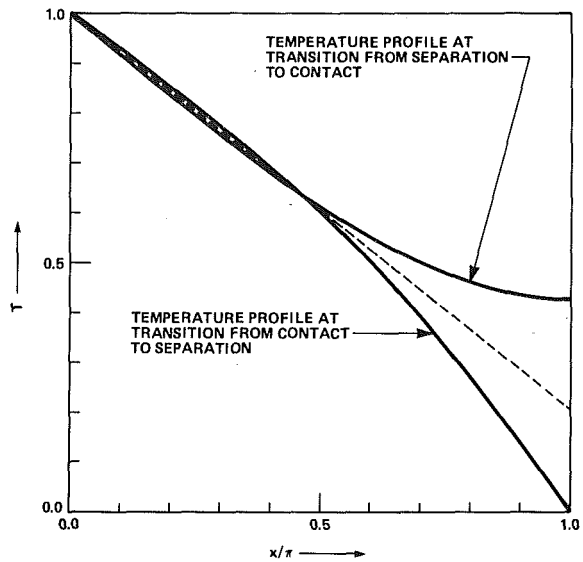


Fig. 2 Temperature profiles at the two transition times $D\tau = 0.5$ and $D\tau^* = 1.0$

$$-\frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^n \left(\frac{m^2}{m^2 - (n - \frac{1}{2})^2}\right)[\exp(-m^2 D\tau) + \exp\{-(n - \tfrac{1}{2})^2 D\tau^*\}] C_n^*, \quad m = 1, 2, \ldots, \infty \tag{9}$$

(Cont.)

and

$$\frac{2}{\pi^2}[1 + \exp\{-(m - \tfrac{1}{2})^2 D\tau^*\}] = (-1)^m (m - \tfrac{1}{2})[1 + \exp\{-2(m - \tfrac{1}{2})^2 D\tau^*\}] C_m^* - \frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^n n\left(\frac{(m - \frac{1}{2})^2}{n^2 - (m - \frac{1}{2})^2}\right)[\exp\{-(m - \tfrac{1}{2})^2 D\tau^*\} + \exp(-n^2 D\tau)] C_n, \quad m = 1, 2, \ldots, \infty \tag{10}$$

For the solution of the periodic temperature field, the foregoing system would be truncated to some finite number of equations, say $2N$. Given times for the desired duration of each part of the cycle, $D\tau$ and $D\tau^*$, one would solve the truncated system for the $2N$ coefficients $C_1, C_2, \ldots, C_{2N}, C_1^*, C_2^*, \ldots, C_{2N}^*$. Results are plotted in Figs. 2 and 3 for the single case $D\tau = 0.5, D\tau^* = 1.0$. Note that the temperature profiles undulate, and near the center of the rod one sees a periodic flow and ebb of heat superimposed on the overall temperature gradient.

Mechanical Requirements for Auto-Oscillation

Given the thermal coefficients of expansion for the two parts of the rod, one can compute the extension of the rod as a function of time simply by integrating $\alpha T(x, t)$ over the length of the rod. (Note that if the length a had been transposed to a' in order to eliminate the slope discontinuity in the temperature field, it would be necessary to transpose back to the original length a before integrating over x.) The extension of the rod during each of the two parts of the cycle may be written, respectively, as

$$\epsilon(t) = \alpha_1 \int_0^a T(x, t)dx + \alpha_2 \int_a^{\pi} T(x, t)dx, \quad 0 < t < \tau \tag{11}$$

and

$$\epsilon^*(t^*) = \alpha_1 \int_0^a T^*(x, t^*)dx + \alpha_2 \int_a^{\pi} T^*(x, t^*)dx, \quad 0 < t^* < \tau^* \tag{12}$$

In order to satisfy the mechanical condition necessary for auto-os-

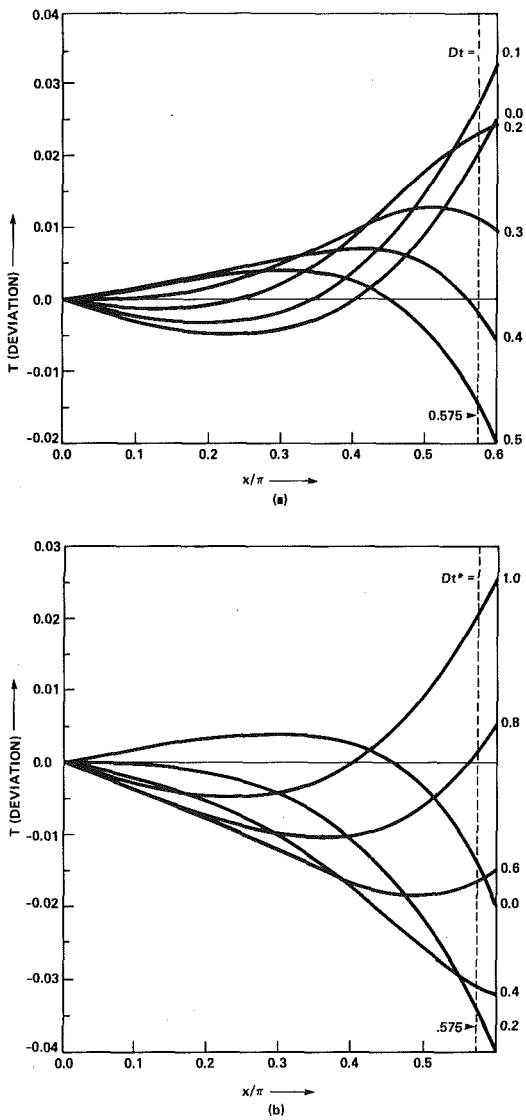


Fig. 3 Detail showing deviation of the temperature from the line $1-0.8x/\pi$ as Dt goes from 0 to 0.5 and as Dt^* goes from 0 to 1.0

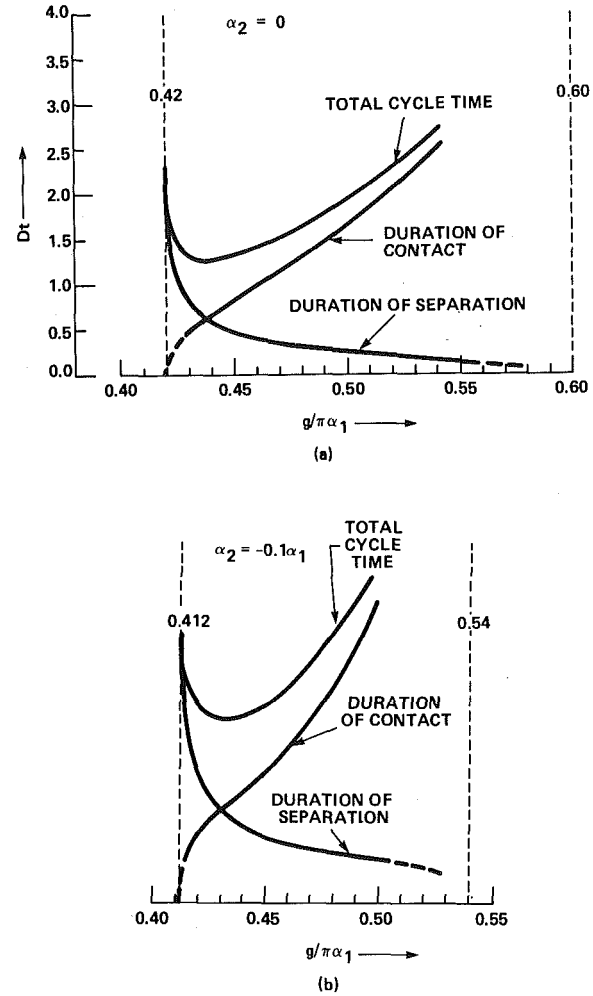


Fig. 4 Period of auto-oscillation plotted against initial gap for $a = 0.6\pi$

cillation one must not only close the gap at the beginning and end of each part of the cycle,

$$\epsilon(0) = \epsilon^*(0) = g \quad \text{or} \quad \epsilon(\tau) = \epsilon^*(\tau^*) = g \tag{13}$$

but one must also satisfy the inequalities

$$\epsilon(t) > g, \quad 0 < t < \tau \tag{14}$$

and

$$\epsilon^*(t^*) < g, \quad 0 < t^* < \tau^* \tag{15}$$

which require that the gap not open/close during that part of the cycle when the gap is to remain closed/open. Throughout the remaining analysis it is assumed that this set of inequalities is satisfied as long as α_2 is nonpositive and equation (13) is satisfied. A rigorous proof of this statement will not be presented, but its validity will be demonstrated in the next section by considering a sample problem.

When applying equation (13) it is sufficient to use only one of the two versions persented. In consideration of improved numerical convergence it is advantageous to use the right-hand version. Thus, in conjunction with equations (11) and (12), one finds

$$\alpha_1\left(a - \tfrac{1}{2}a^2/\pi + \sum_{n=1}^{\infty} n^{-1} C_n \exp(-n^2 D\tau)\,[1 - \cos na]\right)$$
$$+ \alpha_2\left((\pi - a) - \tfrac{1}{2}(\pi^2 - a^2)/\pi + \sum_{n=1}^{\infty} n^{-1} C_n \exp(-n^2 D\tau)\right.$$
$$\left.\times\,[\cos na - (-1)^n]\right) = \alpha_1\left(a + \sum_{n=1}^{\infty} (n - \tfrac{1}{2})^{-1} C_n{}^*\right.$$
$$\left.\times \exp\{-(n - \tfrac{1}{2})^2 D\tau^*\}\,[1 - \cos(n - \tfrac{1}{2})a]\right) + \alpha_2\left((\pi - a)\right.$$
$$\left.+ \sum_{n=1}^{\infty} (n - \tfrac{1}{2})^{-1} C_n{}^* \exp\{-(n - \tfrac{1}{2})^2 D\tau^*\}\,[\cos(n - \tfrac{1}{2})a]\right) = g \tag{16}$$

The most practical means of solving the foregoing equation is to proceed in three stages. First, choose values for $D\tau$ and $D\tau^*$, and solve equations (9) and (10) for the Fourier coefficients C_n and $C_n{}^*$. Then choose a value for a and determine the ratio α_2/α_1 from the first equality of equation (16). Once α_2/α_1 has been found it is possible to determine g/α_1 from the second equality of equation (16). In this way it is possible to develop a plot of oscillation period versus initial gap for a rod of given configuration. Two examples of such plots are given in Fig. 4.

Satisfying the Gap Inequalities

As previously stated, it has been assumed, but not proven, that

requiring $\alpha_2 \leq 0$ and satisfying equation (13) is sufficient to guarantee that the inequalities of equations (14) and (15) are also satisfied. In order to demonstrate the validity of that assumption, consider the temperature profiles shown in Fig. 3 where $D\tau = 0.5$ and $D\tau^* = 1.0$. By iterating on equation (16), it is found that the time $D\tau = 0.5$ and $D\tau^* = 1.0$ result in auto-oscillation when $\alpha_2 = 0$ and $a = 0.575\pi$. Since $\alpha_2 = 0$ in this example, the elongation of the rod is equal to α_1 times the integral of the temperature (the area under the curve) from $x = 0$ to $x = 0.575\pi$. Referring to Fig. 3, one sees that the rod does continue to expand for some time after making contact with the wall, and continues to contract for some time after breaking contact.

Another way to verify this behavior is to consider the time rate of elongation of the rod. Differentiating equations (11) and (12) with respect to t, one finds

$$\frac{d}{dt}\epsilon(t) = \alpha_1 \int_0^a \frac{\partial}{\partial t} T(x, t)dx + \alpha_2 \int_0^\pi \frac{\partial}{\partial t} T(x, t)\, dx \qquad (17)$$

and

$$\frac{d}{dt}\epsilon^*(t^*) = \alpha_1 \int_0^a \frac{\partial}{\partial t} T^*(x, t^*)dx + \alpha_2 \int_a^\pi \frac{\partial}{\partial t} T^*(x, t^*)\, dt \qquad (18)$$

Using the heat conduction equation, $D\partial T/\partial t = \partial^2 T/\partial x^2$, one finds that

$$D\frac{d}{dt}\epsilon(t) = \alpha_1 \frac{\partial T}{\partial x}\bigg|_0^a + \alpha_2 \frac{\partial T}{\partial x}\bigg|_a^\pi \qquad (19)$$

and

$$D\frac{d}{dt}\epsilon^*(t^*) = \alpha_1 \frac{\partial T^*}{\partial x}\bigg|_0^a + \alpha_2 \frac{\partial T}{\partial x}\bigg|_a^\pi \qquad (20)$$

Thus, in the present example, since $\alpha_2 = 0$, one can determine whether or not the rod is expanding by comparing the slopes of the temperature profiles at the points $x = 0$ and $x = 0.575\pi$. To insure that the rod is expanding when contact is made, and that it is contracting when contact is broken, one must have

$$\frac{\partial T(x, t)}{\partial x}\bigg|_{x=0.575\pi} > \frac{\partial T(x, t)}{\partial x}\bigg|_{x=0}, \quad Dt \simeq 0 \qquad (21)$$

and

$$\frac{\partial T^*(x, t^*)}{\partial x}\bigg|_{x=0.575\pi} < \frac{\partial T^*(x, t^*)}{\partial x}\bigg|_{x=0}, \quad Dt^* \simeq 0 \qquad (22)$$

Indeed one sees from Fig. 3 that these conditions are met. Moreover, since the temperature profiles change so smoothly in time, one can convincingly argue that the extension of the rod changes in a simple cyclic manner. Finally, since the particular example apparently contains no anomalies, it is reasonable to generalize the conclusions obtained here to all other cases.

Conclusions

In the aforementioned work the author examines a simple thermomechanical system for the existence of auto-oscillation; the system being a thin thermoelastic rod projecting from the warmer of two parallel rigid walls, their separation being slightly wider than the length of the rod. It is found that given a physical configuration which excludes the two limit cases (intimate contact with the cooler wall and absolute noncontact with the cooler wall) either an oscillatory solution or a steady-state solution may develop.

The actual solution which does develop depends solely on α_2, the thermal coefficient of expansion of the part of the rod near the cooler wall. If $\alpha_2 > 0$, one must employ Barber's condition of imperfect contact. In this case the rod expands by an amount to just close the gap, develops a condition of imperfect contact, and remains in contact with the wall. Notice that this is just the behavior expected if one were to allow radiation effects, thus suggesting an equivalence between the two processes. On the other hand if $\alpha_2 \leq 0$, one finds that imperfect contact does not develop. In fact if $\alpha_2 < 0$ one sees that as the end of the rod just touches the cooler wall, the rod actually jumps into intimate contact with the wall. This behavior is due to the fact that for $\alpha_2 < 0$ the extension of the rod is proportional to the heat flow out of the rod, and as contact develops the instantaneous heat flow out of the rod is infinite.

The combined result of these considerations are that for thermoelastic contact between two bodies

1 Imperfect contact *must* develop during the transition from intimate contact to separation if heat flows out of the body with the *higher* distortivity (distortivity is proportional to α; see Dundurs and Panek [4]).

2 Imperfect contact *cannot* develop during the transition from intimate contact to separation if heat flows out of the body with the *lower* distortivity.

These results refer to a transition in time, but the validity of a similar set of results applicable to transition on a boundary surface has been recently proven by Comninou and Dundurs [5].

Acknowledgments

The author wishes to thank Prof. J. Dundurs and Prof. G. Djukic for their helpful comments. The author also wishes to acknowledge that part of this work was completed while he was a member of the faculty at the University of Cincinnati.

References

1 Kochina, N. N., "A Periodic Solution of the Diffusion Equation," *Journal of Applied Mechanics and Technical Physics*, Vol. 8, No. 1, 1964, pp. 51–56.

2 Gardner, M., "Mathematical Games," *Scientific American*, Vol. 224, No. 2, 1971, pp. 112–117.

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

4 Dundurs, J. and Panek, C., "Heat Conduction Between Bodies With Wavy Surfaces," *International Journal of Heat and Mass Transfer*, Vol. 19, 1976, pp. 731–736.

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

H. Saito
Professor,
Department of Mechanical Engineering,
Tohoku University,
Sendai, Japan

T. Terasawa
Research Engineer,
Hitachi Central Research Laboratory,
Hitachi, Ltd.,
Kokubunji, Japan

Steady-State Vibrations of a Beam on a Pasternak Foundation for Moving Loads

The response of an infinite beam supported by a Pasternak-type foundation and subjected to a moving load is investigated. It is assumed that the load is uniformly distributed over the finite length on a beam and moves with constant velocity. The equations of motion based on the two-dimensional elastic theory are applied to a beam. Steady-state solutions are determined by applying the exponential Fourier transform with respect to the coordinate system attached to the moving load. The results are compared with those obtained from the Timoshenko and the Bernoulli-Euler beam theories, and the differences between the displacement and stress curves obtained from the three theories are clarified.

Introduction

The dynamic response of an infinite beam on an elastic foundation under moving loads is of great interest in many fields of engineering, and many papers have been published in the field of such problems. In these studies, the simplest model considered was a Bernoulli-Euler beam model supported by a Winkler-type foundation. Some of the previous studies of this kind have been the works of Dörr [1], Kenney [2], Mathews [3, 4], and Keer [5]. Holder and Michalopoulos [6] introduced an inertial foundation instead of massless Winkler-type foundation that reacts only on the local deflection of the mounting body and investigated the effects on the system response of foundation mass. It is well known, however, that the Bernoulli-Euler beam theory is inadequate for increasing frequencies which result from higher load velocities, and a significant improvement has been obtained by considering the Timoshekno beam theory. Crandall [7] replaces the Bernoulli-Euler beam by the Timoshenko model and obtained the steady-state solutions for the beam on an elastic foundation with a moving concentrated load. The problem of a semi-infinite Timoshenko beam of an elastic foundation with a load moving from the end at a constant velocity was discussed by Steele [8, 9]. Recently Chonan [10, 11] analyzed the response of an elastically supported Timoshenko beam subjected to an axial force and a moving load, and a moving harmonic force. Thus, in the analysis of response of a beam subjected to a moving load, one-dimensional equations of motion based on the Bernoulli-Euler or the Timoshenko beam theories have been widely used. Though such approximations are very useful when one wishes to handle the complex problem of the response of a beam to dynamic loading, it will be necessary to confirm how accurate the results can express the behavior of the beam as compared with the exact ones obtained from the two-dimensional linear elasticity theory. On the other hand, shear interactions between the foundation spring elements are neglected in the Winkler foundation. This is not the case with the real foundation.

The present paper is concerned with an analysis of the response of an infinite beam on an elastic foundation under moving load with constant velocity. The equations of motion based on the two-dimensional elastic theory are applied to a beam and the Pasternak-type foundation is introduced considering shear interactions as well as reaction proportional to displacement. For simplicity a mass of foundation is neglected. If the moving load is assumed to be a concentrated force acting on a straight boundary, the stress components become infinite at the point of application of the load and a further investigation of the problem is necessary. In practice the load will distribute over a finite area. To consider the local effect near the load, it is assumed that the load is distributed over the narrow finite length on a beam. The solutions are obtained by using Fourier transforms and the numerical results are compared with those obtained from the Bernoulli-Euler beam theory and the Timoshenko beam theory. The influence of the higher modes of vibration due to two-dimensional effects is conspicuous for load speeds higher than the velocity of Rayleigh waves.

Analysis

Fig. 1 shows an infinite beam with thickness h which is on the Pasternak-type foundation and subjected to a moving load. The x-axis is spaced along the interface between the beam and the foundation, and the y-axis normal to the x-axis. The applied load which is distributed in the region 2ϵ with intensity f is assumed to move with constant velocity c in the x-direction.

Putting the beam displacements in the x,y-directions as

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, May, 1979; final revision, April, 1980.

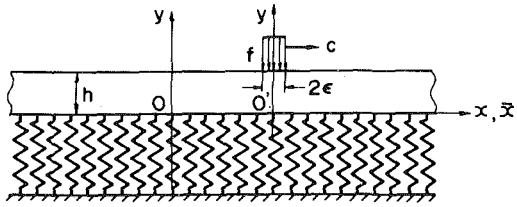


Fig. 1 Geometry of problem and coordinate axes

$$u = \frac{\partial \phi}{\partial x} + \frac{\partial \psi}{\partial y}, \quad v = \frac{\partial \phi}{\partial y} - \frac{\partial \psi}{\partial x} \tag{1}$$

the plane-stress equations of elastodynamics are described as

$$\frac{2}{1-\nu}\nabla^2\phi = \frac{\rho}{G}\frac{\partial^2\phi}{\partial t^2}, \quad \nabla^2\psi = \frac{\rho}{G}\frac{\partial^2\psi}{\partial t^2} \tag{2}$$

and the stress components are, in usual notation,

$$\sigma_x = \frac{2G}{1-\nu}\left(\frac{\partial u}{\partial x} + \nu\frac{\partial v}{\partial y}\right), \quad \sigma_y = \frac{2G}{1-\nu}\left(\frac{\partial v}{\partial y} + \nu\frac{\partial u}{\partial x}\right)$$

$$\tau_{xy} = G\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) \tag{3}$$

where G is the modulus of rigidity, ν is Poisson's ratio, ρ is mass density, t is time, respectively, and ∇^2 is

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

For the Pasternak foundation with viscous damping, the foundation reaction Q is given as [12]

$$Q = kv + d\frac{\partial v}{\partial t} - G_0\frac{\partial^2 v}{\partial x^2}$$

where k is the Winkler foundation modulus, G_0 is the shear foundation modulus, and d is viscous damping coefficient. Then the boundary conditions are

At $y = 0$

$$\tau_{xy} = 0, \quad \sigma_y = kv + d\frac{\partial v}{\partial t} - G_0\frac{\partial^2 v}{\partial x^2}$$

At $y = h$

$$\tau_{xy} = 0, \quad \sigma_y = F(x - ct) = \begin{cases} -f, & |x - ct| \leqq \epsilon \\ 0, & |x - ct| > \epsilon \end{cases} \tag{4}$$

Now it is assumed that the steady state has been reached and deformation pattern is time invariant relative to a following coordinate system attached to the moving load

$$\bar{x} = x - ct$$

Transforming the equations (1)–(4) to the moving coordinate and applying the exponential Fourier transform with respect to $\bar{x}$ defined as

$$\hat{F}(\xi,y) = \int_{-\infty}^{\infty} F(\bar{x},y)e^{-i\xi\bar{x}}\,d\bar{x}, \quad F(\bar{x},y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{F}(\xi,y)e^{i\xi\bar{x}}d\xi$$

give

$$\left(\frac{d^2}{dy^2} - p^2\xi^2\right)\hat{\phi} = 0, \quad \left(\frac{d^2}{dy^2} - q^2\xi^2\right)\hat{\psi} = 0 \tag{5}$$

$$\hat{u} = i\xi\hat{\phi} + \frac{d\hat{\psi}}{dy}, \quad \hat{v} = \frac{d\hat{\phi}}{dy} - i\xi\hat{\psi} \tag{6}$$

$$\hat{\sigma}_x = \frac{2G}{1-\nu}\left\{\left(-\xi^2\hat{\phi} + i\xi\frac{d\hat{\psi}}{dy}\right) + \nu\left(\frac{d^2\hat{\phi}}{dy^2} - i\xi\frac{d\hat{\psi}}{dy}\right)\right\}$$

$$\hat{\sigma}_y = \frac{2G}{1-\nu}\left\{\left(\frac{d^2\hat{\phi}}{dy^2} - i\xi\frac{d\hat{\psi}}{dy}\right) + \nu\left(-\xi^2\hat{\phi} + i\xi\frac{d\hat{\psi}}{dy}\right)\right\}$$

$$\hat{\tau}_{xy} = G\left(2i\xi\frac{d\hat{\phi}}{dy} + \xi^2\hat{\psi} + \frac{d^2\hat{\psi}}{dy^2}\right) \tag{7}$$

and

At $y = 0$

$$\hat{\tau}_{xy} = 0, \quad \hat{\sigma}_y = (k + i\xi cd + G_0\xi^2)\hat{v}$$

At $y = h$

$$\hat{\tau}_{xy} = 0, \quad \hat{\sigma}_y = f(e^{-i\epsilon\xi} - e^{i\epsilon\xi})/i\xi \tag{8}$$

where

$$p^2 = 1 - \left(\frac{c_2}{c_1}\beta\right)^2, \quad q^2 = 1 - \beta^2, \quad \beta = \frac{c}{c_2}$$

$$c_1{}^2 = \frac{2G}{\rho(1-\nu)}, \quad c_2{}^2 = \frac{G}{\rho}, \quad i = \sqrt{-1} \tag{9}$$

and β is the nondimensional load velocity.

The solutions of equations (5) can be obtained as

$$\hat{\phi} = A\cosh p\xi y + B\sinh p\xi y, \quad \hat{\psi} = C\cosh q\xi y + D\sinh q\xi y \tag{10}$$

where A–D are undetermined constants. Substituting equations (10) into equations (6) and (7) and utilizing equations (8) give

$$\begin{bmatrix} 0 & 2ip & 1+q^2 & 0 \\ (p^2-\nu)\eta & -(K + L\beta\eta i + H\eta^2)p & (K + L\beta\eta i + H\eta^2)i & (-1+\nu)iq\eta \\ 2ip\sinh p\eta & 2ip\cosh p\eta & (1+q^2)\cosh q\eta & (1+q^2)\sinh q\eta \\ (p^2-\nu)\cosh p\eta & (p^2-\nu)\sinh p\eta & (-1+\nu)iq\sinh q\eta & (-1+\nu)iq\cosh q\eta \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \Theta(\eta) \end{bmatrix} \tag{11}$$

where

$$\eta = h\xi, \quad K = \frac{(1-\nu)hk}{2G}, \quad L = \frac{(1-\nu)c_2 d}{2G}, \quad H = \frac{(1-\nu)G_0}{2Gh}$$

$$\Theta(\eta) = \frac{(1-\nu)fh^3}{2iG\eta^3}(e^{-i\eta\epsilon/h} - e^{i\eta\epsilon/h}) \tag{12}$$

Substituting $\hat{\phi}$ and $\hat{\psi}$, in which constants A–D are determined from equation (11), into equations (6) and (7) and using the inverse transform, one can get the solutions for the displacements and the stress components of a beam. If one puts $\epsilon \to 0$ and $2\epsilon f = P$ in equations (11) and (12), the case of a beam subjected to a concentrated force P moving with constant velocity can be obtained.

The integrals are transformed to the sum of the residues of the poles defined by the roots of the characteristic equation

$$\begin{vmatrix} 0 & 2ip & 1+q^2 & 0 \\ (p^2-\nu)\eta & -(K + iL\beta\eta + H\eta^2)p & (K + iL\beta\eta + H\eta^2)i & (-1+\nu)iq\eta \\ 2ip\sinh p\eta & 2ip\cosh p\eta & (1+q^2)\cosh q\eta & (1+q^2)\sinh q\eta \\ (p^2-\nu)\cosh p\eta & (p^2-\nu)\sinh p\eta & (-1+\nu)iq\sinh q\eta & (-1+\nu)iq\cosh q\eta \end{vmatrix} = 0 \tag{13}$$

which results from equating the denominators of the integrands to zero.

It is noted that equation (13) corresponds to the dispersion equation which is obtained in the study of the free wave propagation in an in-

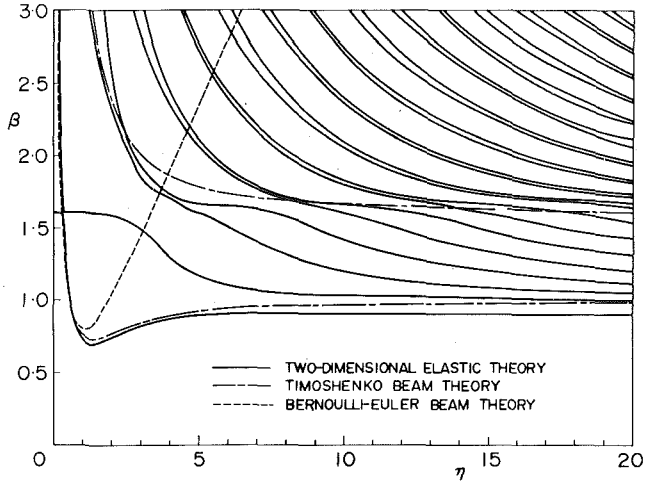


Fig. 2 Nondimensional velocity of the moving load β versus nondimensional integral variable η ($K = 0.1$, $H = 0.05$)

finite beam on an elastic stratum by regarding β and η as the nondimensional phase velocity and wave number of free waves.

Numerical Results

Numerical calculations have been carried out for the case $2\epsilon/h = 0.4$, $\nu = 0.3$, and arbitrarily chosen values of K and H. To simplify the calculations, we will consider only the limiting case $d = 0$. For $d = 0$, the equation (13) has real roots besides complex roots. In this case, however, if infinitesimal amount of damping is introduced, it is possible to ascertain whether the real poles should belong to the upper half plane or the lower half plane at the time of accomplishing the inverse transform.

The variations of the nondimensional velocity of the moving load β versus the nondimensional integral variable η in the case of $K = 0.1$ and $H = 0.05$ are shown by the solid lines in Fig. 2. A broken line and chained lines in the figure are the results from the solutions of the Bernoulli-Euler beam theory and the Timoshenko beam theory, respectively, which are shown in the Appendix. The graph is symmetrical with respect to $\eta = 0$, and only the curves for real and positive values of η are shown. The value of the shear coefficient κ for Timoshenko beams is taken to be 0.85 which is given for a rectangular section by Cowper [13]. It will be found from the figure that infinite numbers of curves exist if the two-dimensional elastic theory is applied to a beam.

When the integrand has poles of the second order on the real axis, the integral values blow up and the displacements along the beam become unbounded. The velocity in this case is associated with critical velocity, the values of which are given by those of the load velocities at the minimum points of the curves in Fig. 2, that is, $\beta_{cr} = 0.704$ for the two-dimensional elastic theory, $\beta_{cr} = 0.727$ for the Timoshenko beam theory and $\beta_{cr} = 0.805$ for the Bernoulli-Euler beam theory.

Figs. 3 (*a*) and (*b*) show the curves of the first mode for various values of K or H taking $H = 0$ or $K = 0.1$. In the figures, solid lines are results obtained from the two-dimensional elastic theory, chained lines from the Timoshenko beam theory and broken lines from the Bernoulli-Euler beam theory. It can be seen from Fig. 3 that critical velocities increase with increasing the values of K and H.

As mentioned previously, Figs. 2 and 3 are also considered as the phase velocity spectrum of free waves along the beam, if β indicates the phase velocity and η the wave number. When $\eta \to 0$ in Fig. 2, the curve of the solid line for the first mode consists with the longitudinal wave speed $c = \sqrt{E/\rho}$ and the rest of curves tend to infinity. On the other hand, as η becomes large, the curve of the solid line for the first mode approaches to the Rayleigh wave speed and curves in the region of $\beta > 1$ to the velocity of distortional wave $\beta = 1$ [14]. The Timoshenko beam theory gives two dispersion curves and the curve for the first mode is in good agreement with that from the two-dimensional

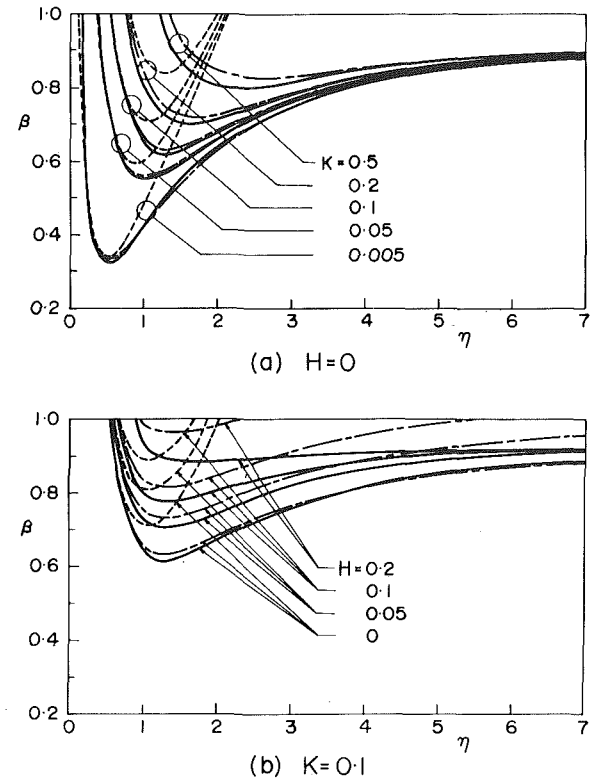


Fig. 3 Nondimensional velocity of the moving load β versus nondimensional integral variable η for lowest mode

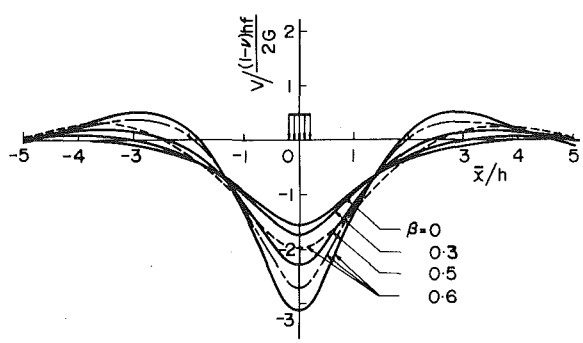


Fig. 4 Displacement of midplane of a beam versus $\bar{x}/h$ ($K = 0.1$, $H = 0.05$)

elastic theory. As the value of H increases, however, the discrepancies between the two become evident and even if $\eta \to \infty$, the curve from the Timoshenko beam theory does not approach to the Rayleigh wave speed. This phenomenon is due to using the shear coefficient obtained by Cowper. If the shear coefficient given by Mindlin's method [15] is used, the curve for the first mode obtained from the Timoshenko beam theory approaches to the Rayleigh wave speed as $\eta \to \infty$. In this case, however, it is noted that the shear coefficient κ becomes the function of the shear foundation modulus H, and the value of κ changes with the value of H.

Now let us return to a consideration of the response of a beam to moving loads. The curves of the solid line in Fig. 4 show the deflection profiles of the midplane of the beam ($y/h = 0.5$) for various values of load velocities lower than the critical velocity in the case of $K = 0.1$, $H = 0.05$. The displacement v is nondimensionalized by $(1 - \nu)hf/2G$ and $\bar{x}$ by h. Fifty terms of residues which belong to each of the upper and the lower half plane were taken in the prediction of the displacement of a beam, where the error of integral value is less than 10^{-5}. For $\beta = 0.6$, results obtained from the Timoshenko beam theory

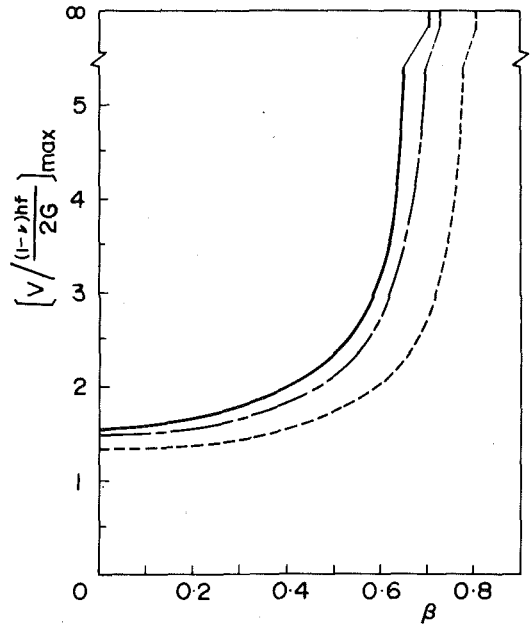


Fig. 5 Maximum value of displacement of midplane of a beam versus β ($K = 0.1$, $H = 0.05$)

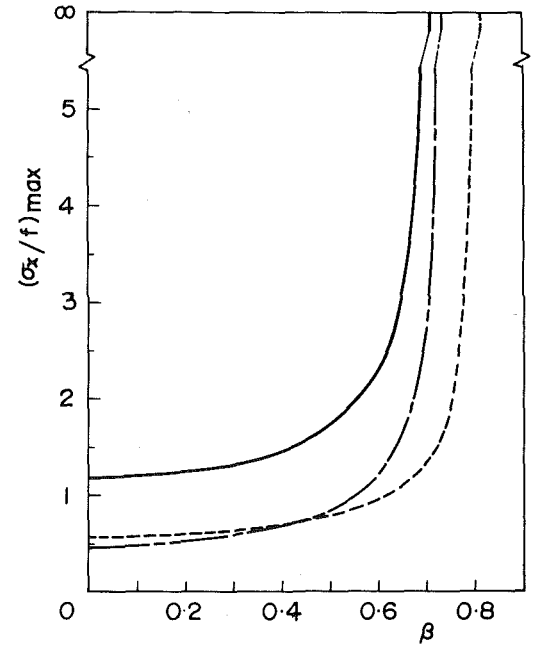


Fig. 6 Maximum value of fiber stress of upper surface of a beam versus β ($K = 0.1$, $H = 0.05$)

and the Bernoulli-Euler beam theory are superimposed by chained and broken lines, respectively. The deflection profile is symmetrical with respect to the load center. It can be observed from the figure that the Bernoulli-Euler beam theory gives the smallest peak deflection and the two-dimensional elastic theory gives the largest peak deflection among three theories.

The deflection is maximum under the load and increases monotonically to infinity as the load velocity approaches to the critical velocity. These maximum values of the deflection of the midplane of a beam are plotted as a function of the load velocity β in Fig. 5.

Fig. 6. is a plot of the fiber stress σ_x at the midposition of the load region on the upper surface of the beam as a function of the load velocity β.

Fig. 7 shows the deflection profile of the median plane of the beam for various values of velocities higher than the critical velocity in the case of $K = 0.1$ and $H = 0.05$. Solid, chained, and broken curves are obtained from the two-dimensional elastic theory, the Timoshenko beam theory, and the Bernoulli-Euler beam theory, respectively. It is seen from the figure that the deflection wave shapes lose their symmetry with respect to the point $\bar{x} = 0$ and the peak deflection occurs behind the moving load. When $\beta = 0.8$, the results from the Timoshenko beam theory are in good agreement with those from the two-dimensional elastic theory, though the curve from the Bernoulli-Euler beam theory is yet symmetrical relative to the position $\bar{x} = 0$ since the velocity $\beta = 0.8$ is just lower than the critical velocity obtained from the Bernoulli-Euler beam theory.

When the load velocity becomes higher than the Rayleigh wave speed, remarkable discripancies occur between the results from the two-dimensional elastic theory and those from the one-dimensional beam theories. In the case of $\beta = 0.95$, the deflection profile in front of the load obtained from the two-dimensional elastic theory seems to be a damped sinusoid and those from the one-dimensional beam theories are undamped sinusoids, though in rear of the load the deflection profiles from the three theories are almost good agreeable. If the shear coefficient for the Timoshenko beam is taken to be that by Mindlin's method, the amplitude of displacement from the Timoshenko beam theory will decrease rapidly with an increase in distance from the load.

When $\beta > 1$, the deflection profile in rear of the load obtained from the two-dimensional elastic theory is very complicated because it is

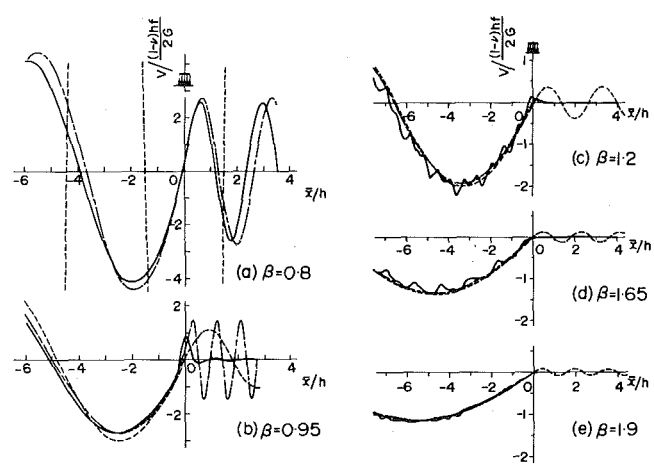


Fig. 7 Displacement of midplane of a beam versus $\bar{x}/h$ ($K = 0.1$, $H = 0.05$)

constructed with many higher modes. In the case of $\beta = 1.2$, the deflection curve in rear of the load obtained from the Bernoulli-Euler beam theory is an undamped sinusoid and that from the Timoshenko beam theory is also an undamped sinusoid except the neighbor region of the load. In front of the load, the deflection curve from the Bernoulli-Euler beam theory is an undamped sinusoid while the amplitudes of waves from the Timoshenko beam theory and the two-dimensional elastic theory decrease rapidly. In the case of $\beta = 1.65$, as the load velocity is between $\sqrt{E/\rho}$ and c_1, the Timoshenko beam is undisturbed in front of the moving load, but the displacement obtained from the two-dimensional elastic theory has negative value and approaches to zero as $\bar{x} \to \infty$. In the case of $\beta = 1.9$ where the load velocity is higher than $c_1 (> c_2)$, the beam is undisturbed in front of the load, i.e., the displacement of midplane of the beam is zero in $\bar{x}/h > \epsilon/h$. The results from the two-dimensional elastic theory and the Timoshenko beam theory can indicate this phenomenon accurately, though that from the Bernoulli-Euler beam theory cannot do so and the deflection profile is an undamped sinusoid in front of the moving load.

On the whole, it may be seen from curves in Fig. 7 that, though there are many sectional differences, the discrepancies between the deflection profiles in rear of the load from the two-dimensional elastic theory and the one-dimensional beam theories are not so serious and beam theories seem to afford a useful approximation in the rear of the load. It is supposed, however, that fiber stress of the beam obtained from the two-dimensional elastic theory is very complicated in rear of the load compared with that from the one-dimensional beam theories, because the beam response from the two-dimensional elastic theory is the sum of the numerous higher mode responses of a beam and the deflection profile is not even, while that from the Bernoulli-Euler beam theory is constructed with one sinusoid and that from the Timoshenko beam theory with two sinusoids.

Conclusions

The responses of an infinite beam on the Pasternak-type foundation to moving loads have been treated exactly by applying the two-dimensional elastic theory and the results obtained have been compared with those from the Timoshenko and the Bernoulli-Euler beam theories. Though the Bernoulli-Euler beam theory gives extremely inconsistent results in the front of the load for the moving velocity higher than the critical velocity, it seems to be reliable for all the velocities excluding the foregoing region. The results obtained from the Timoshenko beam theory may be seen to be in remarkably good agreement with those obtained from the exact two-dimensional elastic theory, but one can also see that higher modes of beam vibration give locally a great influence upon the response of a beam in the region of load velocities higher than the critical velocity.

Acknowledgment

The authors wish to thank Dr. S. Chonan of Tohoku University for his useful discussions on the work.

References

1 Dörr, J., "Der unendliche, federnd gebettete Balken unter dem Einfluss einer gleichförmig bewegten last," *Ingenieur-Archiv*, Vol. 14, 1943, pp. 167–192.

2 Kenney, J. T., Jr., "Steady-State Vibrations of Beam on Elastic Foundation for Moving Load," ASME JOURNAL OF APPLIED MECHANICS, Vol. 21, 1954, pp. 359–364.

3 Mathews, P. M., "Vibrations of a Beam on Elastic Foundation," *Zeitschrift für Angewandte Mathematik und Mechanik*, Vol. 38, 1958, pp. 105–115.

4 Mathews, P. M., "Vibrations of a Beam on Elastic Foundation II," *Zeitschrift für Angewandte Mathematik und Mechanik*, Vol. 39, 1959, pp. 13–19.

5 Keer, A. D., "The Continuously Supported Rail Subjected to an Axial Force and a Moving Load," *International Journal of Mechanical Science*, Vol. 14, 1972, pp. 71–78.

6 Holder, B. W., and Michalopoulos, C. D., "Response of a Beam on an Inertial Foundation to a Traveling Load," *AIAA Journal*, Vol. 15, 1977, pp. 1111–1115.

7 Crandall, S. H., "The Timoshenko Beam on an Elastic Foundation," *Proceedings of the Third Midwestern Conference on Solid Mechanics*, Ann Arbor, Mich., 1957, pp. 146–159.

8 Steele, C. R., "The Timoshenko Beam With a Moving Load," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, 1967, pp. 105–111.

9 Steele, C. R., "Beams and Shells With Moving Loads," *International Journal of Solids and Structures*, Vol. 7, 1971, pp. 1171–1198.

10 Chonan, S., "The Elastically Supported Timoshenko Beam Subjected to an Axial Force and a Moving Load," *International Journal of Mechanical Science*, Vol. 17, 1975, pp. 573–581.

11 Chonan, S., "Moving Harmonic Load on an Elastically Supported Timoshenko Beam," *Zeitschrift für Angewandte Mathematik und Mechanik*, Vol. 58, 1978, pp. 9–15.

12 Kerr, A. D., "Elastic and Viscoelastic Foundation Models," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, 1964, pp. 491–498.

13 Cowper, G. R., "The Shear Coefficient in Timoshenko's Beam Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, 1966, pp. 335–339.

14 Ewing, Jardetzky, and Press, *Elastic Waves in Layered Media*, McGraw-Hill, New York, 1957.

15 Mindlin, R. D., "Influence of Rotatory Inertia and Shear on Flexural Motions of Isotropic Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 18, 1951, pp. 31–38.

APPENDIX

Solution From the Bernoulli-Euler Beam Theory

The equation of motion of the beam with shear deformation and the rotatory inertia being ignored is

$$\frac{1+\nu}{6}Gh^3\frac{\partial^4 v}{\partial x^4}+\rho h\frac{\partial^2 v}{\partial t^2}+kv+d\frac{\partial v}{\partial t}-G_0\frac{\partial^2 v}{\partial x^2}=\begin{cases}-f, & |x-ct|\leqq\epsilon\\ 0, & |x-ct|>\epsilon\end{cases}$$

The resulting desplacement is given by

$$v\Big/\frac{(1-\nu)\,hf}{2G}=\frac{12}{1-\nu^2}\frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{1}{\eta\Delta_1}\{e^{i(\bar{x}/h-\epsilon/h)\eta}-e^{i(\bar{x}/h+\epsilon/h)\eta}\}d\eta$$

where

$$\Delta_1=\eta^4-\frac{6}{1+\nu}\left(\beta^2-\frac{2H}{1-\nu}\right)\eta^2-\frac{12}{1-\nu^2}iL\beta\eta+\frac{12}{1-\nu^2}K$$

Solution From the Timoshenko Beam Theory

The equations of motion of the beam where shear and rotatory inertia are incorporated are

$$\frac{1+\nu}{6}Gh^3\frac{\partial^2\theta}{\partial x^2}-G\kappa h\left(\frac{\partial v}{\partial x}+\theta\right)-\frac{\rho h^3}{12}\frac{\partial^2\theta}{\partial t^2}=0$$

$$\rho h\frac{\partial^2 v}{\partial t^2}-G\kappa h\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial\theta}{\partial x}\right)+kv+d\frac{\partial v}{\partial t}-G_0\frac{\partial^2 v}{\partial x^2}=\begin{cases}-f, & |x-ct|\leqq\epsilon\\ 0, & |x-ct|>\epsilon\end{cases}$$

where θ is the bending slope and κ is the shear coefficient.

The resulting displacement is given by

$$v\Big/\frac{(1-\nu)hf}{2G}=\frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{1}{\eta\Delta_2}[\{\beta^2-2(1+\nu)\}\eta^2-12K]\{e^{i(\bar{x}/h-\epsilon/h)\eta}-e^{i(\bar{x}/h+\epsilon/h)\eta}\}d\eta$$

where

$$\Delta_2=\begin{vmatrix}\{(\beta^2/2)-(1+\nu)\}\eta^2-6\kappa & -6i\kappa\eta\\ -(1-\nu)i\kappa\eta & \times\{(\kappa-\beta^2)(1-\nu)+2H\}\eta^2+2iL\beta\eta+2K\end{vmatrix}$$

C. R. Steele
Professor.
Mem. ASME

K. E. Barry
Research Assistant.

Applied Mechanics Division,
Stanford University,
Stanford, Calif. 94305

Asymptotic Integration Methods Applied to Rotating Beams

The in-plane vibrational characteristics of an off-axis clamped beam subjected to either compressive or tensile forces arising from steady rotation are studied. The differential equations of motion are cast into state vector form and solved using asymptotic matrix integration methods. The general theory of these methods is described in this paper and their application to the analysis of rotating beams is made. The advantages inherent in these methods with regard to accuracy, reduction of analytical complexity, and savings in computational costs are discussed.

1 Introduction

For mechanical systems with coefficients whose spatial variation is smooth, asymptotic integration methods can, for example, yield all higher modes for little computational effort. Since so much information can be gained from the simplest one-term asymptotic solution, i.e., a WKB solution, it is natural to seek to extend the range through including more terms of the asymptotic series, through recasting the asymptotic series, through alternate forms of the expansion utilizing special functions, etc. However, these methods are not well suited for general computer applications, particularly since the higher terms of an asymptotic series invariably involve successively higher derivatives of the system coefficients.

In this article the WKB method is investigated and applied to the analysis of a linear fourth-order differential equation with variable coefficients. For the first and second modes, WKB supplies a rough approximation which can be corrected using a convergent asymptotic matrix integration method. One such method, derived by Keller and Keller [1], has the WKB solution as its leading term. A second method termed WKB-direct integration iterates the WKB solution through an integral equation to converge to the exact solution. For modes three and above, the WKB one-term solution produces satisfactory results.

The simplicity of these three solutions, WKB, Keller–Keller, and WKB-direct integration, as well as the accuracy they afford are studied. Other methods available for comparative purposes include approximations based on energy principles, Myklestad finite differencing, perturbation expansions, and integrating matrices.

In order to fully explore the range of capabilities to be exhibited in the three solutions, both the buckling instability as well as the general response for a uniform cantilevered beam rotating at constant speed are addressed. Critical rotation speeds are determined for those beams oriented radially inward and subjected to compressive stresses. Tensile stresses alter the modal response of those beams oriented radially outward and yield modal frequencies which are dependent on the rotation speed.

2 Theory

To apply the WKB method to a solid mechanics problem, it is advantageous for both the analysis and the numerical evaluations to use an n-dimension system of first-order equations. The Hellinger-Reissner-Washizu formulation [2] is convenient for formulating the problem in state vector form.

2.1 WKB First Approximation. A discussion of the WKB method, its application to solid mechanics problems, and how successive corrections can be obtained is made by Steele [3]. In this section, the first-order approximation using the WKB method and state vector representation is presented. The general nth-order equation can be written in matrix form as

$$\mathbf{y}'(s) = \mathbf{A}(s)\mathbf{y}(s) \tag{1}$$

The first approximation to the general solution of (1) is [1, 3]

$$\mathbf{y}_1 = \mathbf{X} \exp(\mathbf{\Lambda})\mathbf{C} \tag{2}$$

The eigenvector $\mathbf{X}_n$ corresponding to an eigenvalue ξ_n' satisfies the matrix equation

$$[\mathbf{A} - \xi_n'\mathbf{I}]\mathbf{X}_n = \mathbf{0} \tag{3}$$

We consider only the situation in which all eigenvalues are distinct. Each eigenvalue is integrated to give the diagonal element of $\mathbf{\Lambda}$,

$$\xi_n = \int \xi_n' ds = \Lambda_{nn} \tag{4}$$

For convenience, the eigenvector $\mathbf{X}_n$ is decomposed into two parts,

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, June, 1980. Paper No. 80-WA/APM-32.

$$\mathbf{X}_n(\xi_n') = d_n(\xi_n')\mathbf{x}_n(\xi_n') \tag{5}$$

where d_n can be viewed as a normalizing or an amplitude modification scalar function. Denote by $\mathbf{Z}$ the matrix of eigenvectors associated with the transposed matrix. Where

$$[\mathbf{A}^t - \xi_n'\mathbf{I}]\mathbf{Z}_n = \mathbf{0} \tag{6}$$

$$\mathbf{Z}_n(\xi_n') = d_n(\xi_n')\mathbf{z}_n(\xi_n') \tag{7}$$

The determination of d_n follows from

$$(d_n)^2\mathbf{z}_n{}^t \times \mathbf{x}_n = \text{constant} \tag{8}$$

2.2 Keller-Keller Method. A thorough description of an expansion which is convergent and has the first approximation as its leading term is found in Keller and Keller [1]; the general steps taken in the derivation are outlined as follows.

By differentiating the first approximation given by the WKB method in (2), we find the exact differential equation which satisfies, $\mathbf{y}_1$,

$$\mathbf{y}_1' - (\mathbf{A} + \mathbf{X}'\mathbf{X}^{-1})\mathbf{y}_1 = 0 \tag{9}$$

The exact differential equation for $\mathbf{y}$ can be rewritten in the form

$$\mathbf{y}' - (\mathbf{A} + \mathbf{X}'\mathbf{X}^{-1})\mathbf{y} = -\mathbf{X}'\mathbf{X}^{-1} \tag{10}$$

The complementary solution of (10), treating the right-hand side as known, is given exactly by

$$\mathbf{y}_h = \mathbf{y}_1 \tag{11}$$

The particular solution is given by integration with respect to the appropriate Kernel function $\mathbf{K}(s, t)$

$$\mathbf{y}_p = \int \mathbf{K}(s, t)(-\mathbf{X}'\mathbf{X}^{-1})\mathbf{y}dt \tag{12}$$

where the Kernel matrix function is given

$$\mathbf{K}(s, t) = \mathbf{y}_h(s)\mathbf{y}_h{}^{-1}(t) \tag{13}$$

and the entire expression written with reference to (11) yields

$$\mathbf{y} = \mathbf{y}_1[\mathbf{I} - \int \mathbf{y}_1{}^{-1}(t)\mathbf{X}'\mathbf{X}^{-1}\mathbf{y}(t)dt] \tag{14}$$

or letting

$$\mathbf{y} = \mathbf{y}_1\mathbf{\Omega} \tag{15}$$

gives the equation for the correction Ω

$$\mathbf{\Omega} = \mathbf{I} - \int \mathbf{B}(t)\mathbf{\Omega}(t)dt \tag{16}$$

in which

$$\mathbf{B}(t) = \mathbf{y}_1{}^{-1}(t)\mathbf{X}'\mathbf{X}^{-1}\mathbf{y}_1(t) \tag{17}$$

Expanding this yields the convergent series

$$\mathbf{y} = \mathbf{y}_1[\mathbf{I} - \int \mathbf{B}(t)dt + \int \mathbf{B}(t)\int \mathbf{B}(u)du\, dt - \ldots] \tag{18}$$

The term in the brackets is usually rapidly convergent [3].

2.3 WKB-Direct Integration. Since (2) represents a first approximation to (1), then by substituting the expression for $\mathbf{y}$ in (2) into the right-hand side of the differential equation of (1) and integrating numerically, one obtains a new expression for $\mathbf{y}$ which is denoted as $\mathbf{y}_2$.

$$\mathbf{y}_2 = \int \mathbf{A}\mathbf{y}_1 ds \tag{19}$$

Repeating this procedure until convergence is obtained will yield

$$\mathbf{y}_3 = \int \mathbf{A}\mathbf{y}_2 ds \tag{20}$$

$$\vdots$$

$$\mathbf{y}_n = \int \mathbf{A}\mathbf{y}_{n-1} ds \tag{21}$$

$$\mathbf{y}_n = \mathbf{y}_{n-1} + \epsilon \tag{22}$$

where ϵ represents some small allowable difference in the $\mathbf{y}_n$ and $\mathbf{y}_{n-1}$ vectors. This is the basis of the WKB-direct integration procedure.

The noticeable absence of inverting matrices or taking derivatives lends to the analytical simplicity of this method. An understanding of some of the particulars present in the convergent process is important, however. A discussion of some of these characteristics can be found in the first part of Section 4.

3 Application to Rotating Beams

As stated previously, two analyses are conducted. The one with the beam oriented inward is representative of many machine elements and/or turbine blades. The second study, with the beam directed radially outward can be viewed as a model of satellite appendages, windmill, or helicopter blades.

Pictured in Fig. 1 is the relevant geometry associated with the two configurations. The beam of length l is cantilevered to a hub of radius R rotating at constant speed ω. The axis of rotation is perpendicular to the beam and passes through the hub center. The axial coordinate s is the distance from the rim of the hub to an arbitrary point on the beam and is considered positive directed toward the center of the hub.

For bending in the plane perpendicular to the axis of rotation, the equation for small lateral deflection w is

$$EI\, w'''' - (Tw')' = \rho A(\omega^2 + \beta^2)w \quad 0 < s < l \tag{23}$$

where the term $\rho A(\omega^2 + \beta^2)w$ represents the inertial force due to the angular velocity of the rotating frame and of vibration. The distribution of axial force is determined from the equilibrium equation,

$$T' = \rho A\omega^2(R - s) \tag{24}$$

Integration of (24) with the stipulation that the tip ($s = l$) be free from stress yields,

$$T = \rho A\omega^2[(R - l)^2 - (R - s)^2]/2 \tag{25}$$

Note that only a linear analysis is conducted and no other stabilizing or destabilizing effects, due to rotation, are considered.

In the matrix formulation, the dependent variable vector is taken to be

$$\mathbf{y} = [M\, H\, \psi\, w]^t \tag{26}$$

The Hellinger-Reissner-Washizu formulation gives

$$M' = H + T\psi$$

Nomenclature

$\mathbf{A}$ = prescribed square matrix whose components may depend on s
$\mathbf{C}$ = column vector of arbitrary constants
EI = bending stiffness
H = shear force
l = beam length
M = bending moment
R = radius of hub
s = independent variable, distance along beam from clamped end
T = tensile force
w = displacement amplitude of the beam (in the plane of rotation)
$\mathbf{X}$ = square matrix of eigenvectors of $\mathbf{A}$
$x = s/l$
$\mathbf{y}$ = dependent variable, column vector
$\alpha = R/l$
β = modal frequency
$\mathbf{\Lambda}$ = diagonal matrix of integrated eigenvalues of $\mathbf{A}$
ζ_n' = nondimensional eigenvalue of $\mathbf{A}$ matrix (ξ_n'/l)
$\eta = \rho A\beta^2 l^4/EI$ dimensionless frequency parameter
$\lambda = \rho A\omega^2 l^4/EI$ dimensionless spin parameter
ξ_n' = eigenvalue of $\mathbf{A}$ matrix
ρA = mass per unit length
ψ = bending angle
ω = rotation speed of hub
$(\ )'$ = spatial derivative d/ds or d/dx

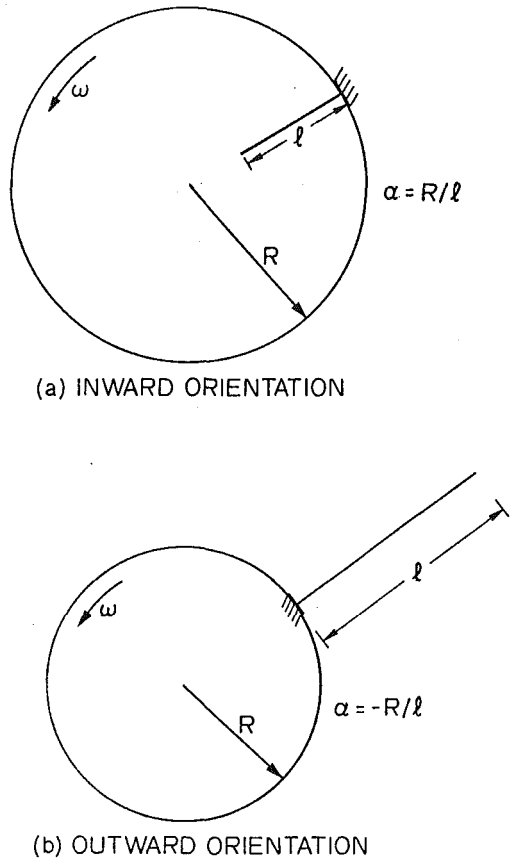


Fig. 1 Rotating beam geometry

$$H' = -\rho A(\omega^2 + \beta^2)w \tag{27}$$
$$\psi' = M/EI$$
$$w' = -\psi$$

The prime denotes differentiation with respect to the axial coordinate. These equations are in the form of (1), where the **A** matrix is given by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & T & 0 \\ 0 & 0 & 0 & -\rho A(\omega^2 + \beta^2) \\ l/EI & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{28}$$

To determine the eigenvalues of **A**, the determinant of (3) is set equal to zero which yields the following fourth-degree equation:

$$\xi'^4 - (T/EI)\xi'^2 - (\rho A/EI)(\omega^2 + \beta^2) = 0 \tag{29}$$

In the dimensionless terms given in the Nomenclature, (29) transforms to

$$\zeta'^4 - \lambda f(x)\zeta'^2 - (\lambda + \eta) = 0 \tag{30}$$

where

$$f(x) = [(\alpha - 1)^2 - (\alpha - x)^2]/2 \tag{31}$$

The four eigenvalues obtained from the roots of (30) are separate and distinct.

$$\zeta_1' = \zeta' = [\lambda f/2 + [(\lambda f/2)^2 + \lambda + \eta]^{1/2}]^{1/2}$$
$$\zeta_2' = -\zeta' \tag{32}$$
$$\zeta_3' = \zeta^{*\prime} i = [-\lambda f/2 + [(\lambda f/2)^2 + \lambda + \eta]^{1/2}]^{1/2}$$
$$\zeta_4' = -\zeta^{*\prime} i$$

The nth eigenvector corresponding to the nth eigenvalue is determined,

$$\mathbf{X}_n = d(\zeta_n') \begin{bmatrix} -(EI/l^2)\zeta_n'^2 \\ -(EI(\lambda + \eta)/l^3)/\zeta_n' \\ -\zeta_n'/l \\ l \end{bmatrix} \tag{33}$$

The eigenvector associated with the $\mathbf{A}^t$ matrix gives

$$\mathbf{Z}_n = d(\zeta_n') \begin{bmatrix} \zeta_n'/l \\ l \\ (EI/l^2)\zeta_n'^2 \\ -(EI(\lambda + \eta)/l^3)/\zeta_n' \end{bmatrix} \tag{34}$$

Letting

$$\mathbf{Z}_n{}^t(\zeta_n') \cdot \mathbf{X}_n(\zeta_n') = -1 \tag{35}$$

gives

$$d_n = d(\zeta_n') = [\zeta_n' l^3/2EI(\zeta_n'^4 + \lambda + \eta)]^{1/2} \tag{36}$$

The first approximation given in (2) can now be assembled; all necessary expressions having been appropriately identified. Before completing the last step of the analysis, which will give the frequency equation and mode shapes, it is useful to transform the expression

$$\mathbf{X} \exp \mathbf{\Lambda}\, \mathbf{C} = \mathbf{Y}\, \mathbf{C}^* \tag{37}$$

by making use of sine and hyperbolic sine functions in place of exponential functions. Boundary conditions for a cantilever beam are now applied,

$$\begin{aligned} &M(l) = H(l) = 0 && \text{Moment and shear vanish at the tip} \\ &\psi(0) = w(0) = 0 && \text{Bending angle and deflection are zero at the root} \end{aligned} \tag{38}$$

to give

$$\mathbf{Y}_p \mathbf{C}^* = 0 \tag{39}$$

The determinant of $\mathbf{Y}_p$ gives the following transcendental equation for the frequency:

$$\frac{\Xi^{*2}}{\Phi^*} + \frac{\Xi^2}{\Phi} + \left[\frac{\Xi^*\Xi}{\Phi^*\Phi}\right]^{1/2} [(\Xi^* + \Xi) \cos \textstyle\int_0^1 \zeta^{*\prime} dx \cosh \int_0^1 \zeta' dx + (\Xi^* - \Xi) \sin \int_0^1 \zeta^{*\prime} dx \sinh \int_0^1 \zeta' dx] = 0 \tag{40}$$

where

$$\Xi, \Xi^* = \zeta'(0), \quad \zeta^{*\prime}(0)$$
$$\Phi, \Phi^* = \Xi^4 + \lambda + \eta, \quad \Xi^{*4} + \lambda + \eta \tag{41}$$

The form of this equation is simplified by the fact that $\zeta'(1) = \zeta^{*\prime}(1)$. It is also of interest to note that for $\omega = 0$, $\Xi = \Xi^*$, and (40) reduces to the familiar frequency equation for the nonrotating beam.

$$1 + \cos \textstyle\int_0^1 \zeta^{*\prime} dz \cosh \int_0^1 \zeta' dx = 0 \tag{42}$$

For the instability study, α is considered positive and the search is for λ-critical which drives η, a measure of the beam natural frequency to zero. In the analysis of frequency modification for the outward oriented beam, α is negative and η is studied as λ takes on various prescribed values.

4 Results

The results presented in the next three subsections deal with the buckling problem of the inward oriented beam and with the general modal determination of a cantilevered rotating beam. Although the methods presented are valid for variable mass and stiffness properties, uniform properties are used in this analysis for the purposes of comparison. In order to investigate the range of accuracy afforded by the three solutions, α, the ratio of hub radius to beam length, and ω the rotation speed are varied.

Application of the solutions presented involve a search for either λ or η which will yield zero for the determinant of the $\mathbf{Y}_p$ matrix. The secant interpolation scheme gave very satisfactory results in all cases and usually zeroed in on the correct value within one or two trials.

Use of the WKB first approximation given in (40) is straightforward and requires no further comment. However, a few remarks concerning the integration procedures involved in the Keller-Keller and WKB-direct integration solutions are made, as they should be taken into consideration when formulating the problem.

Choice of the limits of integration for the eigenvalues of the $\mathbf{A}$ matrix and the eigenvectors $\mathbf{X}_n$ should be made such that the exp $[\Lambda]$ will yield values consistent with one another (i.e., of the same numerical order). In this study all integration is carried forward from 0.0 to x except in the evaluation of Λ_{11} and $\mathbf{X}_2$ where integration is backward from 1.0 to x. Use of Simpson quadrature with an integration step size of 0.02 is used throughout.

Orthogonality of the eigenvectors is applied during computation of the WKB-direct integration method. Upon iteration of the sinusoidal eigenvectors, any exponential residue will tend to amplify and produce a dependent system. To insure orthogonality and independence of the eigenvectors, a normalized dot product is subtracted from the sinusoidal eigenvectors prior to each integration step. This procedure not only proved useful when iterating on the eigenvectors, but also provides another check on the accuracy of the converged solution.

In deriving the current results using the Keller-Keller expansion, only a second approximation or one term correction is used. This is found by approximating $\mathbf{y}(t)$ in the first integrand as $\mathbf{y}_1(t)$ and ignoring all subsequent terms (see (18)).

Two methods are used to check the convergence of the WKB-direct integration method. Iteration is stopped when the sum of the squares of the integration endpoint of each eigenvector matches the previous two derived values within 1 percent or less. Also, the Rayleigh quotient is used to insure the last two integrations have yielded a converged solution as well as a general check on the value of λ or η produced.

4.1 Inward Oriented Beam—Instability Study. Various direct numerical methods have been applied to the subject of instability of a rotating beam within the past 10 years [4–11]. An approximate analysis [4] using quartic polynomial shape functions which satisfy all the boundary conditions obtained information for the range of $0.5 < \alpha < 2.0$. In the interest of determining the existence of a stability boundary, reference [5] obtained results with an analysis based on the extended Galerkin procedure using two Lengendre polynomials satisfying only the displacement boundary conditions as shape functions. Fig. 2 presents a comparison of the results found in [5] with those derived using the three methods presented herein toward the evaluation of the frequency Ω, derived from λ by the following:

$$\Omega = \alpha^2(\lambda)^{1/2} = \omega R^2(\rho A/EI)^{1/2} \tag{43}$$

In the range of $0.5 < \alpha < 2.0$ the values determined by Wang [5] and Mostaghel and Tajbaksh [4] overlay the iterative integral data. In the lower range of α, $0.0 < \alpha < 0.2$, values produced in references [7, 8, 18] support the data derived by the WKB-direct integration method; results given by Wang appear to be high. Table 1 provides a comparison of the three solutions for a few representative data points.

It is generally felt that the WKB method is most useful in the analysis of higher mode calculations; by utilizing this approximation in the iteratuve integral evaluation, exact results are achieved for all values of α, $0.05 < \alpha < 5.0$.

4.2 Transition Point—Axis Clamped Case. A notable point of interest to many researchers in this area has been the buckling instability for a rotating beam subjected entirely to tensile forces, $0.0 < \alpha < 0.5$. Reference [4] was the first to express surprise at this behavior and subsequent investigations [7, 8, 10] have established that $\alpha = 0$, the axis clampled beam, indeed represents the stability boundary for in-plane buckling of the first mode.

The only physical explanation thus far given for this behavior is credited to Weber [11] who cites the existence of centrifugal forces as the factor responsible for this buckling phenomenon. Centrifugal forces, for the in-plane bending problem, increase with increasing deflection, and therefore do represent a source of instability. However, the same centrifugal forces are present in both inward and outward oriented geometry ($+\alpha$ and $-\alpha$). A more complete explanation for the cause of the known stability boundary can be found with a little investigation of the WKB one-term solution.

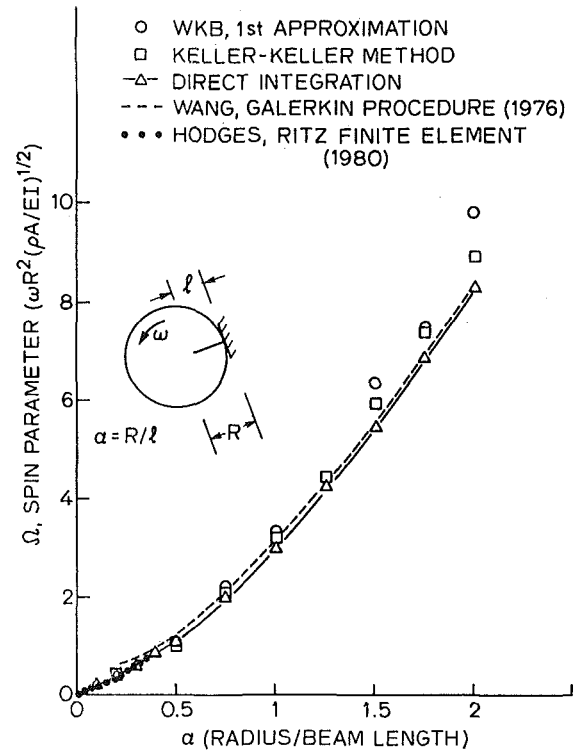


Fig. 2 Critical buckling speed

Table 1 Critical buckling speed (Ω)

α	WKB-Direct Integration (converged)	WKB (1-term)	Keller-Keller (2-term)	Hodges[18] Numerical Integration
.2	.35	.52	.42	.35
.5	1.13	1.06	1.07	1.13
1.0	2.96	3.32	3.25	2.99
1.5	5.35	6.36	5.93	5.38
2.0	8.27	9.80	8.98	8.19
3.0	14.91	19.77	16.06	14.87
4.0	22.70	30.15	24.24	22.77

It can be determined (see Fig. 2) that as α tends to zero, λ is rapidly increasing. Since it is known that the WKB approximation yields good results for large parameter values, it is surprising that the one-term approximation does not produce the expected results. An examination of the governing equations yields an understanding of the situation.

For large values of $\lambda(\eta = 0)$, the eigenvalues of (32) can be approximated as

$$\begin{aligned} \zeta_1' &\simeq (\lambda f)^{1/2} \\ \zeta_2' &= -\zeta_1' \\ \zeta_3', \zeta_4' &\simeq 0 \end{aligned} \tag{44}$$

It is important to note that ζ_1' is purely real for all values of x in the range under consideration.

The inability of the one-term approximation to give accurate predictions arises from two of the roots, corresponding to the oscillatory eigenvectors, being zero. To remedy this situation, the first two eigenvectors associated with ζ_1' and ζ_2' are kept and two more solutions

of the differential equation are sought. Rewriting (23) in dimensionless terms gives

$$w'''' - \lambda(fw')' - (\lambda + \eta)w = 0 \tag{45}$$

A linear variation in x is found to satisfy this equation exactly for $\eta = 0$.

$$w = D(x + \alpha) \tag{46}$$

where D is some arbitrary constant.

Boundary conditions at the tip dictate that the contribution of the exponentially increasing eigenvector must be zero. Focusing attention on the satisfaction of the root boundary conditions will give

$$w(0) = C_2 d_1[1] + d\alpha = 0 \tag{47}$$

$$w'(0) = C_2 d_1 \zeta_1'(0)[1] + D = 0 \tag{48}$$

which requires

$$D \simeq -(\lambda f(0))^{1/2} C_2 d_1 \tag{49}$$

$$D = -C_2 d_1/\alpha \tag{50}$$

Equations (49) and (50) can only both be true if α is positive, i.e., inward oriented beam. For the outward oriented geometry, α is negative, and the inability for the root boundary conditions to be satisfied excludes the possibility of buckling in the region. Physically, the buckled beam is conforming to a straight line rotation of the beam about the spin axis. This is possible for the inward oriented beam as the exponentially decreasing solution allows for a smooth transition in the deflection; however this is not possible for the outbound beam.

Also from equations (49) and (50) an approximate estimate for the spin parameter λ can be obtained as α tends toward zero.

$$\lambda^{1/2} \simeq (\alpha f(0)^{1/2})^{-1} = 1.414/\alpha(1 - \alpha) \tag{51}$$

This approximation is good to the order of $\lambda^{-1/2}$ and the percent error is of order λ^{-1}. Reference [10] using singular perturbation techniques gives the approximation as

$$\lambda^{1/2} \simeq 1.414/\alpha \tag{52}$$

In Table 2 a comparison is made with exact results using the WKB direct integration method to illustrate the accuracy of these approximations.

A study of the vibration response for higher modes in this region yielded similar behavior to that exhibited in the outward oriented geometry discussed in the next section. This is in agreement with the statement in reference [10] that the beam has exactly one buckled mode. From the analysis presented this follows fairly naturally, the absence of the oscillatory eigenvector solutions excludes the possibility of any other buckling configuration shape.

At this point it is appropriate to insert a note about the nature of the out-of-plane buckling problem. In this case, the influence of inertial forces is absent and only the first two terms of (45) remain in the differential equation [12]. A similar analysis to the one given for the in-plane problem results in a second-order matrix equation. While a full scale treatment will not be presented, a few salient points which can be easily verified using the WKB one-term approximation are made.

Table 2 Critical buckling speed ($\lambda^{1/2} = \Omega/\alpha^2$) for small values of α

α	WKB Direct Integration	Equation (51)	Equation (52)
.05	30.428	29.773	28.284
.075	20.067	20.385	18.856
.1	15.149	15.713	14.142
.15	11.109	11.092	9.428
.2	8.686	8.839	7.071

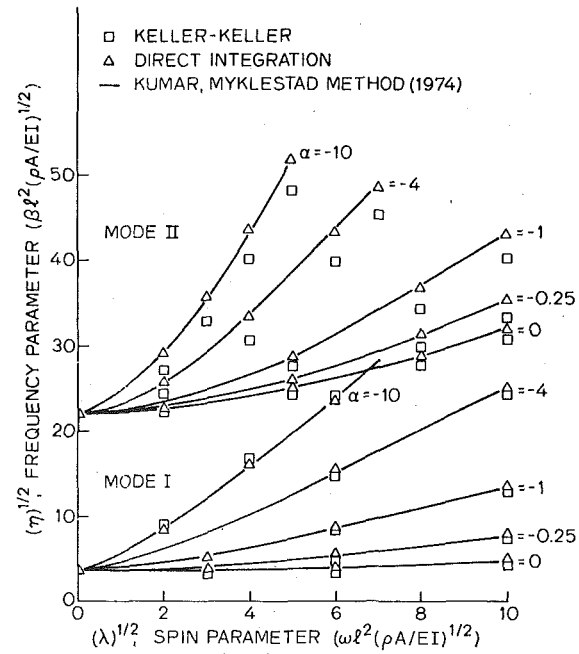


Fig. 3 Variation in the first and second natural frequencies

1 The buckling boundary present at $\alpha = 0.5$ is due to the inability of the resulting exponential solutions (purely real arguments depending on the large parameter λ) to satisfy the boundary conditions. This is similar to the transition point problem described in the foregoing, only in this event no solution varying linear with x is available.

2 The nature of the solution for the instability region $0.5 < \alpha$ must be divided into two subregions to be handled correctly.

(*a*) $0.5 < \alpha < 1$: In this range two turning points exist, the first at $x = \alpha$ and the second at $x = 1$. At $x = \alpha$ the tensile function f changes sign and causes a switch from oscillatory to exponential behavior, very similar to the behavior of the Airy function at $x = 0$. This type of change can be easily accommodated by the WKB method working with appropriately described beam sections and taking care to match the values at the turning point. The turning point at the boundary, $x = 1$, has not been a source of difficulty in any of the analysis thus far made.

(*b*) $1 < \alpha$: For this region, the beam is entirely under compression, and the two oscillatory solutions which result describe the nature of the buckling problem for the first and subsequent higher modes.

4.3 Outward Oriented Beam—Frequency Modification Study. The effect of rotation speed on the outward oriented beam natural frequencies has been a major subject of interest beginning with the design features of an airplane propeller blade. Early studies usually relied on the use of Rayleigh energy principles [12, 13] to investigate the bending frequencies of vibration for a given spin rate. As pointed out by Lo and Renbauger [14] a law of similarity exists which relates the natural frequency of vibrations occurring in any plane to that of another. The form this takes with regard to the in-plane or out-of-plane studies is as follows:

$$\beta^2_{\text{out-of-plane}} = \beta^2_{\text{in-plane}} + \omega^2 \tag{53}$$

Recent studies by Kumar [15] using the Myklestad method and by Nguyen and Hughes [16] and Peters [17] employing perturbation techniques are available for evaluating the calculations obtained herein.

Plotted in Fig. 3 are the frequencies for the first two modes found using the convergent expansion method and the iterative integral evaluation. Direct integration is seen to reproduce data extracted from Kumar's study for both the first and second modes as well as the third mode (not shown). For comparison of the three solutions, values are

Table 3 Modal frequencies ($\eta^{1/2}$) for moderate spin rates α = 0.0, −1.0, −4.0

$-\alpha$	$\lambda^{1/2}$	WKB-Direct Integration (converged)	WKB (1-term)	Keller-Keller (2-term)	Hodges[18] Numerical Integration
MODE I					
0.0	10.	5.08	-	4.22	5.05
1.0	10.	13.21	-	12.99	13.26
4.0	10.	24.98	-	24.07	24.92
MODE II					
0.0	10.	32.04	29.57	30.70	32.12
1.0	10.	43.24	38.90	40.26	43.23
4.0	7.	48.88	43.13	45.33	48.84
MODE III					
0.0	10.	74.23	73.07	73.13	73.98
1.0	10.	88.93	87.59	87.26	88.59
4.0	7.	96.62	93.72	94.62	96.20
MODE IV					
0.0	10.	135.41	134.05	134.25	(NA)
1.0	10.	153.62	150.72	151.02	(NA)
4.0	7.	162.38	159.84	160.57	(NA)

given in Table 3 along with data provided by Hodges [18] obtained by numerical integration to yield five place accuracy, rounded here to two decimal places. A dash in the tables signifies when a method is unable to give any useful information for a particular choice of parameters; (NA) denotes values were not available for inclusion in the table.

This analysis is valid also at higher rotation speeds. Fig. 4 presents data determined by WKB-direct integration and curves given by Nguyen and Hughes for zero hub radius (α = 0.). Good agreement is achieved for Modes II or higher, using either the Keller-Keller (not shown) or WKB-direct integration (see Table 4). At higher spin rates, an increased number of iterations are required for convergence. Iterations are necessary for the determination of the first two modes; the WKB 1st approximation is sufficient for obtaining information for the third mode and above even at the higher spin rates.

5 Conclusion

It has been proposed that the WKB perturbation method would be a useful analytical technique in conjunction with other numerical methods for the complete investigation of a problem typical of solid mechanics. From the results produced herein, it is found that the WKB method along with some means for obtaining successive corrections, such as the Keller-Keller or WKB-direct integration, provide all the desired information over the full range of parameter variation. The simplicity, accuracy, and low computer costs afforded by this method, especially when applied to the generation of higher modes, should not be ignored in the analysis of problems in solid mechanics.

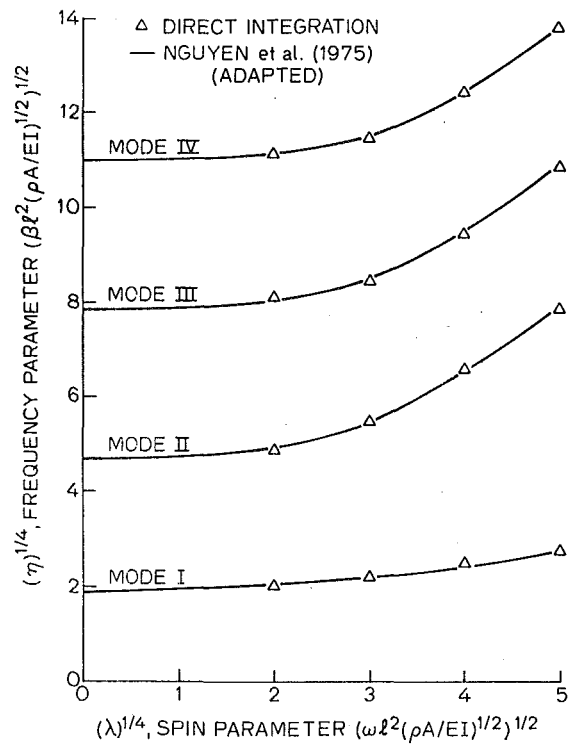


Fig. 4 Variation in the first four natural frequencies, α = 0

Table 4 Modal frequencies ($\eta^{1/2}$) for high spin rates (α = 0)

$\lambda^{1/2}$	WKB-Direct Integration (converged)	WKB (1-term)	Keller-Keller (2term)	Hodges[18] Numerical Integration
MODE I				
4.	3.91	-	3.56	(NA)
9.	4.85	-	4.12	(NA)
16.	6.16	-	4.33	(NA)
25.	7.60	-	-	(NA)
36.	9.16	-	-	(NA)
MODE II				
4.	23.92	23.49	23.60	23.94
9.	30.51	28.34	29.20	30.47
16.	43.26	37.83	40.14	43.26
25.	62.37	51.46	60.52	61.82
36.	85.36	68.45	85.15	85.50
MODE III				
4.	64.07	63.68	63.70	63.84
9.	72.08	71.08	71.13	71.83
16.	89.45	87.40	87.44	89.49
25.	117.92	113.11	114.87	117.54
36.	155.31	146.96	148.68	155.00
MODE IV				
4.	124.82	123.12	123.12	123.20
9.	133.53	131.66	131.70	132.05
16.	154.71	151.91	152.03	153.01
25.	189.78	186.38	186.56	188.70
36.	239.70	234.22	234.53	238.50

References

1 Keller, H. B., and Keller, J. B., "Exponential-Like Solutions of Systems of Linear Ordinary Differential Equations," *Journal of the Society of Industrial and Applied Mathematics*, Vol. 10, 1962, pp. 246–259.

2 Reissner, E., "Variational Methods and Boundary Conditions in Shell Theory," *Studies in Optimization, Proc. Symp. on Optimization*, Vol. 1, *Soc Indust and Applied Math.*, Philadelphia, Pa., 1970, pp. 79–94.

3 Steele, C. R., "Application of the WKB Method in Solid Mechanics," *Mechanics Today*, Vol. 3, 1976, pp. 243–295.

4 Mostaghel, N., and Tadjbaksh, I., "Buckling of Rotating Rods and Plates," *International Journal of Mechanical Science*, Pergamon Press, Vol. 15, 1973, pp. 429–434.

5 Wang, J. T. S., "On the Buckling of Rotating Rods," *International Journal of Mechanical Science*, Vol. 18, 1976, pp. 407–411.

6 Nachman, A., "The Buckling of Rotating Rods," TRANS. ASME. Mar. 1973, pp. 222–224.

7 Peters, D. A., and Hodges, D. H., "In-Plane Vibration and Buckling of a Rotating Beam Clamped Off the Axis of Rotation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, June 1980, pp. 398–402.

8 Kaza, K. R. V., Kvaternik, R. G., and White, W. F., "Buckling of Rotating Beams," *International Journal of Mechanical Science*, Vol. 21, 1979, pp. 739–745.

9 Lakin, W. D., and Nachman, A., "Unstable Vibrations and Buckling of Rotating Flexible Rods," *Quarterly of Applied Mathematics*, Jan. 1978, pp. 479–493.

10 Lakin, W. D., and Nachman, A., "Vibration and Buckling of Rotating Flexible Rods at Transitional Parameters Values," *Journal of Engineering Mathematics*, Vol. 13, No. 1, October 14, 1979.

11 Weber, H. I., "A Note of the Stability of a Rod Subjected to Compression by Centrifugal Forces," *Journal of Sound and Vibration,* Vol. 46, 1976, pp. 105–111.

12 Yntema, T., "Simplified Procedures and Charts for the Rapid Estimation of Bending Frequencies of Rotating Beams," NACA 3499, 1955.

13 Boyce, W. E., DiPrima, R. C., and Handelman, G. N., "Vibrations of Rotating Beams of Constant Section," Second U. S. National Congress of Applied Mechanics, 1956, pp. 165–173.

14 Lo, H., and Renbauger, J. L., "Bending Vibrations of a Rotating Beam," *Proceedings, First U. S. National Congress of Applied Mechanics,* ASME, 1952, pp. 75–79.

15 Kumar, R., "Vibrations of Space Booms Under Centrifugal Force Field," *Canadian Aeronautics and Space Institute Transactions,* Vol. 7, No. 1, Mar. 1974.

16 Nguyen, P. K., and Hughes, P. C., "Approximate Mode Shapes and Frequencies for Rotating Cantilevered Beams and Application to Flexible Spinning Spacecraft, *C.A.S.I. Transactions,* Vol. 8, No. 1, Mar. 1975.

17 Peters, D. A., "An Approximate Solution for the Free Vibrations of Rotating Uniform Cantilever beams," NASA Tm X-62,299, Sept. 1973.

18 Hodges, D. A., personal communication, 1980.

A. W. Leissa
Department of Engineering Mechanics,
Ohio State University,
Columbus, Ohio 43210
Mem. ASME

P. A. A. Laura
R. H. Gutierrez
Instituto de Mecanica Aplicada,
Base Naval Puerto Belgrano,
8111, Argentina

Vibrations of Rectangular Plates With Nonuniform Elastic Edge Supports

Two methods are introduced for the solution of free vibration problems of rectangular plates having nonuniform, elastic edge constraints, a class of problems having no previous solutions in the literature. One method uses exact solutions to the governing differential equation of motion, and the other is an extension of the Ritz method. Numerical results are presented for problems having parabolically varying rotational constraints.

Introduction

There is a wealth of literature dealing with the free vibrations of elastic, rectangular plates. Indeed, various summaries [1–4] identify at least 500 references, about half of which deal with plates governed by the classical differential equation

$$D\nabla^4 w + \rho \frac{\partial^2 w}{\partial t^2} = 0 \tag{1}$$

that is, plates which are isotropic, homogeneous, thin, uniform thickness, which are not subjected to in-plane forces or surrounding fluids, and which undergo small amplitude vibrations.

The vast majority of the aforementioned references deal with the classical boundary conditions representing clamped, simply supported or free edges, and a much smaller number can be found which treat edges which are elastically restrained against translation and/or rotation. A notable early effort was made by Lundquist and Stowell [5], who studied the problem of *buckling* of simply supported rectangular plates having two opposite edges elastically restrained against rotation, and obtained extensive numerical results which, by analogy, can be directly utilized in the vibration problem. Other relatively early works dealing with the vibrations of elastically supported, rectangular plates include those by Das [6], Carmichael [7], Joga-Rao and Kantham [8], Chulay [9], Stokey, Zorowski, and Appl [10], and Hoppmann and Greenspon [11]. More recently the effects of elastic constraints upon the free vibrations of rectangular plates have been extensively studied by Laura, et al. [12–17], as well as by others [18–20].

However, all of the foregoing references consider only elastic constraints which are *uniform* along a given boundary. Elastic constraints represent stiffness coupling with surrounding support structure and, in practical application, the stiffness of such surrounding structure along the common boundary will not be constant, but will vary from point to point. Indeed, the variation in stiffness of the support structure is often known, either from experimental tests or from other numerical evaluations, particularly finite-element analyses.

Recently, the present authors [21] demonstrated how one can solve the vibration problem for *nonuniform* elastic translational and/or rotational constraints in the case of a *circular* plate. The method utilized the exact, classical solution of the differential equation in polar coordinates expressed in Bessel functions, and expressed the arbitrarily varying spring coefficients in terms of Fourier series. The procedure was subsequently utilized by Narita and Leissa [22, 23] to analyze the vibrations of simply supported and free circular plates having uniform elastic constraints along part of an edge.

In the present work the free vibrations of rectangular plates having nonuniform elastic constraints are studied. Two methods for dealing with such problems are demonstrated. The first, utilizing exact solutions to equation (1) in rectangular coordinates, treats problems having two opposite edges simply supported. The second procedure shows how the well known Ritz [24] method can be applied to a more general class of problems. Numerical results are obtained in both instances for the case when two edges have parabolically varying rotational springs, and are compared with each other and with other results in certain limiting cases.

Exact Solutions

The classical governing differential equation of transverse motion for a plate is given by equation (1). Exact solutions for equation (1) in rectangular coordinates may be taken as (cf., ref. [1, p. 45])

$$w(x, y, t) = W(x, y) \sin \omega t \tag{2}$$

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, May, 1980; final revision, June, 1980. Paper No. 80-WA/APM-28.

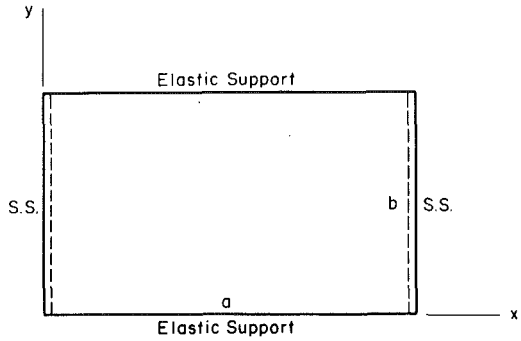


Fig. 1 Rectangular plate, simply supported on two opposite edges, elastically supported (translational and rotational) on the others

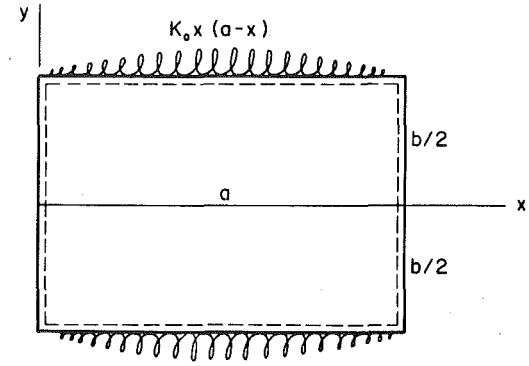


Fig. 2 Simply supported plate with rotational springs of parabolically varying stiffness along two opposite sides

$$W(x, y) = \sum_{m=1}^{\infty} [A_m \sin \sqrt{k^2-\alpha_m{}^2}\, y + B_m \cos \sqrt{k^2-\alpha_m{}^2}\, y + C_m \sinh \sqrt{k^2+\alpha_m{}^2}\, y + D_m \cosh \sqrt{k^2+\alpha_m{}^2}\, y] \sin \alpha_m x \quad (3)$$

where

$$\alpha_m = m\pi/a \quad (m = 1, 2 \ldots)$$
$$k^2 = \omega \sqrt{\rho/D}$$
$$k^2 - \alpha_m{}^2 > 0 \quad (4)$$

If $k^2 - \alpha_m{}^2 < 0$, then $\sin \sqrt{k^2-\alpha_m{}^2}\, y$ and $\cos \sqrt{k^2-\alpha_m{}^2}\, y$ must be replaced by $\sinh \sqrt{\alpha_m{}^2-k^2}\, y$ and $\cosh \sqrt{\alpha_m{}^2-k^2}\, y$.

Consider now a rectangular plate having dimensions $a \times b$, simply supported along the boundaries $x = 0$ and $x = a$, and elastically supported along $y = 0$ and $y = b$, as shown in Fig. 1. It is well known, and easily shown, that the classic solution of Voigt [25] given previously satisfies the boundary conditions at $x = 0, a$ exactly. The elastic support conditions on the other two edges are (cf., [1, p. 114]):

$$M_y(x, 0) = -K_1 \frac{\partial W}{\partial y}(x, 0) \quad (5a)$$

$$M_y(x, b) = K_2 \frac{\partial W}{\partial y}(x, b) \quad (5b)$$

$$V_y(x, 0) = K_3\, W(x, 0) \quad (5c)$$

$$V_y(x, b) = -K_4\, W(x, b) \quad (5d)$$

where now the translational and rotational spring coefficients $K_1, \ldots, K_4$ are not necessarily constants, but in general are functions of x; that is, *nonuniform*. The bending moment and edge reaction are, as usual, given by

$$M_y = -D\left(\frac{\partial^2 W}{\partial y^2} + \upsilon \frac{\partial^2 W}{\partial x^2}\right) \quad (6a)$$

$$V_y = Q_y + \frac{\partial M_{xy}}{\partial x} = -D\frac{\partial}{\partial y}\nabla^2 W - D(1-\upsilon)\frac{\partial^2 W}{\partial x^2 \partial y} \quad (6b)$$

Now expand each K_i into a Fourier cosine series. For example,

$$K_1(x) = \sum_{n=0}^{\infty} a_n \cos \alpha_n x \quad (\alpha_n = n\pi/a) \quad (7)$$

where

$$a_n = \frac{2}{a}\int_0^a K_1(x) \cos \alpha_n x dx, \quad a_0 = \frac{1}{a}\int_0^a K_1(x) dx \quad (8)$$

and similarly for K_2, K_3, and K_4. Substituting equations (3), (6a), and (7) into equation (5a) yields

$$-D \sum_{m=1}^{\infty} \{A_m \cdot 0 + B_m [-(k^2 - \alpha_m{}^2) - \upsilon\alpha_m{}^2] + C_m \cdot 0 + D_m [+(k^2 + \alpha_m{}^2) - \upsilon\alpha_m{}^2]\} \sin \alpha_m x \quad (9)$$

$$= -\left\{\sum_{n=0}^{\infty} a_n \cos \alpha_n x\right\} \sum_{m=1}^{\infty} \{A_m \sqrt{k^2-\alpha_m{}^2} - B_m \cdot 0 + C_m \sqrt{k^2+\alpha_m{}^2} + D_m \cdot 0\} \sin \alpha_m x \quad (9)$$

(*Cont.*)

The products of two infinite sums, as seen on the right-hand-side of equation (9) can be reduced to a single infinite sum by utilizing the identity

$$\sin \alpha_m x \cdot \cos \alpha_m x = \tfrac{1}{2} \sin (\alpha_m - \alpha_n)x + \tfrac{1}{2} \sin (\alpha_m + \alpha_n)x \quad (10)$$

Repeating this procedure for equations (5b), (5c), and (5d) and equating coefficients of like terms $\sin \alpha_m x$ in the resulting four equations yields a four fold infinite set of homogeneous equations in $A_m, \ldots, D_m$. The eigenvalues, or nondimensional frequency parameters, $(ka)^2 = \omega a^2 \sqrt{\rho/D}$ are determined by setting the determinant of the coefficient matrix equal to zero. Convergent solutions are determined to any desired degree of accuracy by successive truncation of the infinite determinant into a set of determinants or order $4p$ ($p = 1, 2, \ldots$). Eigenfunctions are obtained, in the usual manner, by back-substitution of the eigenvalues to determine the amplitude ratios B_m/A_m, C_m/A_m, and D_m/A_m.

Example 1: Simply Supported Plate With Symmetric, Nonuniform, Rotational Constraints

To demonstrate the solution procedure previously described, consider the rectangular plate simply supported along all four edges, restrained by nonuniform rotational springs of parabolically varying stiffness coefficient along two opposite edges (see Fig. 2). Specifically, take

$$K_1 = K_2 = K_0 x(a - x) \quad (11)$$

where K_0 is a constant, which yields a problem having twofold symmetry. A new choice of coordinate origin is shown in Fig. 2 to take advantage of the symmetry.

The problem uncouples into four symmetry classes, corresponding to vibration modes which are either symmetric or antisymmetric with respect to the two symmetry axes of the problem. We will restrict ourselves to finding the fundamental (i.e., lowest) frequency, which is associated with the first doubly symmetric mode. Therefore, $A_m = C_m = 0$, and m are odd integers in equation (3), and the two boundary conditions to be satisfied are

$$W(x, b/2) = 0 \quad (12a)$$

$$M_y(x, b/2) = K_2 \frac{\partial W}{\partial y}(x, b/2) \quad (12b)$$

Substituting equation (3) into (12a) permits exact satisfaction of the latter, provided that (for $k^2 > \alpha_m{}^2$)

$$B_m \cos \sqrt{k^2 - \alpha_m{}^2} \cdot \frac{b}{2} + D_m \cosh \sqrt{k^2 + \alpha_m{}^2} \cdot \frac{b}{2} = 0 \quad (13)$$

hence

$$W(x, y) = \sum_{m=1}^{\infty} B_m \left[\cos \sqrt{k^2 - \alpha_m{}^2}\, y - \left(\frac{\cos \sqrt{k^2 - \alpha_m{}^2} \cdot \frac{b}{2}}{\cosh \sqrt{k^2 + \alpha_m{}^2} \cdot \frac{b}{2}}\right) \cosh \sqrt{k^2 + \alpha_m{}^2}\, y\right] \sin \alpha_m y \quad (14)$$

The last boundary condition is satisified by expanding $K_2(x)$ into a Fourier cosine series corresponding to the parabolic variation given in equation (11) and substituting it and equation (14) into (12b). The resulting characteristic determinant of infinite order will be

$$|M_{ij}| = 0 \quad (15)$$

where

$$M_{ii} = 2E_{2i-1} + (2\, a_0{}^1 - a_{4i-2}{}^1)F_{2i-1} \quad (16a)$$

$$M_{ij} = \begin{cases} [a_{2(i-j)}{}^1 - a_{2(i+j-1)}{}^1]F_{2j-1} & (i > j) \\ [a_{2(j-1)}{}^1 - a_2{}^1(i + j - 1)]F_{2j-1} & (i < j) \end{cases} \quad (16b)$$

and where, for $(ka)^2 - m^2\pi^2 > 0$,

$$E_m = 2(ka)^2 \cos\left[\frac{\lambda'}{2}\sqrt{(ka)^2 - m^2\pi^2}\right] \quad (17a)$$

$$F_m = \sqrt{(ka)^2 - m^2\pi^2} \sin\left[\frac{\lambda'}{2}\sqrt{(ka)^2 - m^2\pi^2}\right] + \sqrt{(ka)^2 + m^2\pi^2} \cos\left[\frac{\lambda'}{2}\sqrt{(ka)^2 - m^2\pi^2}\right] \cdot \tanh\left[\frac{\lambda'}{2}\sqrt{(ka)^2 + m^2\pi^2}\right] \quad (17b)$$

and for $(ka)^2 - m^2\pi^2 < 0$,

$$E_m = -2(ka)^2 \cosh\left[\frac{\lambda}{2}\sqrt{m^2\pi^2 - (ka)^2}\right] \quad (18a)$$

$$F_m = \sqrt{m^2\pi^2 - (ka)^2} \sinh\left[\frac{\lambda}{2}\sqrt{m^2\pi^2 - (ka)^2}\right] - \sqrt{m^2\pi^2 + (ka)^2} \cosh\left[\frac{\lambda}{2}\sqrt{m^2\pi^2 - (ka)^2}\right] \times \tanh\left[\frac{\lambda}{2}\sqrt{m^2\pi^2 + (ka)^2}\right] \quad (18b)$$

with

$$\lambda = \frac{b}{a}, \quad a_0{}^1 = \frac{K_0 a^3}{6D}, \quad a_n{}^1 = -\frac{K_0 a^3}{n^2\pi^2 D} \quad (n = 2, 4 \ldots).$$

Ritz Solutions

Consider next, the rectangular plate simply supported all around, restrained by uniform rotational springs along the edges $\bar{x} = \pm a/2$ and nonuniform rotational springs along $\bar{y} = \pm b/2$ (see Fig. 3). The previous procedure, utilizing exact solutions, is not suitable for this problem, but the method of Ritz [24] can be extended.

Assuming simple harmonic motion in time as before (equation (2)), the functional to be minimized is

$$I(W) = V_{\max} - T_{\max} = \frac{D}{2}\iint_A \left(\frac{\partial^2 W}{\partial \bar{x}^2} + \frac{\partial^2 W}{\partial \bar{y}^2}\right) d\bar{x}d\bar{y} + \frac{1}{2}\oint_B m(s)\frac{\partial W}{\partial n} ds - \frac{\rho w}{2}\iint_A W^2 d\bar{x}d\bar{y} \quad (19)$$

where $V_{\max}$, the maximum potential energy, consists of the first area integral and the line integral around the boundary, $m(s)$ is the boundary value of the edge moment, and $T_{\max}$, the maximum kinetic energy, is represented by the last area integral. It is generally known that the strain energy, given by the first integral, reduces to the relatively simple form given above for polygonal plates having zero boundary displacement.

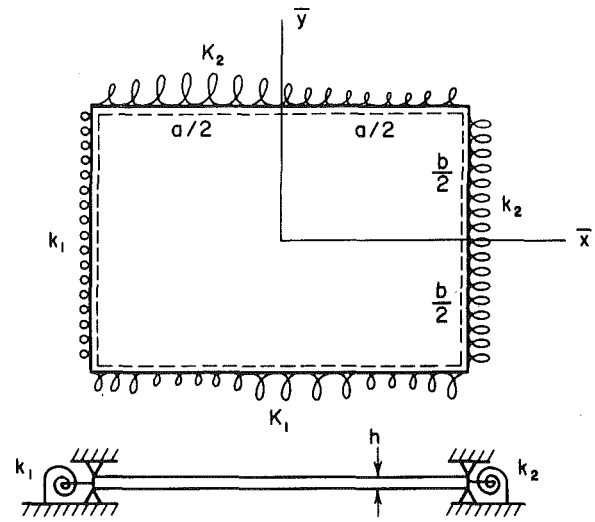


Fig. 3 Simply supported plate with uniform (k_1, k_2) and nonuniform (K_1, K_2) rotational springs

The geometric boundary conditions, which must be satisfied exactly when using the method, are

$$W(-a_1, \bar{y}) = W(a_1, \bar{y}) = W(\bar{x}, -b_1) = W(\bar{x}, b_1) = 0 \quad (20)$$

where $a_1 \equiv a/2$ and $b_1 \equiv b/2$. The boundary conditions of elastic, rotational constraint are given by

$$k_1 \frac{\partial W}{\partial \bar{x}}(-a_1, \bar{y}) = D\left[\frac{\partial^2 W}{\partial \bar{x}^2}(-a_1, \bar{y}) + \upsilon \frac{\partial^2 W}{\partial \bar{y}^2}(-a_1, \bar{y})\right] \quad (21a)$$

$$k_2 \frac{\partial W}{\partial \bar{x}}(a_1, \bar{y}) = -D\left[\frac{\partial^2 W}{\partial \bar{x}^2}(a_1, \bar{y}) + \upsilon \frac{\partial^2 W}{\partial \bar{y}^2}(a_1, \bar{y})\right] \quad (21b)$$

$$K_1(\bar{x}) \frac{\partial W}{\partial \bar{y}}(\bar{x}, -b_1) = D\left[\frac{\partial^2 W}{\partial \bar{y}^2}(\bar{x}, -b_1) + \upsilon \frac{\partial^2 W}{\partial \bar{x}^2}(\bar{x}, -b_1)\right] \quad (21c)$$

$$K_2(\bar{x}) \frac{\partial W}{\partial \bar{y}}(\bar{x}, b_1) = -D\left[\frac{\partial^2 W}{\partial \bar{y}^2}(\bar{x}, b_1) + \upsilon \frac{\partial^2 W}{\partial \bar{x}^2}(\bar{x}, b_1)\right] \quad (21d)$$

Equations (21) need not be satisfied by the choice of trial functions, $W(\bar{x}, \bar{y})$, but if they can be, the resulting solution will usually be improved.

At this point nondimensional variables will be adopted; that is,

$$x = \frac{\bar{x}}{a_1}, \quad y = \frac{\bar{y}}{b_1} \quad (22)$$

A two-term trial function will be taken as

$$W(x, y) = A_1 X_1(x) Y_1(y) + A_2 X_2(x) Y_2(y) \quad (23)$$

$$X_1(x) = \alpha_{41}x^4 + \alpha_{31}x^3 + \alpha_{21}x^2 + \alpha_{11}x + 1 \quad (24a)$$

$$X_2(x) = \alpha_{62}x^6 + \alpha_{52}x^5 + \alpha_{42}x^4 + \alpha_{32}x^3 + x^2 \quad (24b)$$

$$Y_1(y) = \beta_{41}y^4 + \beta_{31}y^3 + \beta_{21}y^2 + \beta_{11}y + 1 \quad (24c)$$

$$Y_2(y) = \beta_{62}y^6 + \beta_{52}y^5 + \beta_{42}y^4 + \beta_{32}y^3 + y^2 \quad (24d)$$

where A_1, A_2 and the α_{ij} and β_{ij} are arbitrary constants. The α_{ij} can be chosen in such a manner that six boundary condition equations (20), (21a) and (21b) are satisfied *exactly*. Performing the necessary substitutions yields

$$\alpha_{11} = \frac{2(k_1' - k_2')}{k_1' k_2' + 4(k_1' + k_2') + 15} \quad (25)$$

Table 1 Frequency parameters $\omega a^2 \sqrt{\rho/D}$ for simply supported plates having parabolically varying rotational contraints on two opposite edges (N = order of determinant, using exact solutions)

$\frac{a}{b}$	$\frac{K_0 a^3}{D}$	N: 1	2	3	4	---	12	Ritz method
1	0	(all sides simply supported, [1])					($2\pi^2$ = 19.739)	---
	0.1	19.757					19.757	19.765
	1	19.915					19.915	20.114
	10	21.236	21.235				21.235	24.068
	100	25.802	25.799				25.799	28.193
	∞	(two opposite sides clamped, [26])					(28.951)	---
0.5	0	(all sides S.S., [1])					($1.25\pi^2$ = 12.337)	---
	0.1	12.341					12.341	12.344
	1	12.372					12.372	12.401
	10	12.621					12.621	12.975
	100	13.320	13.320	13.320	13.319		13.319	13.598
	∞	(two opposite sides clamped, [27])					(13.688)	---
2	0	(all sides S.S., [1])					($5\pi^2$ = 49.348)	---
	0.1	49.405					49.405	49.428
	1	49.914					49.914	50.981
	10	54.388	54.223				54.387	71.545
	100	74.241	74.223				74.223	91.890
	∞	(two opposite sides clamped, [27])					(95.256)	---

$$\alpha_{21} = \frac{k_1' + 3}{k_1' + 5}(\alpha_{11} - 2)$$

$$\alpha_{31} = -\alpha_{11}$$

$$\alpha_{41} = -1 - \alpha_{21}$$

$$\alpha_{32} = \frac{2(k_1' - k_2')}{k_1' k_2' + 8(k_1' + k_2') + 63}$$

$$\alpha_{42} = \frac{k_1' + 7}{k_1' + 9}(\alpha_{32} - 2)$$

$$\alpha_{51} = -\alpha_{32}$$

$$\alpha_{62} = -1 - \alpha_{42}$$

where

$$k_1' \equiv \frac{k_1 a_1}{D}, \quad k_2' \equiv \frac{k_2 a_1}{D} \tag{25}$$

(*Cont.*)

The β_{ij} can be chosen so as to *approximate* the remaining two boundary condition equations (21c) and (21d). Two possible procedures are as follows:

1 Boundary collocation at selected points along the edges $y = \pm b_1$.

2 Setting the integrals of equations (21c) and (21d), integrated along the edges $y = \pm b_1$, equal to zero.

The approximate eigenvalues (and nondimensional frequencies) and eigenfunctions (and amplitude ratios A_2/A_1) are determined from the Ritz minimizing equations

$$\frac{\partial I(W)}{\partial A_i} = 0 \quad (i = 1, 2) \tag{26}$$

Example 2: Simply Supported Plate With Both Uniform and Nonuniform, Rotational Constraints

Let the uniform spring coefficients along the edges $\bar{x} = \pm a_1$ remain as k_1 and k_2, and choosing the nonuniform coefficients as in the previous example to vary parabolically, that is, in terms of the present coordinates,

$$K_1(\bar{x}) = K_2(\bar{x}) = K_0(a_1^2 - \bar{x}^2) \tag{27}$$

where, again, K_0 is a constant. Satisfying boundary conditions (21c) and (21d) exactly at the points $(\bar{x}, \bar{y}) = (0, \pm b_1)$ determines the β_{ij} as

$$\beta_{21} = -2\,\frac{\overline{K}_0 + 3}{\overline{K}_0 + 5}, \quad \beta_{41} = -1 - \beta_{21}$$

$$\beta_{42} = -2\,\frac{\overline{K}_0 + 7}{\overline{K}_0 + 9}, \quad \beta_{62} = -1 - \beta_{42}$$

$$\beta_{31} = \beta_{11} = \beta_{52} = \beta_{32} = 0$$

where

$$\overline{K}_0 = \frac{K_0 a_1^3 \lambda}{D}, \quad \lambda \equiv \frac{b}{a} = \frac{b_1}{a_1} \tag{28}$$

At this point all the constants save A_1 and A_2 in equations (23) and (24) are determined. Substituting equations (23) into (26) yields the a characteristic determinant in the form

$$\begin{vmatrix} (L_{11} - \Omega^2 J_{11}) & (L_{12} - \Omega^2 J_{12}) \\ \text{symmetric} & (L_{22} - \Omega^2 J_{13}) \end{vmatrix} = 0 \tag{29}$$

where Ω is the nondimensional frequency parameter given by

$$\Omega = \omega a^2 \sqrt{\rho/D} \tag{30}$$

and L_{ij} and J_{ij} are simple polynomials involving λ, the α_{ij} and the β_{ij} and are given in detail in reference [28].

Numerical Results and Discussion

The problem of the simply supported rectangular plate having parabolically varying rotational constraints on two opposite edges was solved using both the exact solutions to the differential equation and the Ritz method. In the latter approach the present problem is a special case when $k_1 = k_2 = 0$ in the preceding two sections.

Table 2 Frequency parameters for plates having two opposite edges clamped and parabolically varying rotational constraints on the simply supported edges (Ritz method)

$\frac{a}{b}$	$\frac{K_0 a^3}{D}$	$\omega a^2 \sqrt{\rho/D}$
1	0	(28.951)*
	0.1	28.969
	1	29.219
	10	32.179
	100	35.379
	∞	(35.992)*
0.5	0	(23.814)**
	0.1	23.844
	1	23.876
	10	24.136
	100	24.561
	∞	(24.56)**

* from [26]
** from [27]

In Table 1 numerical results for the nondimensional frequency parameter $\omega a^2 \sqrt{\rho/D}$ are presented for three aspect ratios—a/b = 1, 0.5, and 2. The nondimensional spring stiffness parameter $K_0 a^3/D$ is also varied between its limiting values of 0 and ∞. The former limit corresponds to the problem of all sides simply supported, the most simple, classic case having an exact solution. The latter limit yields two opposite sides clamped, also having a relatively well-known exact solution.

The method using exact solutions is seen to converge rapidly as the order of the truncated determinant is increased, yielding at least five significant figure accuracy for the frequencies from second-order determinants. The second-order determinant (29) arising from the Ritz method is seen to yield less accurate results, with accuracy decreasing as either $K_0 a^3/D$ or a/b increases. The Ritz method yields upper bounds on the frequencies, as it is to be expected. The two procedures discussed previously for determining the constants β_{ij} yielded virtually identical results when the boundary points $(\bar{x}, \bar{y}) = (0, \pm b/2)$ (see Fig. 3) are used for the collocation. The frequency parameters are independent of Poisson's ratio.

Table 2 displays numerical results for the case when $k_1 = k_2 = \infty$; that is, a rectangular plate having its edges $\bar{x} = \pm a/2$ clamped, while the remaining two are simply supported and have parabolically varying rotational constraints. The results are obtained by the Ritz method and, again, do not depend upon Poisson's ratio. Comparisons are once again made with limiting cases of $K_0 a^3/D = 0$ (simply supported edges only) and ∞ (all four edges fully clamped) for which results are available in the literature cited.

References

1 Leissa, A. W., *Vibration of Plates,* NASA SP-160, U.S. Govt. Printing Office, 1969.

2 Leissa, A. W., "Recent Research in Plate Vibrations: Classical Theory," *The Shock and Vibration Digest,* Vol. 9, No. 10, 1977, pp. 13–24.

3 Leissa, A. W., "Recent Research in Plate Vibrations: 1973–1976: Complicating Effects," *The Shock and Vibration Digest,* Vol. 10, No. 12, 1978, pp. 21–35.

4 Leissa, A. W., Notes for Short Course in "Vibrations of Beams, Plates and Shells," Ohio State University, Sept. 1979.

5 Lundquist, E. E., and Stowell, E., "Critical Compressive Stress for Flat Rectangular Plates Supported Along All Edges and Elastically Restrained Against Rotation Along the Unloaded Edges," NACA Rept. 733, 1942.

6 Das, Y. C., "On the Transverse Vibrations of Rectangular Isotropic Plates," *Journal of the Aeronautical Society of India,* Vol. 13, No. 4, Nov. 1961, pp. 111–117.

7 Carmichael, T. E., "The Vibration of a Rectangular Plate With Edges Elastically Restrained Against Rotation," *Quarterly Journal of Mechanics and Applied Mathematics,* Vol. 12, Pt. 1, 1959, pp. 29–42.

8 Joga-Rao, C. V., and Kantham, C. L., "Natural Frequencies of Rectangular Plates With Edges Elastically Restrained Against Rotation," *Journal of Aeronautical Sciences,* Vol. 24, No. 4, Nov. 1957, pp. 855–856.

9 Chulay, S. J., "Vibration of Elastically Restrained Rectangular Plates," MS Thesis, University of Wisconsin, 1957.

10 Stokey, W. F., Zorowski, C. F., and Appl, F. C., "Prevention of Mechanical Vibrations in Electronic Chassis-Design Manual," Report to Rome Air Development Center, Contract AF30(602)-913, Sept. 1955.

11 Hoppmann, W. H., II, and Greenspon, J., "An Experimental Device for Obtaining Elastic Rotational Constraints on the Boundary of a Plate," *Proceedings of the U.S. National Congress of Applied Mechanics,* 1954, pp. 187–191.

12 Laura, P. A. A., and Romanelli, E., "Vibrations of Rectangular Plates Elastically Restrained Against Rotation Along All Edges and Subjected to a Biaxial State of Stress," *Journal of Sound and Vibration,* Vol. 37, No. 3, 1974, pp. 367–377.

13 Laura, P. A. A., Luisoni, L. E., and Filipich, C., "A Note on the Determination of the Fundamental Frequency of Vibration of Thin, Rectangular Plates With Edges Possessing Different Rotational Flexibility Coefficients," *Journal of Sound and Vibration,* Vol. 55, No. 3, 1977, pp. 327–333.

14 Laura, P. A. A., Luisoni, L. E., and Ficcadenti, G., "On the Effect of Different Edge Flexibility Coefficients on Transverse Vibrations of Thin, Rectangular Plates," *Journal of Sound and Vibration,* Vol. 57, No. 3, 1978, pp. 333–340.

15 Laura, P. A. A., and Grossi, R., "Transverse Vibration of a Rectangular Plate Elastically Restrained Against Rotation Along Three Edges and Free on the Fourth Edge," *Journal of Sound and Vibration,* Vol. 59, No. 3, 1978, pp. 355–368.

16 Laura, P. A. A., Grossi, R. O., and Soni, S. R., "Free Vibrations of a Rectangular Plate of Variable Thickness Elastically Restrained Against Rotation Along Three Edges and Free on the Fourth Edge," *Journal of Sound and Vibration,* Vol. 62, No. 4, 1979, pp. 493–503.

17 Laura, P. A. A., and Luisoni, L. E., "Vibrations of Orthotropic Rectangular Plates With Edges Possessing Different Rotational Flexibility and Subjected to In-Plane Forces," *Computers Structures,* Vol. 9, No. 6, 1978, pp. 527–532.

18 Egle, D. M., "The Resonant Response of a Rectangular Plate With an Elastic Edge Restraint," ASME *Journal of Engineering for Industry,* Vol. 94, No. 2, 1972, pp. 517–525.

19 Hanawa, T., et al., "Vibration of an Elastically Supported Cantilever Plate," *Proceedings of the 16th Japan National Congress on Applied Mechanics, Tokyo, October 19–20, 1966, 1967, pp. 284–290.*

20 Nassar, E. E. M., "On the Dynamic Characteristics of Beams, Plates, and Shells," PhD Thesis, Georgia Institute of Technology, 1973.

21 Leissa, A. W., Laura, P. A. A., and Gutierrez, R. H., "Transverse Vibrations of Circular Plates Having Nonuniform Edge Constraints," *Journal of the Acoustical Society of America,* Vol. 66, No. 1, 1979, pp. 180–184.

22 Narita, Y., and Leissa, A. W., "Transverse Vibration of Simply Supported Circular Plates Having Partial Elastic Constraints," *Journal of Sound and Vibration,* Vol. 70, No. 1, 1980, pp. 103–116.

23 Narita, Y., and Leissa, A. W., "Flexural Vibrations of Free Circular Plates Elastically Constrained Along Parts of the Edge," *International Journal of Solids and Structures* (forthcoming).

24 Ritz, W., "Theorie der Transversalschwingungen, einer quadratischen Platte mit freien Rändern," *Ann. Physik,* Bd. 28, 1909, pp. 737–786.

25 Voigt, W., "Bemerkungen zu dem Problem der transversalen Schwingungen rechteckiger Platten," *Nachr. Ges. Wiss.,* Göttingen, No. 6, 1893, pp. 225–230.

26 Leissa, A. W., "The Free Vibration of Rectangular Plates," *Journal of Sound and Vibration,* Vol. 31, No. 3, 1973, pp. 257–293.

27 Iguchi, S., "Die Eigenwertprobleme für die elastische rechteckige Platte," Mem. Fac. Eng., Hokkaido University, 1938, pp. 305–372.

28 Leissa, A. W., Laura, P. A. A., and Gutierrez, R. H., "Vibraciones de Placas Rectangulares con Condiciones Elasticas, no Uniformes en los Bordes," Instituto de Mécanica Applica, Publicacion IMA No. 80-18, 1980.

T. Irie
Professor.

G. Yamada
Associate Professor.

S. Aomura
Graduate Student.

Department of Mechanical Engineering,
Hokkaido University,
North-13, West-8,
Sapporo 060, Japan

The Steady-State Response of a Rotating Damped Disk of Variable Thickness

The stress distribution and steady-state response of a rotating damped annular disk of variable thickness are determined by means of the matrix method. The equation of equilibrium and the equations for the flexural vibration of the rotating disk are written as a respective coupled set of first-order differential equations by use of the matrices of the disk. The elements of the matrices are calculated by numerical integration of the equations, and the stress components and the driving-point impedance and force transmissibility of the disk are obtained by using these elements. The method is applied to free-clamped rotating disks with linearly, exponentially, and hyperbolically varying thickness driven by a harmonic force at the free outer edge, and the effects of the angular velocity and the variable thickness are studied.

Introduction

In the past, failure of rotating disk wheels due to flexural vibration has frequently occurred in rotodynamic machinery such as steam turbines and gas turbines. In 1924, Campbell [1] studied vibration problems of steam turbine disk wheels and recommended some protective measures. Lamb and Southwell [2], and many other authors [3–6] analyzed the flexural vibrations of rotating uniform disks theoretically. For rotating nonuniform disks, Ehrich [7] analyzed the vibration modes by a matrix method, and Kirkhope and Wilson [8], and Kennedy and Gorman [9] studied nonuniform disks subjected to centrifugal and thermal stresses by the finite-element method. The steady-state response of damped circular plates has been extensively studied by Snowdon [10, 11]. However, all of these studies have been based upon the classical plate theory in which the rotatory inertia and shear deformation of the plate are not taken into account.

This paper applies the matrix method to an analysis of the stress distribution and the steady-state vibration of a rotating damped disk of variable thickness in response to a sinusoidally varying force. The Mindlin theory [12, 13] is adopted for the analysis of the vibration of the disk, in which both of the rotatory inertia and shear deformation are taken into consideration. The radial displacement and the centrifugal stress are written in a matrix differential equation and the equations of flexural vibration of the rotating disk are expressed as a coupled set of first-order differential equations by use of the transfer matrix. The elements of the matrices are determined by numerical integration of the equations. From the matrices thus obtained and the boundary conditions, the centrifugal stress components, the driving-point impedance, and force transmissibility are determined numerically for free-clamped disks excited along the free outer edge. In this paper, Young's modulus and the shear modulus of internally damped disks are assumed to be complex quantities. This assumption has been justified by the results of experimental measurement [14].

By the application of the present method, the centrifugal stress distribution and the steady-state response of rotating annular disks of linearly, exponentially, and hyperbolically varying thickness are calculated numerically, and the effects of the angular velocity and the variable thickness on the vibration of the disks are discussed.

Centrifugal Stress Distribution of a Rotating Annular Disk

When an annular disk rotates at constant velocity Ω, the equation of equilibrium of the forces in radial direction can be expressed as

$$\frac{d}{dr}(hr\sigma_r^*) - h\sigma_\theta^* + \rho\Omega^2 hr^2 = 0 \tag{1}$$

where ρ is the mass per unit volume, h is the thickness, and the polar coordinates (r, θ) are taken in the neutral surface of the disk. The radial and circumferential stress components are given by

$$\sigma_r^* = \frac{E}{1-\nu^2}\left(\frac{du^*}{dr} + \nu\frac{u^*}{r}\right)$$
$$\sigma_\theta^* = \frac{E}{1-\nu^2}\left(\nu\frac{du^*}{dr} + \frac{u^*}{r}\right) \tag{2}$$

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, April 1980.

in terms of radial displacement u^*, where E is Young's modulus, and ν is Poisson's ratio. When the inner edge ($r = a$) is clamped and the outer edge ($r = b$) is free, the boundary conditions are written as

$$u^* = 0 \quad \text{at} \quad r = a$$
$$\sigma_r^* = 0 \quad \text{at} \quad r = b \tag{3}$$

Upon eliminating σ_θ^* from (1) and (2), the following matrix differential equation is derived:

$$\frac{d}{d\eta}\begin{Bmatrix} u \\ \sigma_r \end{Bmatrix} = \begin{bmatrix} -\dfrac{\nu}{\eta} & q_0 \\ \dfrac{1}{q_0}\dfrac{1-\nu^2}{\eta^2} - \dfrac{1-\nu}{\eta} & -\dfrac{d'}{d} \end{bmatrix}\begin{Bmatrix} u \\ \sigma_r \end{Bmatrix} - \begin{Bmatrix} 0 \\ \Lambda^2\eta \end{Bmatrix} \tag{4}$$

where the symbol $'$ denotes differentiation with respect to η, and where the following dimensionless variables are introduced:

$$\eta = \frac{r}{b} \quad \left(\beta = \frac{a}{b}\right) \quad d = \frac{h}{h_0}$$
$$q_0 = \frac{1}{12}\left(\frac{h_0}{b}\right)^2, \quad \Lambda = \sqrt{\frac{\rho h_0 b^4}{D_0}}\,\Omega \quad \left[D_0 = \frac{Eh_0^3}{12(1-\nu^2)}\right] \tag{5}$$

and

$$u = \frac{u^*}{b}, \quad (\sigma_r, \sigma_\theta) = \frac{bh_0^2}{D_0}(\sigma_r^*, \sigma_\theta^*) \tag{6}$$

where h_0 is the thickness at the inner edge.

Since the exact solution of (4) cannot usually be obtained in a convenient form for a nonuniform disk, the equation has to be integrated numerically on $[\beta, 1]$ with $u = 0$ and an appropriately assumed starting value $\sigma_r(\beta)$ for the radial stress at the clamped inner edge. The numerical integration should be repeated until the radial stress σ_r (1) becomes zero in value at the free outer edge.

The equations governing axisymmetrical vibrations of a rotating disk subjected to centrifugal stresses can be written as

$$\frac{\partial Q_r^*}{\partial r} + \frac{Q_r^*}{r} + \frac{1}{r}\frac{\partial}{\partial r}\left(rh\sigma_r^*\frac{\partial W^*}{\partial r}\right) + C_1\frac{\partial W^*}{\partial t} = \rho h\frac{\partial^2 W^*}{\partial t^2}$$
$$\frac{\partial M_r^*}{\partial r} + \frac{M_r^* - M_\theta^*}{r} - Q_r^* + C_2\frac{\partial \psi_r^*}{\partial t} = \frac{1}{12}\rho h^3\frac{\partial^2 \psi_r^*}{\partial t^2} \tag{7}$$

where C_1 and C_2 are the viscous damping coefficients with respect to the translational and rotational motion of the disk, respectively. The components of the moment and shearing force are given by

$$M_r^* = \tilde{D}\left(\frac{\partial \psi_r^*}{\partial r} + \nu\frac{\psi_r^*}{r}\right), \quad M_\theta^* = \tilde{D}\left(\nu\frac{\partial \psi_r^*}{\partial r} + \frac{\psi_r^*}{r}\right)$$
$$Q_r^* = K\tilde{G}h\left(\psi_r^* + \frac{\partial W^*}{\partial r}\right) \tag{8}$$

in terms of the transverse deflection W^* and the angular rotation ψ_r^* of the normal to the neutral surface in radial direction. The quantity K is the shear coefficients which assumes usually the value $\pi^2/12$, and $\tilde{D}$ is the flexural rigidity of internally damped disk expressed by

$$\tilde{D} = \frac{\tilde{E}h^3}{12(1-\nu^2)} \tag{9}$$

Young's modulus and the shear modulus of the internally damped disk are considered to be the complex quantities

$$\tilde{E} = E(1 + j\delta_E) \quad \tilde{G} = G(1 + j\delta_G) \tag{10}$$

where E and G express the real parts of $\tilde{E}$ and $\tilde{G}$, respectively, and δ_E and δ_G are constants representing the ratios of the imaginary to the real parts of them at any frequencies ω. The steady-state deflection, slope, bending moment, and shearing force of the disk driven by a sinusoidally varying force F^* at the outer edge are written as

$$(M_r^*, M_\theta^*) = \frac{D_0}{b}(M_r, M_\theta)e^{j\omega t}$$
$$Q_r^* = \frac{D_0}{b^2}Q_r e^{j\omega t} \tag{11}$$

$$\psi_r^* = \psi_r e^{j\omega t}; \quad W^* = bWe^{j\omega t}$$
$$F^* = \frac{D_0}{b^2}Fe^{j\omega t} \tag{11}$$

(*Cont.*)

where

$$\lambda = \sqrt{\frac{\rho h_0 b^4}{D_0}}\,\omega, \quad \tau = \sqrt{\frac{D_0}{\rho h_0 b^4}}\,t \tag{12}$$

and the variables M_r, M_θ, Q_r, ... without the asterisk * are the respective dimensionless quantities. Upon eliminating M_θ^* from (7) and (8), the Mindlin equations can be written as a matrix differential equation as follows:

$$\frac{d}{d\eta}\{Z(\eta)\} = [U(\eta)]\{Z(\eta)\} \tag{13}$$

by using the state vector $\{Z(\eta)\} = \{M_r Q_r \psi_r W\}^T$ and the coefficient matrix $[U(\eta)]$ with the elements

$$U_{11} = -\frac{1-\nu}{\eta}, \quad U_{12} = 1,$$
$$U_{13} = \left\{(1 + j\delta_E)\frac{1-\nu^2}{\eta^2} - q_0\lambda^2\right\}d^3 - 2j\zeta_2\lambda, \quad U_{14} = 0,$$
$$U_{21} = \frac{\sigma_r(1 + j\delta_G)}{(1 + j\delta_E)(1 + j\delta_G + k_0\sigma_r)d^2},$$
$$U_{22} = -\frac{1}{1 + j\delta_G + k_0\sigma_r}\left(k_0\sigma_r' + k_0\frac{\sigma_r}{\eta} + \frac{1 + j\delta_G}{\eta}\right),$$
$$U_{23} = \frac{(1 + j\delta_G)d}{1 + j\delta_G + k_0\sigma_r}\left\{\sigma_r' + \left(\frac{1-\nu}{\eta} + \frac{d'}{d}\right)\sigma_r\right\},$$
$$U_{24} = -\frac{(\lambda^2 d + 2j\lambda\zeta_1)(1 + j\delta_G)}{1 + j\delta_G + k_0\sigma_r},$$
$$U_{31} = \frac{1}{(1 + j\delta_E)d^3}, \quad U_{32} = 0, \quad U_{33} = -\frac{\nu}{\eta}, \quad U_{34} = 0,$$
$$U_{41} = 0, \quad U_{42} = \frac{k_0}{(1 + j\delta_G)d}, \quad U_{43} = -1, \quad U_{44} = 0 \tag{14}$$

where ζ_1, ζ_2, and k_0 are the dimensionless parameters expressed as

$$\zeta_1 = \frac{b^2 C_1}{2\sqrt{\rho h_0 D_0}}, \quad \zeta_2 = \frac{C_2}{2\sqrt{\rho h_0 D_0}}$$
$$k_0 = \frac{2q_0}{K(1-\nu)} \tag{15}$$

Analysis by Use of the Transfer Matrix

Since the analytical solution of (13) cannot be obtained for rotating disk of variable thickness, the transfer matrix approach is adopted here. In general, the state vector $\{Z(\eta)\}$ is expressed as

$$\{Z(\eta)\} = [T(\eta)]\{Z(\beta)\} \tag{16}$$

by using the transfer matrix $[T(\eta)]$ in radial direction. From (13) and (16), the following equation is derived:

$$\frac{d}{d\eta}[T(\eta)] = [U(\eta)][T(\eta)] \tag{17}$$

For nonuniform disk, the matrix $[T(\eta)]$ is obtained by integrating (17) numerically. Since the matrix $[U(\eta)]$ and also the transfer matrix $[T(\eta)]$ are complex quantities, $[T(\eta)]$ should be calculated numerically in practice by the equation

$$\frac{d}{d\eta}\begin{bmatrix} T_R(\eta) \\ T_I(\eta) \end{bmatrix} = \begin{bmatrix} U_R(\eta) & U_I(\eta) \\ -U_I(\eta) & U_R(\eta) \end{bmatrix}\begin{bmatrix} T_R(\eta) \\ T_I(\eta) \end{bmatrix} \tag{18}$$

which is obtained by dividing the matrices $[U(\eta)]$ and $[T(\eta)]$ into the real and imaginary parts, respectively. The matrices $[T_R(\eta)]$ and $[T_I(\eta)]$ are obtained by integrating (18) numerically on $[\beta, \eta]$ under the initial value.

$$\begin{bmatrix} T_R(\beta) \\ T_I(\beta) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad ([1]\text{: unit matrix}) \tag{19}$$

In the case of a free-clamped annular disk, the boundary conditions are written as

$$\begin{aligned} \psi_r = W = 0 \quad &\text{at} \quad \eta = \beta \\ M_r = 0, \quad Q_r = F \quad &\text{at} \quad \eta = 1 \end{aligned} \tag{20}$$

by which (16) is written as

$$\begin{Bmatrix} 0 \\ F \\ \psi_r \\ W \end{Bmatrix}_{(1)} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \\ T_{31} & T_{32} \\ T_{41} & T_{42} \end{bmatrix}_{(1)} \begin{Bmatrix} M_r \\ Q_r \end{Bmatrix}_{(\beta)} \tag{21}$$

with only the elements of $[T(\eta)]$ necessary for the calculation. By solving (21), unknown quantities $\{M_r Q_r\}^T{}_{(\beta)}$ and $\{\psi_r W\}^T{}_{(1)}$ at both edges are obtained in the following form:

$$\begin{Bmatrix} M_r \\ Q_r \end{Bmatrix}_{(\beta)} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}_{(1)}^{-1} \begin{Bmatrix} 0 \\ F \end{Bmatrix} \tag{22}$$

and

$$\begin{Bmatrix} \psi_r \\ W \end{Bmatrix}_{(1)} = \begin{bmatrix} T_{31} & T_{32} \\ T_{41} & T_{42} \end{bmatrix}_{(1)} \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}_{(1)}^{-1} \begin{Bmatrix} 0 \\ F \end{Bmatrix} \tag{23}$$

The steady-state responses of the disk are determined by (16), (22), and (23). The normalized driving-point impedance at the free outer edge is determined by

$$\frac{Z(b)}{j\omega M} = \frac{F}{\lambda^2 W(1) \int_\beta^1 d \cdot \bar{\eta}\, d\bar{\eta}} \tag{24}$$

where M is the mass of the whole disk. The transfer impedance $Z(\eta)$ between the driving-point (the outer edge) and a concentric circle of any radius is given by replacing $W(1)$ of (24) with $W(\eta)$. The force transmissibility of the disk at the clamped edge is expressed in the simple form

$$T_F = \left| \frac{\beta Q_r(\beta)}{F} \right| \tag{25}$$

Numerical Calculations and Discussions

In this section, the present method is applied to free-clamped annular disks of variable thickness of the following three profiles:

An Annular Disk of Linearly Varying Thickness.

$$h = h_0 \left\{ 1 - \left(1 - \frac{h_1}{h_0} \right) \left(\frac{r-a}{b-a} \right) \right\} \tag{26}$$

An Annular Disk of Exponentially Varying Thickness.

$$h = h_0 (h_1/h_0)^{(r-a)/(b-a)} \tag{27}$$

An Annular Disk of Hyperbolically Varying Thickness.

$$h = h_0 (r/a)^{-\log_\beta (h_1/h_0)} \tag{28}$$

and the axisymmetrical responses to a harmonic force are calculated numerically. Here, h_1 is the thickness at the outer edge. In the numerical calculations reported here, the transfer matrix is obtained by integrating (18) numerically by the Runge-Kutta-Gill method with step size 1/50, which assures sufficient accuracy in practice. For metals and nonmetals, the assumption that δ_E is nearly equal to δ_G in value through a broad frequency range is well justified from the results of experimental measurement [14]. Though the dynamic moduli and damping factors depend upon the frequencies for practical materials, it is assumed for the calculations presented here that $\delta_E = \delta_G =$ *constant* for all frequencies.

Figs. 1 shows the centrifugal stress distributions of free-clamped rotating disks of linearly and hyperbolically varying thickness. The disk of $h_1/h_0 = 1.0$ is a uniform disk. The stress components become

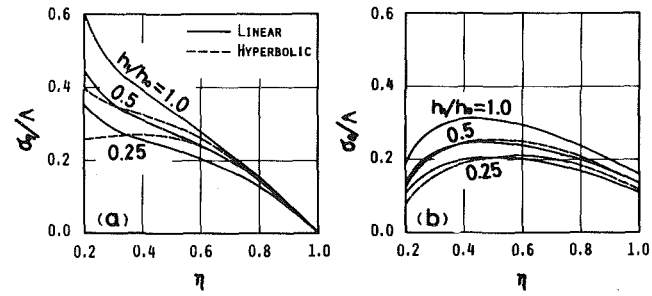


Fig. 1 Stress distribution of free-clamped rotating annular disks of variable thickness; (*a*) Radial stresses: (*b*) circumferential stresses; $\nu = 0.3$, $\beta = 0.2$

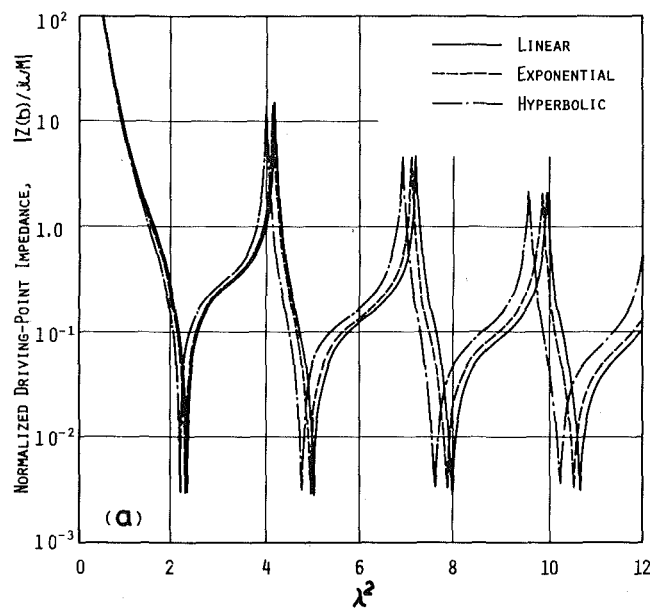


Fig. 2(*a*)

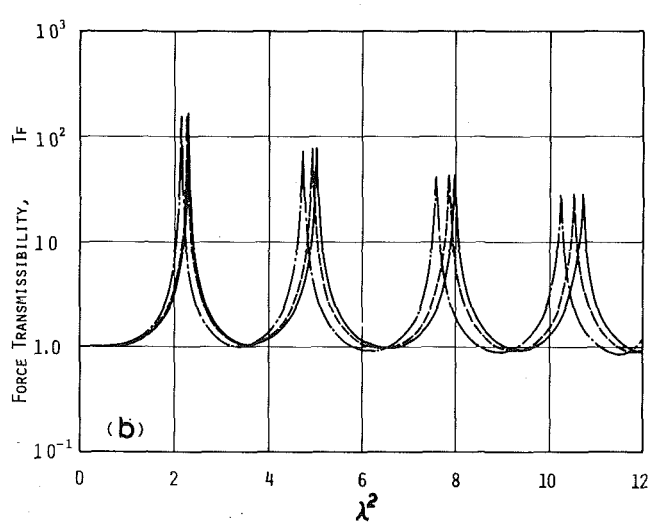


Fig. 2(*b*)

Fig. 2 Steady-state response of free-clamped rotating disks of variable thickness driven by a harmonic force at the free outer edge; (*a*) Normalized driving-point impedance: (*b*) force transmissibility; $\nu = 0.3$, $\delta_E = \delta_G = 0.01$, $\beta = 0.2$, $h_1/h_0 = 0.5$, $\Lambda^2 = 1.0$

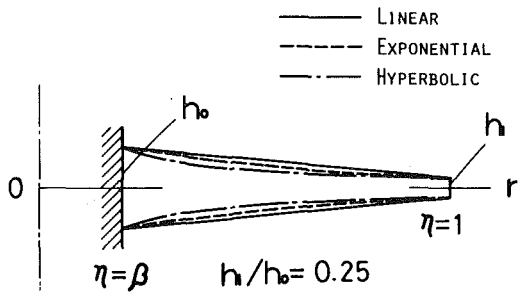


Fig. 3 Profiles of annular disks

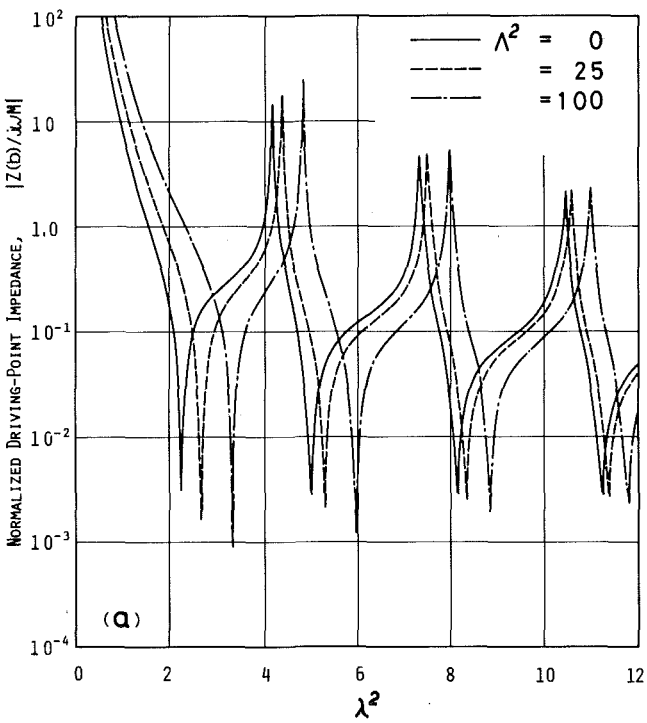


Fig. 4(a)

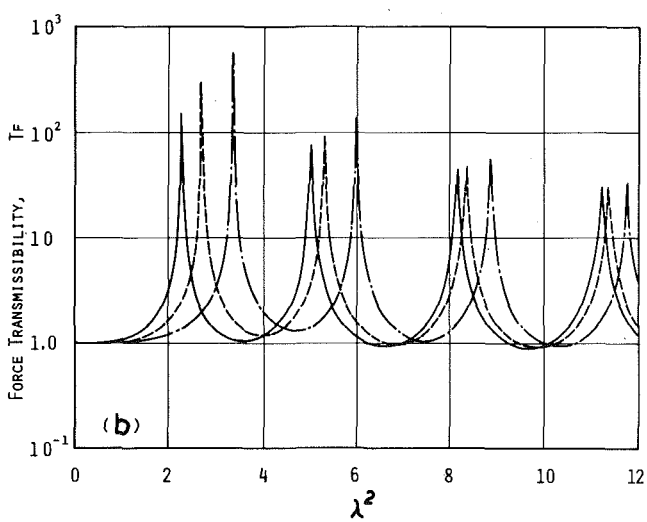


Fig. 4(b)

Fig. 4 Steady-state response of free-clamped rotating disks of exponentially varying thickness driven by a harmonic force at the free outer edge; (a) Normalized driving-point impedance: (b) force transmissibility; $\nu = 0.3$, $\delta_E = \delta_G = 0.01$, $\beta = 0.2$, $h_0/a = 0.05$, $h_1/h_0 = 0.5$

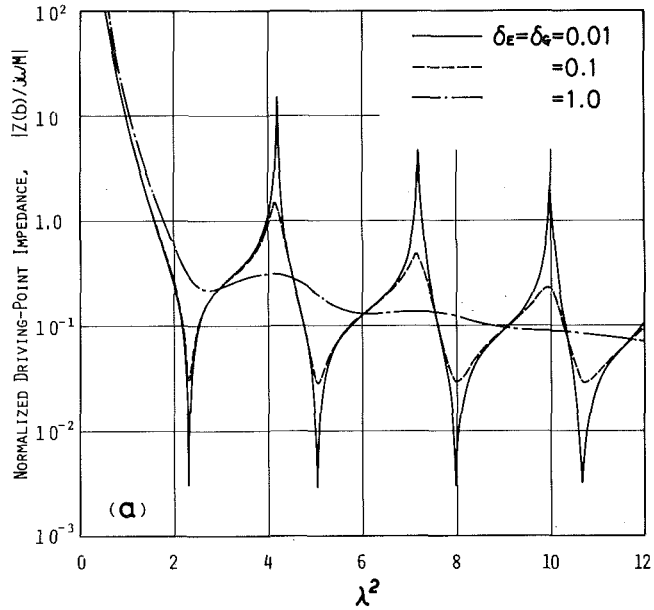


Fig. 5(a)

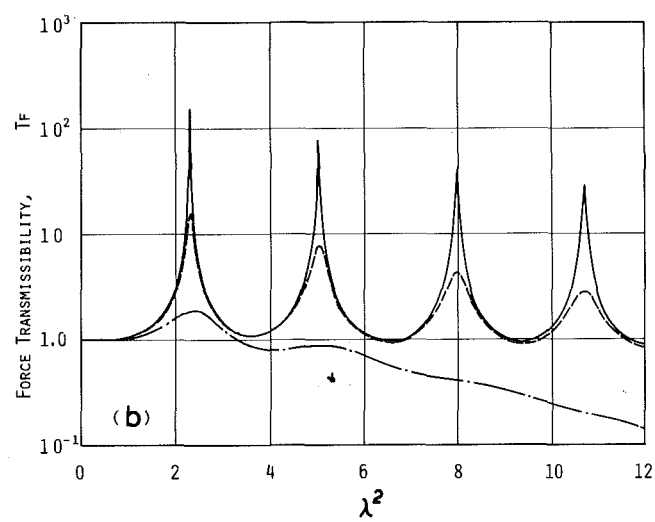


Fig. 5(b)

Fig. 5 Steady-state response of free-clamped rotating disks of linearly tapered thickness driven by a harmonic force at the free outer edge; (a) Normalized driving-point impedance: (b) force transmissibility; $\nu = 0.3$, $\beta = 0.2$, $h_0/a = 0.1$, $h_1/h_0 = 0.5$, $\Lambda^2 = 1.0$

smaller, with a decrease of the ratio h_1/h_0. As shown in Fig. 1(b), the maximum circumferential stress arises inside the disk and with a decrease of the ratio h_1/h_0, the location of the maximum stress shifts outward. The stress components of the disk of exponentially varying thickness are close to those of the disk of linearly varying thickness.

Figs. 2(a) and (b) show the normalized driving-point impedance and force transmissibility, respectively, of free-clamped rotating disks of linearly, exponentially, and hyperbolically varying thickness which are driven by a harmonic force at the free outer edge. The distinctive resonant and antiresonant peaks appear on these response curves, among which the antiresonant peaks vanish on the transmissibility curves. Though the magnitude of response curves is not affected so much by the nature of the function expressing the profile variation, the resonant and antiresonant frequencies become smaller in that

Table 1 Resonant and antiresonant frequencies of free-clamped undamped disks of variable thickness; $\nu = 0.3$, $\beta = 0.2$, $h_1/h_0 = 0.5$

(a) $\Omega^2 = 0$

Profile	h_0/b = 0		0.1		0.2	
	$\lambda_{res.}$	$\lambda_{antir.}$	$\lambda_{res.}$	$\lambda_{antir.}$	$\lambda_{res.}$	$\lambda_{antir.}$
Linear	2.295	4.248	2.279	4.150	2.233	3.913
	5.166	7.546	5.011	7.160	4.661	6.428
	8.215	10.888	7.963	9.974	7.084	8.606
	11.781	14.237	10.697	12.552	9.160	10.496
Exponential	2.256	4.183	2.241	4.093	2.198	3.872
	5.056	7.415	4.916	7.059	4.591	6.367
	8.284	10.686	7.828	9.838	6.996	8.532
	11.541	13.966	10.537	12.397	9.068	10.420
Hyperbolical	2.124	4.037	2.113	3.962	2.081	3.774
	4.829	7.147	4.714	6.850	4.438	6.243
	8.082	10.274	7.564	9.560	6.831	8.385
	11.066	13.408	10.220	12.073	8.893	10.263

(b) $\Omega^2 = 25$

Profile	h_0/b = 0		0.1		0.2	
	$\lambda_{res.}$	$\lambda_{antir.}$	$\lambda_{res.}$	$\lambda_{antir.}$	$\lambda_{res.}$	$\lambda_{antir.}$
Linear	2.679	4.427	2.654	4.319	2.588	4.060
	5.431	7.712	5.271	7.316	4.915	6.575
	8.630	11.017	8.127	10.094	7.243	8.721
	11.910	14.342	10.818	12.649	9.278	10.587
Exponential	2.691	4.385	2.668	4.285	2.607	4.041
	5.344	7.594	5.198	7.228	4.865	6.525
	8.469	10.825	8.005	9.969	7.167	8.656
	11.680	14.079	10.667	12.501	9.195	10.518
Hyperbolically	2.644	4.287	2.626	4.201	2.574	3.989
	5.174	7.362	5.052	7.053	4.765	6.431
	8.165	10.442	7.776	9.717	7.032	8.531
	11.233	13.544	10.376	12.198	9.042	10.380

order for disks of linearly, exponentially, and hyperbolically varying thickness. The reason for this is that the disk of hyperbolically varying thickness is the thinnest, while the disk of linearly tapered thickness is the thickest among disks of the specified ratio h_1/h_0, as seen in Fig. 3, and therefore the effects of an increase of the flexural rigidity is larger than that of the mass.

Figs. 4 show the response curves of free-clamped disks of exponentially varying thickness, where the angular velocity is assigned as a parameter. The magnitude of the driving-point impedance and force transmissibility of lower frequencies is affected to some extent by the angular velocity. The resonant and antiresonant frequencies become larger with an increase of the velocity.

Table 1 shows the resonant and antiresonant frequencies of rotating undamped disks of variable thickness. The frequencies of the disks of $h_0/b = 0$ present those obtained by the classical theory. With an increase of the ratio h_0/b, the frequencies become smaller by the effect of the rotatory inertia and shear deformation of the disk.

Figs. 5 show the response curves of free-clamped disks of linearly tapered thickness, where the internal damping factors δ_E and δ_G are taken as a parameter. For the disks with large damping factors, the distinctive resonant and antiresonant behaviors vanish and the resonant curves become flat through a broad frequency range. The force transmissibility become smaller than unity in value for the disk of $\delta_E = \delta_G = 1.0$, that indicates the possibility of vibration isolation. The effects of the external damping factors on the responses are essentially the same as those of the internal damping factors.

Fig. 6 shows the transverse deflections at the first three resonances and antiresonances of free-clamped undamped disks of hyperbolically varying thickness driven at the free outer edge. In general, the deflections of the disks are infinite at the resonance, and therefore these figures show the vibration modes without presenting the resonant

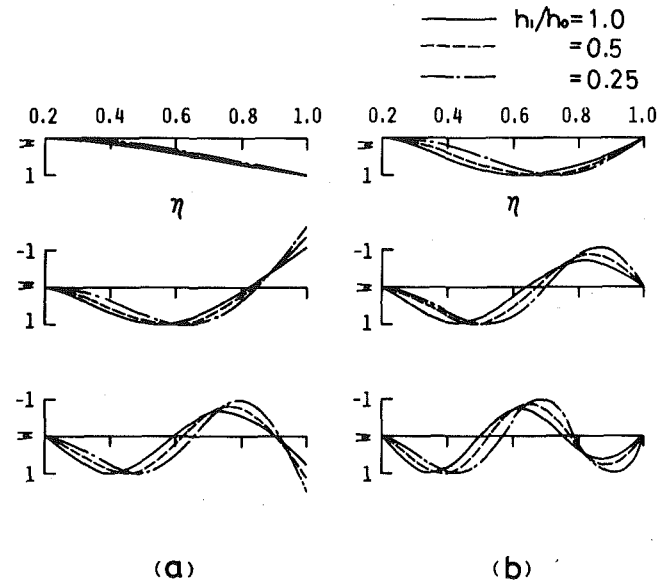


Fig. 6 Transverse deflections of rotating undamped disks of hyperbolically varying thickness; (*a*) Resonant modes: (*b*) antiresonant modes; $\nu = 0.3$, $\beta = 0.2$, $h_0/a = 0.1$, $\Lambda^2 = 25$

deflections. In Fig. 6, the deflections of the first mode are plotted to have a unit value at the free outer edge and those of other modes are also plotted as unity at the loop near the clamped edge where the maximum deflection arises. With a decrease of the ratio h_1/h_0, the deflections near the free edge become larger and the loops causing the maximum deflection shift to the free outer edge. The antiresonant deflections of these undamped disks are the same as those of the disks simply supported at the outer edge.

References

1 Campbell, W., "The Protection of Steam-Turbine Disk Wheels From Axial Vibration," TRANS. ASME, Vol. 46, 1924, pp. 31–160.

2 Lamb, H., and Southwell, R. V., "The Vibrations of a Spinning Disk," *Proceedings of the Royal Society,* London, Vol. 99, 1921, pp. 272–280.

3 Southwell, R. V., "On the Free Transverse Vibrations of a Uniform Circular Disc Clamped at Its Centre: and on the Effects of Rotation," *Proceedings of the Royal Society,* London, Vol. 101, 1922, pp. 133–153.

4 Eversman, W., and Dodson, R. D., Jr., "Free Vibration of a Centrally Clamped Spinning Circular Disk," *AIAA Journal,* Vol. 7, 1969, pp. 2010–2012.

5 Barasch, S., and Chen, Y., "On the Vibration of a Rotating Disk," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, 1972, pp. 1143–1144.

6 Loh, H. C., and Carney, J. F., III, "Vibration and Stability of Spinning Annular Plates Reinforced With Edge Beams," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 499–501.

7 Ehrich, F. F., "A Matrix Solution for the Vibration Modes of Nonuniform Disks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 23, 1956, pp. 109–115.

8 Kirkhope, J., and Wilson, G. J., "Vibration and Stress Analysis of Thin Rotating Discs Using Annular Finite Elements," *Journal of Sound and Vibration,* Vol. 44, 1976, pp. 461–474.

9 Kennedy, W., and Gorman, D., "Vibration Analysis of Variable Thickness Discs Subjected to Centrifugal and Thermal Stresses," *Journal of Sound and Vibration,* Vol. 53, 1977, pp. 83–101.

10 Snowdon, J. C., "Forced Vibration of Internally Damped Circular Plates With Supported and Free Boundaries," *Journal of the Acoustical Society of America,* Vol. 47, 1970, pp. 882–891.

11 Snowdon, J. C., "Forced Vibration of Internally Damped Circular and Annular Plates With Clamped Boundaries," *Journal of the Acoustical Society of America,* Vol. 50, 1971, pp. 846–858.

12 Mindlin, R. D., and Deresiewicz, H., "Thickness-Shear and Flexural Vibrations of a Circular Disk," *Journal of Applied Physics,* Vol. 25, 1954, pp. 1329–1332.

13 Deresiewicz, H., and Mindlin, R. D., "Axially Symmetric Flexural Vibrations of a Circular Disk," ASME JOURNAL OF APPLIED MECHANICS, Vol. 22, 1955, pp. 86–88.

14 Kerlin, R. L., and Snowdon, J. C., "Driving-Point Impedances of Cantilever Beams-Comparison of Measurement and Theory," *Journal of the Acoustical Society of America,* Vol. 47, 1970, pp. 220–228.

J. R. Hutchinson

Professor,
Department of Civil Engineering,
University of California,
Davis, Calif. 95616
Mem. ASME

Vibrations of Solid Cylinders

A series solution of the general three-dimensional equations of linear elasticity is developed and used to find accurate natural frequencies for the vibrations of solid elastic cylinders with traction-free surfaces. The series solution is found to converge to accurate frequencies with the use of very few terms. Results are given for height-to-diameter ratios from zero to two and a frequency parameter $\omega R/C_s$ from zero to five and for modes of circumferential order from zero to four. Comparisons of these analytical results with previous experimental results shows excellent agreement.

Introduction

A three-dimensional series solution of the vibrations of a solid isotropic elastic circular cylinder with traction-free surfaces is developed, and the natural frequencies are evaluated over a representative range of the parameters. The solution includes both axisymmetric and nonaxisymmetric behavior and is evaluated over and beyond the full range of parameters investigated experimentally by McMahon [1].

The problem of the vibrating cylinder (or rod) was first investigated in terms of the general elastic equations by Pochhammer in 1876 and independently by Chree in 1889. An account of this treatment can be found in Love [2]. The Pochhammer-Chree solution is for an infinitely long circular rod which is traction-free on its circumferential boundary, but the solution does not permit traction-free ends. Many authors have considered approximate solutions which hold when the cylindrical solid approaches either a thin disk or a slender rod. These approximate solutions are discussed at length in McMahon's [1] paper, and he compares the various approximate solutions to his experimental results.

The solution for the axisymmetric vibrations of free cylinders has been presented by several authors. In 1962 McNiven and Perry [3] presented a solution based on approximate equations which take into account the coupling between longitudinal, axial shear, and radial modes of propagation in a rod of infinite length. In 1971 Rumerman and Raynor [4] developed frequency spectrum using the Rayleigh-Ritz procedure with displacement functions corresponding to pure radial and axial modes of the infinite cylinder. In 1972 Hutchinson [5] presented a solution for this problem using the method used in this paper and in a 1967 paper [6] for a similar problem. Rumerman and Raynor compared their results with those of McNiven and Perry and with the experimental results of McMahon. Their comparison showed good correlation with the experimental results of McMahon but differ with the results of McNiven and Perry. The results of Hutchinson compare in all significant aspects with the results of Rumerman and Raynor. A detailed comparison of the axisymmetric solution with the thick plate theory developed by Mindlin [7] was made in 1979 by Hutchinson [8]. The axisymmetric case has therefore, been adequately treated in the literature and is included in this paper simply for completeness.

With the exception of approximate solutions which treat the cylinder as either a slender rod or a thin plate, the only work on an analytic solution for the nonaxisymmetric vibrations of finite-length free cylinders appears to be the 1975 paper by Rasband [9]. Rasband's work is an extension of Hutchinson's 1972 paper to the nonaxisymmetric case. Rasband's paper contains no numerical results; although, he mentions exploratory numerical computations in which a somewhat larger number of terms is required for equivalent accuracy to Hutchinson's results.

In this paper, an approach which is slightly different from Rasband's and closer to the method in reference [5] is developed. The final form for numerical computation is much more concise than Rasband's and requires no more terms in each series than in reference [5]. Further, numerical results over the range of parameters investigated experimentally by McMahon [1] show complete agreement with the experimental results.

The method of solution used in this paper has an interesting history in that it has apparently been rediscovered a number of times. Mathieu [10] first applied the method in 1890 to the in plane loading of rectangular plates. Taylor [11] applied the method in 1933 to the buckling of clamped rectangular plates. Tomotika [12] applied it in 1936 to the vibrations of clamped rectangular plates (he did reference Taylor's work). Timoshenko [13] in 1938 applied it to the problem of transverse loading on a clamped rectangular plate, and in 1944 Pickett [14] applied the method to axial compression of circular cylinders. The method has been referred to as an "exact" infinite series solution by some authors (see, e.g., reference [4]), others because of the truncation error prefer to classify it as an approximate solution of the type where the differential equation is identically satisfied and the boundary conditions approximated.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, November, 1979; final revision, May, 1980. Paper No. 80-WA/APM-20.

Table 1 Solutions of elasticity equations. $J_n(\ )$ denotes the nth-order Bessel function of the first kind. Primes denote differentiation with respect to the argument. $\psi_n(x) = xJ_{n-1}(x)/J_n(x) - (n+1)$. The dimensionless wave numbers α, β, and δ and the frequency parameter ω are related by $\alpha^2 + \delta^2 = \omega^2(1-2\nu)/2(1-\nu)$ and $\alpha^2 + \beta^2 = \omega^2$. Each column represents a solution form

	(1)	(2)	(3)
u	$2\alpha J_n'(\alpha r) \begin{Bmatrix}\cos \delta z\\ \sin \delta z\end{Bmatrix} \cos n\theta$	$-\alpha\beta J_n'(\alpha r) \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \cos n\theta$	$-\frac{n}{r} J_n(\alpha r) \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \cos n\theta$
v	$-\frac{2n}{r} J_n(\alpha r) \begin{Bmatrix}\cos \delta z\\ \sin \delta z\end{Bmatrix} \sin n\theta$	$\frac{\beta n}{r} J_n(\alpha r) \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \sin n\theta$	$\alpha J_n'(\alpha r) \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \sin n\theta$
w	$2\delta J_n(\alpha r) \begin{Bmatrix}-\sin \delta z\\ \cos \delta z\end{Bmatrix} \cos n\theta$	$\alpha^2 J_n(\alpha r) \begin{Bmatrix}-\sin \beta z\\ \cos \beta z\end{Bmatrix} \cos n\theta$	0
σ_r	$\frac{2J_n(\alpha r)}{r^2} \{(n^2-1)+\frac{r^2}{2}(2\delta^2-\omega^2)-\psi_n(\alpha r)\} \times \begin{Bmatrix}\cos \delta z\\ \sin \delta z\end{Bmatrix} \cos n\theta$	$-\frac{\beta}{r^2} J_n(\alpha r) \{(n^2-1)-(\alpha r)^2-\psi_n(\alpha r)\} \times \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \cos n\theta$	$-\frac{n}{r^2} J_n(\alpha r)\, \psi_n(\alpha r) \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \cos n\theta$
σ_z	$(2\alpha^2-\omega^2)\, J_n(\alpha r) \begin{Bmatrix}\cos \delta z\\ \sin \delta z\end{Bmatrix} \cos n\theta$	$-\alpha^2\beta J_n(\alpha r) \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \cos n\theta$	0
$\tau_{r\theta}$	$-\frac{4n}{r^2} J_n(\alpha r)\, \psi_n(\alpha r) \begin{Bmatrix}\cos \delta z\\ \sin \delta z\end{Bmatrix} \sin n\theta$	$\frac{2n\beta}{r^2} J_n(\alpha r)\, \psi_n(\alpha r) \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \sin n\theta$	$\frac{2}{r^2} J_n(\alpha r) \{(n^2-1)-\frac{(\alpha r)^2}{2}-\psi_n(\alpha r)\} \times \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix} \sin n\theta$
τ_{rz}	$4\alpha\delta J_n'(\alpha r) \begin{Bmatrix}-\sin \delta z\\ \cos \delta z\end{Bmatrix} \cos n\theta$	$\alpha(\alpha^2-\beta^2) J_n'(\alpha r) \begin{Bmatrix}-\sin \beta z\\ \cos \beta z\end{Bmatrix} \cos n\theta$	$\frac{n\beta}{r} J_n(\alpha r) \begin{Bmatrix}\sin \beta z\\ -\cos \beta z\end{Bmatrix} \cos n\theta$
$\tau_{\theta z}$	$-\frac{4\delta n}{r} J_n(\alpha r) \begin{Bmatrix}-\sin \delta z\\ \cos \delta z\end{Bmatrix} \sin n\theta$	$-(\alpha^2-\beta^2)\, \frac{n}{r} J_n(\alpha r) \begin{Bmatrix}-\sin \beta z\\ \cos \beta z\end{Bmatrix} \sin n\theta$	$\beta\alpha\, J_n'(\alpha r) \begin{Bmatrix}-\sin \beta z\\ \cos \beta z\end{Bmatrix} \sin n\theta$

The method of solution, as applied to this problem, involves combining exact solutions of the governing equations in three series which term by term satisfy three of the six boundary conditions. The remaining three boundary conditions are satisfied by orthogonalization on the boundaries. This leads to an eigenvalue matrix the size of which is equal to the number of terms in each series. Because many of the submatrices in the eigenvalue matrix are diagonal, it can be condensed to the size of the number of the terms in one series without performing any matrix inversions. The frequencies at which the determinant of this condensed matrix is zero are the natural frequencies of the cylinder. The solution converges as more terms in the series are chosen, and it is shown that a very few terms produce excellent results for the lower frequencies.

Formulation

All symbols used are dimensionless. The displacements u, v, and w and the coordinates r and z are made dimensionless by dividing by the radius. All stress quantities are made dimensionless by dividing by the shear modulus. The wave numbers α, β, and δ are made dimensionless by multiplying by the radius. The time-dependence is removed by assuming that all displacements and stresses vary sinusoidally in phase at the same frequency. The frequency ω is made dimensionless by multiplying by the radius and dividing by the shear wave velocity. The thickness parameter h is the half height of the cylinder divided by the radius (i.e., the thickness-to-diameter ratio).

The solution is formed from a sum of basic solutions which satisfy the governing differential equations. The basic solution forms were given by Love [2]. They are tabulated in Table 1 in a similar form to Rasband's [9] Table 1. The terms in braces in Table 1 represent two different solution forms. I will refer to the top term in the braces as the "even" solution and the bottom term as "odd" solution. This is an arbitrary choice in that for the upper term u, v, σ_r, σ_θ, σ_z, and $\tau_{r\theta}$ are even functions of z while w, τ_{rz}, and $\tau_{\theta z}$ are odd functions of z.

The basic solutions are grouped into three series in such a way that three of the boundary conditions are identically satisfied term by term. The remaining boundary conditions are then satisfied by orthogonalization on the boundary. To some extent the choice of which boundary conditions to satisfy identically and which to satisfy by orthogonalization is arbitrary. For the method to work, however, one cannot satisfy all of the boundary conditions along one boundary identically. Thus it is not possible to satisfy the three boundary conditions on the cylindrical surface identically and the end conditions by orthogonalization or vice versa.

The solution form for the radial displacement is

$$u = \sum A_i u_{Ai} + \sum B_j u_{Bj} + \sum C_k u_{Ck} \tag{1}$$

with similar expressions for the tangential displacement, v; the axial displacement, w; and the stresses, σ_r, σ_θ, σ_z, $\tau_{r\theta}$, $\tau_{\theta z}$, τ_{rz} with the same three series of constants A_i, B_j, and C_k.

The functions u_{A_i}, u_{B_j}, and u_{C_k} are shown as follows:

$$u_A = [D_1 2\delta J_n'(\delta r) - D_2\alpha\beta J_n'(\beta r)] \begin{Bmatrix}\cos \alpha z\\ \sin \alpha z\end{Bmatrix} \cos n\theta \tag{2}$$

$$u_B = \left[D_3 2\alpha \begin{Bmatrix}\cos \delta z\\ \sin \delta z\end{Bmatrix} - D_4\alpha\beta \begin{Bmatrix}\cos \beta z\\ \sin \beta z\end{Bmatrix}\right] J_n'(\alpha r) \cos n\theta \tag{3}$$

$$u_C = \left[-D_5 \frac{n}{r} J_n(\beta r) + D_6 2\delta J_i'(\delta r)\right] \begin{Bmatrix}\cos \alpha z\\ \sin \alpha z\end{Bmatrix} \cos n\theta \tag{4}$$

The indices i, j, and k have been omitted for simplicity. They are implied on the terms u, D, α, β, and δ in equations (2)–(4), respectively.

The first term in the A series (multiplied by D_1) is from column 1 of Table 1 with α exchanged with δ, and the second term in the A series is from column 2 with α exchanged with β. The first term in the B series is from column 1 and the second term from column 2. The first term in the C series is from column 3 with α exchanged with β, and the second term is from column 1 with α exchanged for δ. The other displacement and stress terms are formed in an identical manner.

The introduction of six constants D_1–D_6 in equations (2)–(4) is merely a convenience. Since the u terms are multiplied by arbitrary constants, A_i, B_j, and C_k only three constants are strictly necessary. The boundary conditions only lead to the relationship between D_1 and D_2, D_3 and D_4, and D_5 and D_6. The convenience comes in being able to arbitrarily choose convenient forms for, say, D_1, D_3, and D_5.

The constants D_1–D_6 and the values of α are chosen so that the

following boundary conditions will be identically satisfied:

$$\tau_{rz}(r, \theta, h) = 0 \tag{5}$$

$$\tau_{rz}(1, \theta, z) = 0 \tag{6}$$

$$\tau_{\theta z}(r, \theta, h) = 0 \tag{7}$$

For the A and C series α is chosen as

$$\alpha_{Ai} = \begin{Bmatrix} i\pi; & i = 0, 1, 2, \ldots \\ \dfrac{2i-1}{2}\pi; & i = 1, 2, 3, \ldots \end{Bmatrix} \tag{8}$$

where the upper form is for the "even" solution and the lower form for the "odd" solution. This choice makes the A and C terms in τ_{rz} and $\tau_{\theta z}$ identically zero at $z = h$.

For the B series, α is chosen as

$$\alpha_{Bj} = \text{zero's of } J_n'(\alpha_{Bj}) \tag{9}$$

This choice makes the B term in τ_{rz} identically zero at $r = 1$.

The constants D_1–D_6 are chosen so that the B term in τ_{rz} and $\tau_{\theta z}$ is identically zero at $z = h$, and the A and C terms in τ_{rz} are identically zero at $r = 1$. The constants D_1–D_6 are

$$D_1 = \beta_A(\beta_A{}^2 - \alpha_A{}^2)\, J_n'(\beta_A)/\cos(\alpha_A h)(\beta_A\delta_A)^n \tag{10}$$

$$D_2 = -4\alpha_A\delta_A\, J_n'(\delta_A)/\cos(\alpha_A h)(\beta_A\delta_A)^n \tag{11}$$

$$D_3 = (\alpha_B{}^2 - \beta_B{}^2)\begin{Bmatrix}\sin(\beta_B h)/\beta_B \\ \cos(\beta_B h)/\delta_B\end{Bmatrix}/J_n(\alpha_B) \tag{12}$$

$$D_4 = -4\begin{Bmatrix}\delta_B \sin(\delta_B h)/\beta_B \\ \cos(\delta_B h)\end{Bmatrix}/J_n(\alpha_B) \tag{13}$$

$$D_5 = 4\delta_A J_n'(\delta_A)/\cos(\alpha_A h)(\beta_A\delta_A)^n \tag{14}$$

$$D_6 = n\, J_n(\beta_A)\cos(\alpha_A h)(\beta_A\delta_A)^n \tag{15}$$

The three remaining boundary conditions to be satisfied by orthogonality are

$$\sigma_r(1, \theta, z) = 0 \tag{16}$$

$$\sigma_z(r, \theta, h) = 0 \tag{17}$$

$$\tau_{r\theta}(1, \theta, z) = 0 \tag{18}$$

Equation (16) is approximately satisfied by setting

$$\int_0^h \sigma_r(1, \theta, z)\begin{Bmatrix}\cos\alpha_{Ai}z \\ \sin\alpha_{Ai}z\end{Bmatrix} dz = 0 \tag{19}$$

where α_{Ai} is defined by equation (8). Performing this integration leads to the following form:

$$a_iA_i + \sum_{j=1}^{NR} b_{ij}B_j + c_iC_i = 0; \quad i = 1, 2, 3, \ldots NZ \tag{20}$$

where NR and NZ are the number of terms chosen in the series in the r and z-directions, respectively. The coefficients a_i, b_{ij}, and c_i are as follows:

$$a_i = 2h \cdot JP(\beta_A) \cdot JB(\delta_A) \cdot (\beta_A{}^2 - \alpha_A{}^2) \cdot [(n^2 - 1) + (2\alpha_A{}^2 - \omega^2)/2 - \psi_n(\delta_A)] + 4h \cdot JP(\delta_A) \cdot JB(\beta_A) \cdot \alpha_A{}^2 \cdot [(n^2 - 1) - \beta_A{}^2 - \psi_n(\beta_A)] \tag{21}$$

$$b_{ij} = 4[(\alpha_B{}^2 - \beta_B{}^2)\,[n^2 + 2(\delta_B{}^2 - \omega^2)/2]/(\delta_B{}^2 - \alpha_A{}^2) + 2(n^2 - \alpha_B{}^2) \cdot \beta_B{}^2/(\beta_B{}^2 - \alpha_A{}^2)]\begin{Bmatrix}(\delta_B h)^2\, SN(\delta_B h)\, SN(\beta_B h) \\ -\cos(\beta_B h)\cos(\beta_B h)\end{Bmatrix} \tag{22}$$

$$c_i = 2hn \cdot JB(\beta_A)\,[-2JP(\delta_A)\psi_n(\beta_A) + JB(\delta_A)\,[(n^2 - 1) + (2\alpha_A{}^2 - \omega^2)/2 - \psi_n(\delta_A)]] \tag{23}$$

where the functions JB, JP, and SN are defined as follows:

$$JB(x) = J_n(x)/x^n \tag{24}$$

$$JP(x) = J_n'(x)/x^{n-1} \tag{25}$$

$$SN(x) = \sin(x)/x \tag{26}$$

where x is a dummy variable. The reason for these definitions is that for x imaginary the functions are still real-valued. That is

$$JB(ix) = I_n(x)/x^n \tag{27}$$

$$JP(ix) = I_n'(x)/x^{n-1} \tag{28}$$

$$SN(ix) = \sinh(x)/x \tag{29}$$

where i is the modified Bessel function of the first kind. Further

$$JB(0) = 1/(2^n n!) \tag{30}$$

$$JP(0) = 1/[2^n(n-1)!] \tag{31}$$

$$SN(0) = 1 \tag{32}$$

The necessity for these definitions becomes evident from the relationship of the dimensionless wave numbers, α, β, and δ with the dimensionless frequency ω. From Table 1, that relationship is

$$\alpha^2 + \beta^2 = \omega^2 \tag{33}$$

$$\alpha^2 + \delta^2 = \omega^2(1 - 2\nu)/2(1 - \nu) \tag{34}$$

The values α are chosen (equations (8) or (9) and are positive-real; however, the β and δ may be real or imaginary depending on the value of ω. Note that β and δ are the arguments of the functions JB, JP, and SN.

In equations (21)–(23) the constants α_A, β_A, and δ_A have an implied subscript, i, and the constants α_B, β_B, and δ_B have an implied subscript, j. The same notation will be true for the subsequent equations.

Equation (17) is approximately satisfied by setting

$$\int_0^1 \sigma_z(r, \theta, h) r J_n(\alpha_{Bj}r)\, dr = 0 \tag{35}$$

where α_{Bj} is defined by equation (9). Performing the integration leads to the expression

$$\sum_{i=1}^{NZ} \bar{a}_{ji}A_i + \bar{b}_jB_j + \sum_{i=1}^{NZ} \bar{c}_{ji}C_i = 0; \quad j = 1, 2, 3 \ldots NR \tag{36}$$

where

$$\bar{a}_{ji} = 2JP(\beta_A) \cdot JP(\delta_A) \cdot [(\beta_A{}^2 - \delta_A{}^2)(2\delta_A{}^2 - \omega^2)/(\alpha_B{}^2 - \delta_A{}^2) + 4\alpha_A{}^2\beta_A{}^2/(\alpha_B{}^2 - \beta_A{}^2)] \tag{37}$$

$$\bar{b}_j = h(\alpha_B{}^2 - n^2)/\alpha_B{}^2\left[(\alpha_B{}^2 + \beta_B{}^2)^2\begin{Bmatrix}SN(\beta_B h)\cos(\delta_B h) \\ SN(\delta_B h)\cos(\beta_B h)\end{Bmatrix} + 4\alpha^2\begin{Bmatrix}\delta_B{}^2 \cdot SN(\delta_B h)\cos(\beta_B h) \\ \beta_B{}^2\, SN(\beta_B h)\cos(\delta_B h)\end{Bmatrix}\right] \tag{38}$$

$$\bar{c}_{ij} = 2n \cdot JB(\beta_A) \cdot JP(\delta_A)(2\delta_A{}^2 - \omega^2)/(\alpha_B{}^2 - \delta_A{}^2) \tag{39}$$

Equation (18) is approximately satisfied by setting

$$\int_0^h \tau_{r\theta}(1, \theta, z)\begin{Bmatrix}\cos\alpha_{Ai}z \\ \sin\alpha_{Ai}z\end{Bmatrix} dz = 0 \tag{40}$$

where α_{Ai} is defined by equation (8). Performing the integration leads to the expression

$$\hat{a}_iA_i + \sum_{j=1}^{NR} \hat{b}_{ij}B_j + \hat{c}_iC_i = 0; \quad i = 1, 2, 3 \ldots NZ \tag{41}$$

where

$$\hat{a}_i = -2hn \cdot JP(\beta_A) \cdot JB(\delta_A) \cdot \psi_n(\delta_A) \cdot (\beta_A{}^2 - \alpha_A{}^2) - 4hn \cdot JP(\delta_A) \cdot JB(\beta_A) \cdot \psi_n(\beta_A) \cdot \alpha_A{}^2 \tag{42}$$

$$\hat{b}_{ij} = 4n[(\alpha_B{}^2 - \beta_B{}^2)/(\delta_B{}^2 - \alpha_A{}^2) + 2\beta_B{}^2/(\beta_B{}^2 - \alpha_A{}^2)] \times \begin{Bmatrix}(\delta_B h)^2\, SN(\delta_B h)\, SN(\beta_B h) \\ -\cos(\delta_B h)\cos(\beta_B h)\end{Bmatrix} \tag{43}$$

$$\hat{c}_i = 2h \cdot JB(\beta_A)[2JP(\delta_A)[(n^2-1) - \beta_A{}^2/2 - \psi_n(\beta_A)] - n^2 \cdot JB(\delta_A) \cdot \psi_n(\delta_A)] \quad (44)$$

Arranging equations (20), (36), and (41) in matrix form gives

$$\begin{bmatrix} \lfloor a \rfloor & [b] & \lfloor c \rfloor \\ [\bar{a}] & \lfloor \bar{b} \rfloor & [\bar{c}] \\ \lfloor \hat{a} \rfloor & [\hat{b}] & \lfloor \hat{c} \rfloor \end{bmatrix} \begin{Bmatrix} \{A\} \\ \{B\} \\ \{C\} \end{Bmatrix} = \begin{Bmatrix} 0 \end{Bmatrix} \quad (45)$$

The notation, $\lfloor\ \rfloor$, indicates a diagonal matrix. Because of the diagonal matrices, a condensation of the matrix can be performed which requires no matrix inversion except the trivial inversion of a diagonal matrix. To perform the condensation, solve for C_i from equation (41). In matrix form, this gives

$$\{C\} = -\lfloor \hat{c} \rfloor^{-1} [\lfloor \hat{a} \rfloor \{A\} + [\hat{b}] \{B\}] \quad (46)$$

substitution of equation (48) into equations (20) and (36) yields in matrix form

$$\lfloor a^* \rfloor \{A\} + [b^*] \{B\} = 0 \quad (47)$$

$$[\bar{a}^*] \{A\} + [\bar{b}^*] \{B\} = 0 \quad (48)$$

where

$$\lfloor a^* \rfloor = \lfloor a \rfloor - \lfloor c \rfloor \lfloor \hat{c} \rfloor^{-1} \lfloor \hat{a} \rfloor \quad (49)$$

$$[\bar{a}^*] = [\bar{a}] - [\bar{c}] \lfloor \hat{c} \rfloor^{-1} [\hat{a}] \quad (50)$$

$$[b^*] = [b] - \lfloor c \rfloor \lfloor \hat{c} \rfloor^{-1} [\hat{b}] \quad (51)$$

$$[b^*] = \lfloor \bar{b} \rfloor - [\bar{c}] \lfloor \hat{c} \rfloor^{-1} [\hat{b}] \quad (52)$$

From equation (47)

$$\{[A\} = -\lfloor a^* \rfloor^{-1} [b^*] \{B\} \quad (53)$$

subsitution into equation (48) gives

$$[d] \{B\} = 0 \quad (54)$$

where

$$[d] = [\bar{b}^*] - [\bar{a}]^* \lfloor a^* \rfloor^{-1} [b^*] \quad (55)$$

Note that $[d]$ has a size of NR by NR.

The solution process is to choose a frequency parameter ω, the height to diameter ratio, h, and Poisson's ratio, ν, compute the coefficients in equations (21–23), (37)–(39), (42)–(44), perform the condensation in equations (46)–(55) and evaluate the determinant of $[d]$ (in equation (55)). If the determinant is zero, the frequency parameter ω is a natural frequency for that value of h and ν. In the computer program, h and ν were chosen and steps were taken in ω until the determinant of $[d]$ changed sign then *regula-falsi* was used to determine ω to any desired degree of accuracy.

On finding ω, the mode shapes are found by solving equation (55) for the relative values of $\{B\}$. $\{A\}$ and $\{C\}$ are then found from equations (53) and (46), respectively. The displacement components are then found from summing the terms expressed in equations (2)–(4) times their appropriate coefficients (A_i, B_i, or C_i).

Results

Table 2 shows the convergence of the frequencies as the number of terms in the series are increased. The tabulated values are the lowest six frequencies for the odd modes of circumferential order one, a height-to-diameter ratio of one, and a Poisson's ratio of 0.344. An equal number of terms in each series was used for this example. It is clearly evident that as more terms are chosen, the frequencies converge, and their convergence is to the experimental results found by McMahon. It can further be noted that for the accuracy required to plot frequency spectrum curves, such as presented by McMahon, use of as few as two terms in each series is adequate.

It was found that for values of the height-to-diameter ratio which are different than one, it is desirable to use more terms in the "long" direction. Table 3 is intended to demonstrate this observation. For Table 3, the height-to-diameter ratio was chosen as 0.25. The gravest frequency was found for odd modes of circumferential order one and a Poisson's ratio of 0.344. It can be seen that for this case, it is profitable to have more terms in the radial direction than in the axial direction. In fact, the optimum relation appears roughly to be

$$NZ = NR \cdot h \quad (56)$$

This relationship was used in finding the frequency spectra shown in Figs. 1–10. The smallest number was chosen as two and the largest found from equation (56) (rounded up to an integer). Thus, for $h = 0.1$, $NZ = 2$, and $NR = 20$ was used; whereas, for $h = 2$, $NZ = 4$ and $NR = 2$ was used.

Figs. 1–10 are plots of the frequency spectrum for the odd and even modes of circumferential order from zero to four. These spectra are for a Poisson's ratio, ν, of 0.344 to correspond to McMahon's experimental results for aluminum. Note that on these plots, two frequency parameter scales are given. The one on the left in the plots is ω as defined earlier in this paper and as used in Tables 2 and 3. The one on the right in the plots is ω_E which is the frequency parameter as defined by McMahon. McMahon normalized his frequency with respect to the thin rod velocity rather than the shear velocity as done in this paper. The relation between the two is

$$\omega_E = \omega/[2(1+\nu)]^{1/2} \quad (57)$$

A close comparison of the frequency spectra in Figs. 1–8 shows that they are identical with the spectra found by McMahon. Any differences are within the accuracy to which such plots can be made.

For modes of circumferential order of zero to three, the results shown in Figs. 1–8 only slightly extend the range considered by McMahon. He considered the range $0.6 < \omega_E < 2.6$ and $0 < h < 1.65$

Table 2 Convergence of frequencies for the six lowest odd modes of circumferential order one with $\nu = 0.344$, $h = 1.0$, and an equal number of terms in each series

	Frequency					
Terms	1	2	3	4	5	6
1	2.17351	3.11476	4.48244			
2	2.15716	3.10492	4.07309	4.65399	5.01357	5.55055
3	2.15602	3.10350	4.04677	4.62302	5.00056	5.49844
4	2.15575	3.10297	4.03595	4.61175	4.99758	5.49733
5	2.15565	3.10272	4.03052	4.60645	4.99709	5.49699
6	2.15560	3.10257	4.02744	4.60354	4.99692	5.49684
8	2.15556	3.10243	4.02427	4.60061	4.99681	5.49672
10	2.15554	3.10236	4.02276	4.59924	4.99676	5.49666

Table 3 Convergence of the frequency for the gravest odd mode of circumferential order one with ν 0.344 and h = 0.25 as a function of the number of terms in radial and axial directions

	NZ			
NR	1	2	3	4
1	5.965	5.965	5.965	5.965
2	3.331	3.322	3.321	3.321
3	3.251	3.235	3.234	3.233
4	3.223	3.201	3.198	3.198
5	3.211	3.184	3.181	3.180
6	3.206	3.175	3.171	3.171
7	3.203	3.171	3.166	3.165
8	3.202	3.167	3.162	3.161
9	3.201	3.165	3.160	3.158
10	3.200	3.164	3.158	3.156
11	3.200	3.164	3.157	3.155
12	3.200	3.163	3.156	3.154

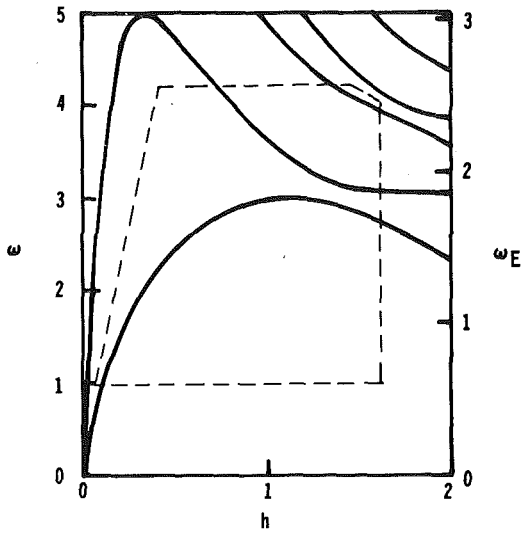


Fig. 1 Frequency spectra for odd modes of circumferential order zero. Dashed lines enclose the region investigated experimentally by McMahon [1]. Note: complete agreement was found.

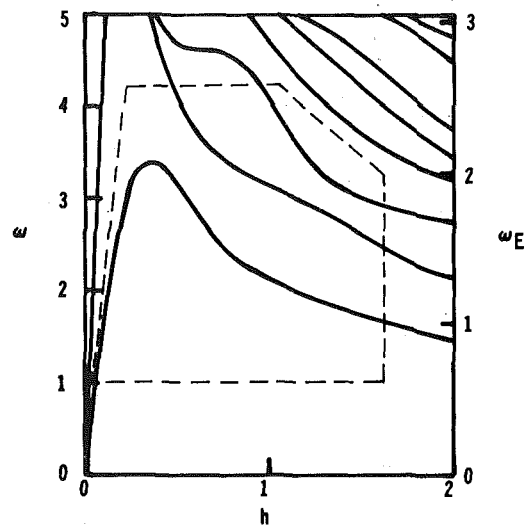


Fig. 3 Frequency spectra for odd modes of circumferential order one. Dashed lines enclose the region investigated experimentally by McMahon [1]. Note: complete agreement was found.

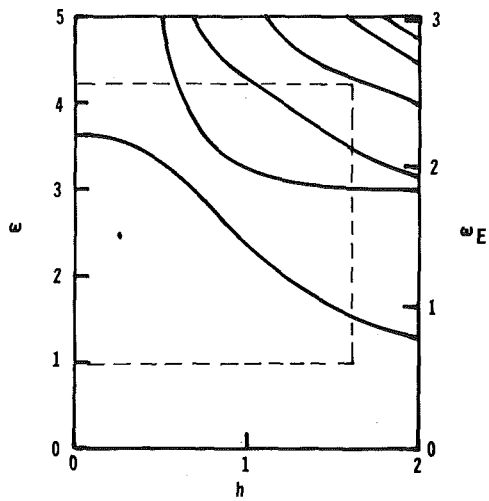


Fig. 2 Frequency spectra for even modes of circumferential order zero. Dashed lines enclose the region investigated experimentally by McMahon [1]. Note: complete agreement was found.

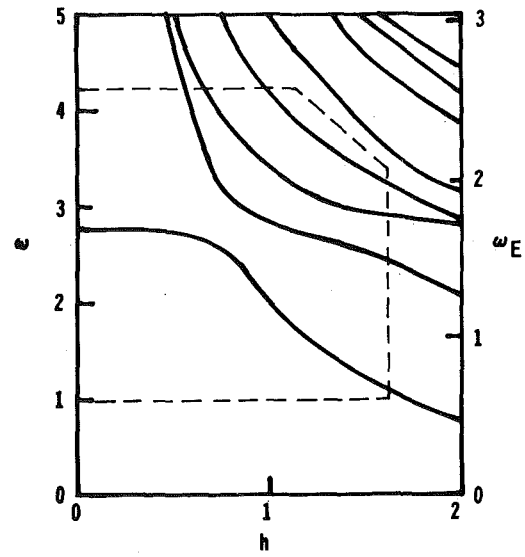


Fig. 4 Frequency spectra for even modes of circumferential order one. Dashed lines enclose the region investigated experimentally by McMahon [1]. Note: complete agreement was found.

whereas the plots shown are for $0 < \omega_E < 3$ and $0 < h < 2$. Figs. 9 and 10 are for circumferential order four and were not reported by McMahon. As noted by McMahon, for the odd modes, (Figs. 1, 3, 5, 7, and 9), there are an infinite number of frequencies corresponding to the flexural disk modes which approach zero as approximately straight lines with steep slopes. McMahon chose to shown only the gravest frequency in this range while I have chosen to omit frequencies in this range (small h) which have a slope (ω/h) greater than 50. It should further be noted that each of these steep curves in the region of small h connect (off scale) with the curves shown for larger h. For example, in Fig. 3, the second highest frequency shown at small h connects off scale with the second highest frequency shown in the range $h > 0.35$. It can also be noted that there are no frequency crossings in any of the plots. From my experience no frequency crossings should occur when the plotting is done for even and odd modes separately.

McMahon also plotted frequency spectra for steel, $\nu = 0.293$. Rather than duplicate these results, a more general investigation into the changes in the frequency parameter with changes in Poisson's ratio was done. The effect of varying Poisson's ratio over its full possible range from zero to one half was considered. Since the behavior is somewhat unremarkable, I have included only two plots which are typical of the changes. Figs. 11 and 12 are for the four gravest frequencies for the odd modes of circumferential order one with a height-to-diameter ratio of one. Both figures are actually the same data plotted using the two different frequency parameter definitions.

Conclusions

The method developed in this paper produces excellent results with a remarkably small amount of computational effort. That is, the computed frequencies converge with use of a very few terms in each of the series. While the analytic formulation (and computer program development) is not easy using this method, the actual computer time

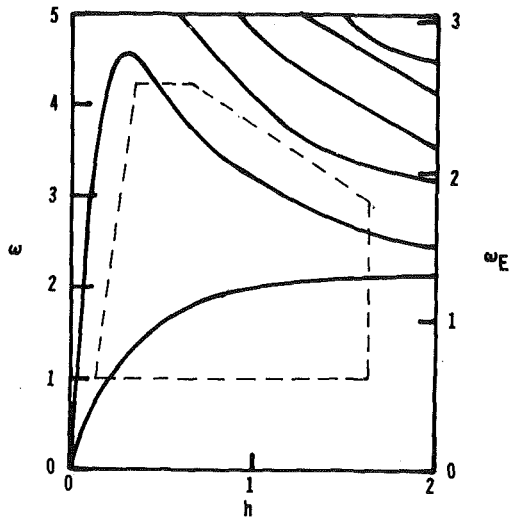


Fig. 5 Frequency spectra for odd modes of circumferential order two. Dashed lines enclose the region investigated experimentally by McMahon [1]. Note: complete agreement was found.

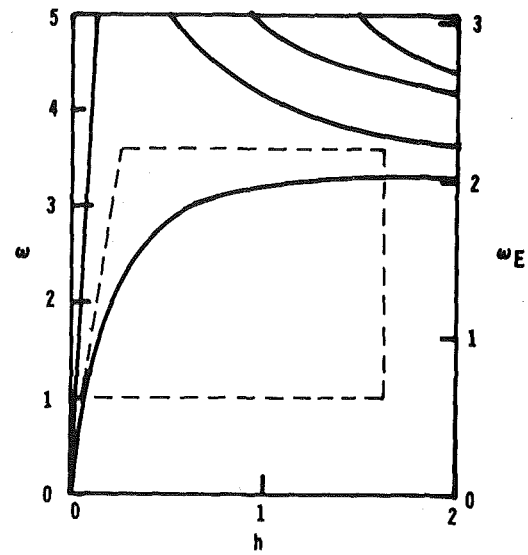


Fig. 7 Frequency spectra for odd modes of circumferential order three. Dashed lines enclose the region investigated experimentally by McMahon [1]. Note: complete agreement was found.

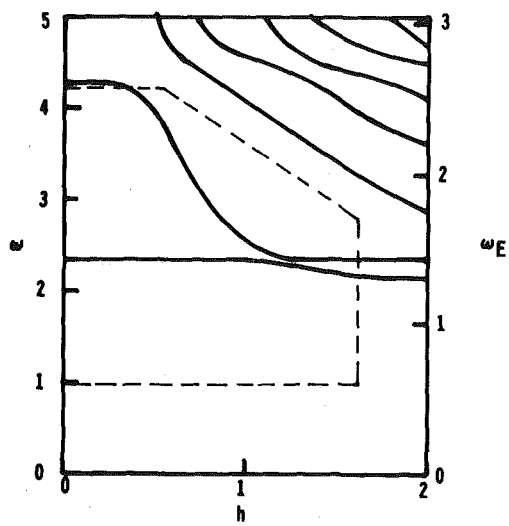


Fig. 6 Frequency spectra for even modes of circumferential order two. Dashed lines enclose the region investigated experimentally by McMahon [1]. Note: complete agreement was found.

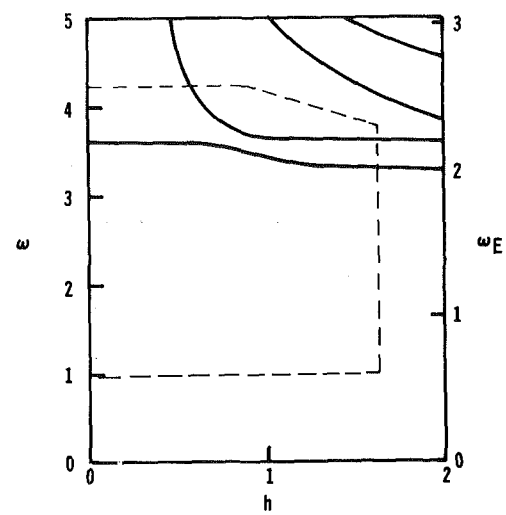


Fig. 8 Frequency spectra for even modes of circumferential order three. Dashed lines enclose the region investigated experimentally by McMahon [1]. Note: complete agreement was found.

in generating the data is quite small, and the frequencies found check completely with McMahon's thorough experimental study. Further, the method allows computation of the frequencies to any desired degree of accuracy.

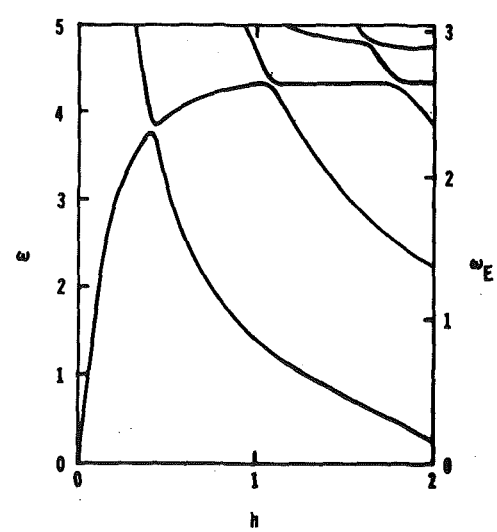


Fig. 9 Frequency spectra for odd modes of circumferential order four

References

1 McMahon, G. W., "Experimental Study of the Vibrations of Solid, Isotropic Elastic Cylinders," *Journal of the Acoustical Society of America*, Vol. 36, 1964, pp. 85–92.

2 Love, A. E. H., *A Treatise on the Mathematical Theory of Elasticity*, Dover, New York, 1944, pp. 287–291.

3 McNiven, H. D., and Perry, D. C., "Axially Symmetric Waves in Finite, Elastic Rods," *Journal of the Acoustical Society of America*, Vol. 34, 1962, pp. 433–437.

4 Rumerman, M., and Raynor, S., "Natural Frequencies of Finite Circular Cylinders in Axially Symmetric Longitudinal Vibration," *Journal of Sound and Vibration*, Vol. 15, 1971, pp. 529–543.

5 Hutchinson, J. R., "Axisymmetric Vibrations of a Free Finite Length Rod," *Journal of the Acoustical Society of America*, Vol. 51, 1972, pp. 233–240.

6 Hutchinson, J. R., "Axisymmetric Vibrations of a Solid Elastic Cylinder Encased in A Rigid Container," *Journal of the Acoustical Society of America*, Vol. 42, 1967, pp. 398–402.

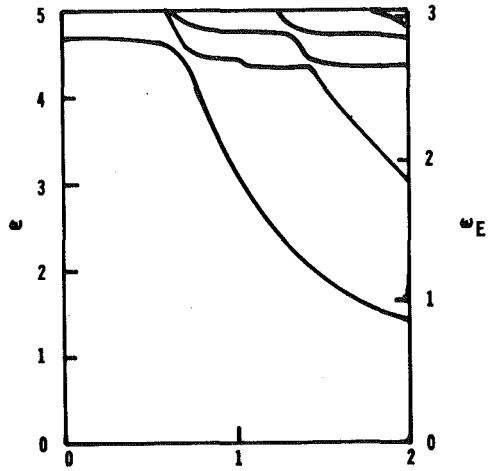


Fig. 10 Frequency spectra for even modes of circumferential order four

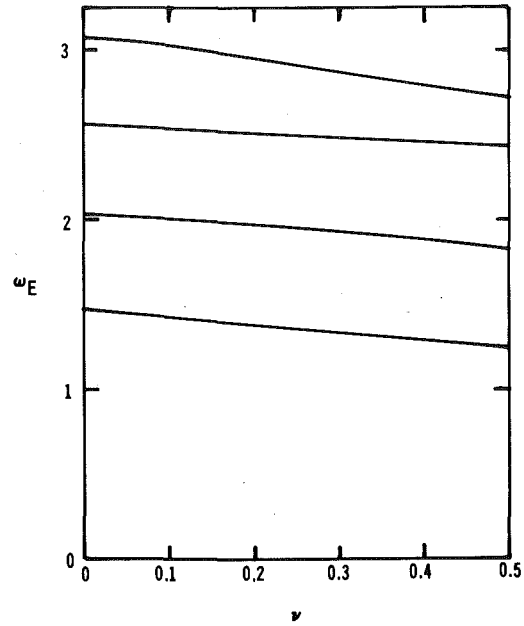


Fig. 12 Frequency parameter ω_E versus Poisson's ratio, ν, for the four lowest odd modes of circumferential order one with $h = 1.0$

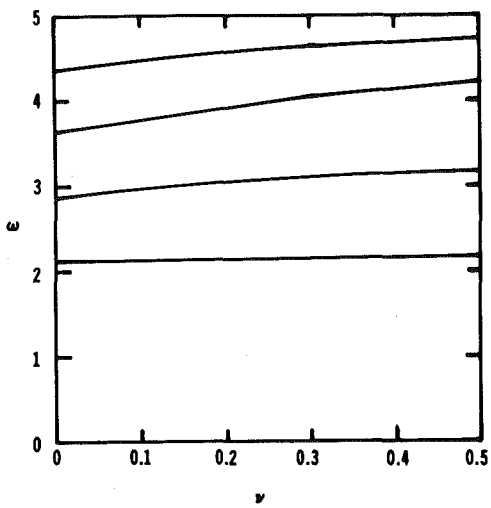


Fig. 11 Frequency parameter ω versus Poisson's ratio, ν, for the four lowest odd modes of circumferential order one with $h = 1.0$

7 Mindlin, R. D., "Influence of Rotary Inertia and Shear on Flexural Motions of Isotropic, Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 18, 1951, pp. 31–38.

8 Hutchinson, J. R., "Axisymmetric Flexural Vibrations of a Thick Free Circular Plate," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 139–144.

9 Rasband, S. N., "Resonant Vibrations of Free Cylinders and Disks," *Journal of the Acoustical Society of America,* Vol. 57, No. 4, Apr. 1975, pp. 899–905.

10 Mathieu, É., *Theorie de l'Élasticite Des Corps Solides,* Gauthier-Villars, Paris, 1890, pp. 140–181.

11 Taylor, G. I., "The Buckling Load for a Rectangular Plate With Four Clamped Edges," *Zeitschrift für Angewandte Mathematik und Mechanik,* Band 13, 1933, pp. 147–152.

12 Tomotika, S., "The Transverse Vibration of a Square Plate Clamped at Four Edges," *Philosophical Magazine,* Vol. XXI, 1936, pp. 745–760.

13 Timoshenko, S. P., "Bending of Rectangular Plates With Clamped Edges," *Proceedings of Fifth International Congress of Applied Mechanics,* Cambridge, Mass., 1938, pp. 40–43.

14 Pickett, G., "Application of the Fourier Method to the Solution of Certain Boundary Problems in the Theory of Elasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 11, 1944, pp. A176–182.

P. Hagedorn
Professor of Mechanics.

W. Teschner
Research Associate.

Institut für Mechanik,
61 Darmstadt, Hochschustrasse 1,
Technische Hochschule Darmstadt,
West Germany

An Instability Theorem for Steady Motions in Free and Restrained Dynamical Systems

The stability of steady motions in dynamical systems with ignorable coordinates is considered. In addition to the original "free" systems "restrained" systems are defined in such a way that the ignorable velocities remain constant along all motions; the stability behavior of the two systems is compared. A previously established instability theorem is generalized and three examples are given.

1 Introduction

In a recent paper [1] by one of the authors, free and restrained systems were defined and some of the results that were obtained earlier by Pascal [2], by Stepanov [3] and others, were generalized.

In particular, an instability theorem was presented, in which the behavior of both types of systems was compared. For this theorem to hold it was required that a certain matrix $\mathbf{B}(\bar{\mathbf{q}})$ should vanish identically. In the present paper the instability theorem is proven under much more general assumptions without the aforementioned condition on the matrix $\mathbf{B}(\bar{\mathbf{q}})$.

In three examples the stability behavior of free and restrained systems is compared.

2 The Free and the Restrained System

Consider the Lagrangian

$$L(\bar{\mathbf{q}}, \dot{\bar{\mathbf{q}}}, \dot{\mathbf{z}}) = \frac{1}{2}\dot{\bar{\mathbf{q}}}^T\mathbf{D}\dot{\bar{\mathbf{q}}} + \dot{\mathbf{z}}^T\mathbf{B}^T\dot{\bar{\mathbf{q}}} + \frac{1}{2}\dot{\mathbf{z}}^T\mathbf{C}\dot{\mathbf{z}} - \dot{\bar{\mathbf{q}}}^T\mathbf{g}(\bar{\mathbf{q}}) - \dot{\mathbf{z}}^T\mathbf{f}(\bar{\mathbf{q}}) - U(\bar{\mathbf{q}}) \quad (1)$$

of a dynamical system, $\bar{\mathbf{q}}$ and $\mathbf{z}$ denoting the n vector of "essential" coordinates and the m vector of ignorable coordinates. The matrices $\mathbf{D}(\bar{\mathbf{q}})$, $\mathbf{B}(\bar{\mathbf{q}})$, and $\mathbf{C}(\bar{\mathbf{q}})$ are $n \times n$, $n \times m$, and $m \times m$ submatrices of the positive-definite matrix

$$\mathbf{A}(\bar{\mathbf{q}}) = \begin{pmatrix} \mathbf{D}(\bar{\mathbf{q}}) & \mathbf{B}(\bar{\mathbf{q}}) \\ \mathbf{B}^T(\bar{\mathbf{q}}) & \mathbf{C}(\bar{\mathbf{q}}) \end{pmatrix} \quad (2)$$

of kinetic energy. In (1), $U(\bar{\mathbf{q}})$ is the potential energy and $\mathbf{g}(\bar{\mathbf{q}})$, $\mathbf{f}(\bar{\mathbf{q}})$ are vectors ($n \times 1$ and $m \times 1$) which define gyroscopic forces. As shown in [1],

$$R(\bar{\mathbf{q}}, \dot{\bar{\mathbf{q}}}) = \frac{1}{2}\dot{\bar{\mathbf{q}}}^T(\mathbf{D} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)\dot{\bar{\mathbf{q}}} + [\mathbf{B}\mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) - \mathbf{g}]^T\dot{\bar{\mathbf{q}}} - W_F(\bar{\mathbf{q}}) \quad (3)$$

is the Routhian of the (free) system, obtained by eliminating the ignorable coordinates, $\mathbf{c}$ being the vector of the corresponding momenta. The dynamic potential is

$$W_F(\bar{\mathbf{q}}) = \frac{1}{2}(\mathbf{c} + \mathbf{f})^T\mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) + U(\bar{\mathbf{q}}). \quad (4)$$

We assume that there is a steady motion defined by $\bar{\mathbf{q}} \equiv \mathbf{0}$, $\dot{\bar{\mathbf{q}}} \equiv \mathbf{0}$ and, without loss of generality, that $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, $\mathbf{g}(\mathbf{0}) = \mathbf{0}$.

In addition to the free system described by (1) we also consider a *restrained* system which differs from the free system in that in it the time derivatives of the ignorable coordinates are constant not only along the steady motions but along *all* motions. This is realized by means of additional generalized forces acting only in the degrees of freedom corresponding to the ignorable coordinates. As shown in [1] the motion of the restrained system follows from the Lagrangian

$$L_R(\bar{\mathbf{q}}, \dot{\bar{\mathbf{q}}}) = \frac{1}{2}\dot{\bar{\mathbf{q}}}^T\mathbf{D}\dot{\bar{\mathbf{q}}} + (\mathbf{B}\mathbf{C}_0^{-1}\mathbf{c} - \mathbf{g})^T\dot{\bar{\mathbf{q}}} - W_R(\bar{\mathbf{q}}) \quad (5)$$

with

$$W_R(\bar{\mathbf{q}}) = U(\bar{\mathbf{q}}) + \mathbf{f}^T\mathbf{C}_0^{-1}\mathbf{c} - \frac{1}{2}\mathbf{c}^T\mathbf{C}_0^{-1}\mathbf{C}\mathbf{C}_0^{-1}\mathbf{c}, \quad (6)$$

where $\mathbf{C}_0 = \mathbf{C}(\bar{\mathbf{q}} = \mathbf{0})$. Free and restrained systems have the same steady motions, however with possibly different stability behavior.

From the expressions for $R(\bar{\mathbf{q}}, \dot{\bar{\mathbf{q}}})$ and $L_R(\bar{\mathbf{q}}, \dot{\bar{\mathbf{q}}})$, the stability and instability theorems given in [1] were deduced. Here a more general form of the instability theorem will be given.

3 The Instability Theorem

In order to apply the results obtained in [4] to a certain system defined by $L(\bar{\mathbf{q}}, \dot{\bar{\mathbf{q}}})$, we need to compute its Hamiltonian

$$H(\bar{\mathbf{q}}, \bar{\mathbf{p}}) = H_2(\bar{\mathbf{q}}, \bar{\mathbf{p}}) + H_1(\bar{\mathbf{q}}, \bar{\mathbf{p}}) + H_0(\bar{\mathbf{q}}). \quad (7)$$

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, November, 1979; final revision, April, 1980.

We suppose that the system has an equilibrium at $\bar{\mathbf{q}} = \mathbf{0}$, $\bar{\mathbf{p}} = \mathbf{0}$. Since

$$\bar{\mathbf{p}} = \frac{\partial L}{\partial \dot{\bar{\mathbf{q}}}}, \tag{8}$$

this implies that in the corresponding Lagrangian the part linear in $\bar{\mathbf{q}}$ vanishes at $\bar{\mathbf{q}} = \mathbf{0}$.

In (3) and (5), this means that $\mathbf{B}_0 = \mathbf{0}$. Requiring this last condition would lead to a considerable loss in generality, since the matrix $\mathbf{A}(\bar{\mathbf{q}})$ of the quadratic form defining the kinetic energy and therefore also the matrix $\mathbf{B}(\bar{\mathbf{q}})$ are given. However, we can always add an exact differential to the Lagrangian without changing the equations of motion. Thus, by replacing (3) and (5), respectively, by

$$R^*(\bar{\mathbf{q}}, \bar{\mathbf{q}}) = \frac{1}{2}\dot{\bar{\mathbf{q}}}^T(\mathbf{D} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)\dot{\bar{\mathbf{q}}} + [\mathbf{B}\mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) - \mathbf{g} - \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}]^T\dot{\bar{\mathbf{q}}} - W_F(\bar{\mathbf{q}}) \tag{9}$$

$$L_R^*(\bar{\mathbf{q}}, \dot{\bar{\mathbf{q}}}) = \frac{1}{2}\dot{\bar{\mathbf{q}}}^T\mathbf{D}\dot{\bar{\mathbf{q}}} + (\mathbf{B}\mathbf{C}_0^{-1}\mathbf{c} - \mathbf{g} - \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c})^T\dot{\bar{\mathbf{q}}} - W_R(\bar{\mathbf{q}}) \tag{10}$$

we obtain a new Routhian of the free system and a new Lagrangian of the restrained system, for which the terms linear in $\dot{\bar{\mathbf{q}}}$ vanish at $\bar{\mathbf{q}} = \mathbf{0}$, without initially requiring $\mathbf{B}_0 = \mathbf{0}$.

From (9) and (10), the Hamiltonians of the free and restrained systems are obtained, respectively, as

$$H_F^* = \frac{1}{2}\{\bar{\mathbf{p}} - \mathbf{B}\mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\}^T(\mathbf{D} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1} \times \{\bar{\mathbf{p}} - \mathbf{B}\mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\} + W_F(\bar{\mathbf{q}}) \tag{11}$$

and

$$H_R^* = \frac{1}{2}\{\bar{\mathbf{p}} - \mathbf{B}\mathbf{C}_0^{-1}\mathbf{c} + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\}^T\mathbf{D}^{-1} \times \{\bar{\mathbf{p}} - \mathbf{B}\mathbf{C}_0^{-1}\mathbf{c} + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\} + W_R(\bar{\mathbf{q}}). \tag{12}$$

In [1] it was shown that

$$\Delta W = (W_F(\bar{\mathbf{q}}) - W_F(\mathbf{0})) - (W_R(\bar{\mathbf{q}}) - W_R(\mathbf{0})) \tag{13}$$

is at least positive-semidefinite, so that a minimum of $W_R(\bar{\mathbf{q}})$ at $\bar{\mathbf{q}} = \mathbf{0}$ implies a minimum of $W_F(\bar{\mathbf{q}})$ at the same point with the corresponding implications on the stability of the steady motions.

On the other hand, the theorem given in [4] guarantees instability if the function $H_0(\bar{\mathbf{q}})$ in (7) has a maximum. By comparing the momenta independent terms H_{F0}^* and H_{R0}^* in (11) and (12) the following can then be proved.

Theorem. If $H_{F0}^*(\bar{\mathbf{q}})$ has a maximum at $\bar{\mathbf{q}} = \mathbf{0}$, then the steady motion of the free system is unstable and it cannot be stabilized by introducing additional generalized forces which maintain constant the ignorable velocities. The restrained system is then also unstable.

In [1] the theorem was proved for $\mathbf{B} \equiv \mathbf{0}$. Here we give a general proof, which holds even when this condition is not fulfilled. Defining

$$\Delta H_0^*(\bar{\mathbf{q}}) = (H_{F0}^*(\bar{\mathbf{q}}) - H_{F0}^*(\mathbf{0})) - (H_{R0}^*(\bar{\mathbf{q}}) - H_{R0}^*(\mathbf{0})) = \Delta V(\bar{\mathbf{q}}) + \Delta W(\bar{\mathbf{q}}), \tag{14}$$

we have

$$\Delta V(\bar{\mathbf{q}}) = \frac{1}{2}\{-\mathbf{B}\mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\}^T \times (\mathbf{D} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1}\{-\mathbf{B}\mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\} - \frac{1}{2}\{-\mathbf{B}\mathbf{C}_0^{-1}\mathbf{c} + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\}^T\mathbf{D}^{-1}\{-\mathbf{B}\mathbf{C}_0^{-1}\mathbf{c} + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\}. \tag{15}$$

Introducing the abbreviations

$$\mathbf{a}(\bar{\mathbf{q}}) = \mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) - \mathbf{C}_0^{-1}\mathbf{c}, \tag{16}$$

$$\mathbf{b}(\bar{\mathbf{q}}) = (\mathbf{D} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1}\{-\mathbf{B}\mathbf{C}^{-1}(\mathbf{c} + \mathbf{f}) + \mathbf{g} + \mathbf{B}_0\mathbf{C}_0^{-1}\mathbf{c}\}, \tag{17}$$

we obtain

$$\Delta W(\bar{\mathbf{q}}) = \frac{1}{2}\mathbf{a}^T\mathbf{C}\mathbf{a}, \tag{18}$$

and, after a short calculation,

$$\Delta H_0^*(\bar{\mathbf{q}}) = \frac{1}{2}(\mathbf{a} - \mathbf{C}^{-1}\mathbf{B}^T\mathbf{b})^T(\mathbf{C} - \mathbf{B}^T\mathbf{D}^{-1}\mathbf{B})(\mathbf{a} - \mathbf{C}^{-1}\mathbf{B}^T\mathbf{b}). \tag{19}$$

From (16) and (17) it follows that $\mathbf{a}(\mathbf{0}) = \mathbf{0}$ and $\mathbf{b}(\mathbf{0}) = \mathbf{0}$, and clearly $\Delta H_0^*(\mathbf{q})$ is positive-semidefinite if the matrix $\mathbf{C} - \mathbf{B}^T\mathbf{D}^{-1}\mathbf{B}$ is at least positive-semidefinite for all $\bar{\mathbf{q}}$. Since the kinetic energy is assumed to be given by the positive-definite quadratic form

$$\frac{1}{2}\begin{Bmatrix}\bar{\mathbf{q}}\\ \dot{\mathbf{z}}\end{Bmatrix}^T\begin{pmatrix}\mathbf{D}(\bar{\mathbf{q}}) & \mathbf{B}(\bar{\mathbf{q}})\\ \mathbf{B}^T(\bar{\mathbf{q}}) & \mathbf{C}(\bar{\mathbf{q}})\end{pmatrix}\begin{Bmatrix}\dot{\bar{\mathbf{q}}}\\ \dot{\mathbf{z}}\end{Bmatrix} \tag{20}$$

it follows that

$$\begin{Bmatrix}-\mathbf{D}^{-1}\mathbf{B}\mathbf{x}\\ \mathbf{x}\end{Bmatrix}^T\begin{pmatrix}\mathbf{D} & \mathbf{B}\\ \mathbf{B}^T & \mathbf{C}\end{pmatrix}\begin{Bmatrix}-\mathbf{D}^{-1}\mathbf{B}\mathbf{x}\\ \mathbf{x}\end{Bmatrix} = \mathbf{x}^T(\mathbf{C} - \mathbf{B}^T\mathbf{D}^{-1}\mathbf{B})\mathbf{x} \tag{21}$$

is positive-definite in $\mathbf{x}$ for all values of the parameter $\bar{\mathbf{q}}$. The matrix $\mathbf{C} - \mathbf{B}^T\mathbf{D}^{-1}\mathbf{B}$ is therefore positive-definite and $\Delta H_0^*(\bar{\mathbf{q}})$ is at least positive-semidefinite.

The positive-semidefiniteness of $\Delta H_0^*(\bar{\mathbf{q}})$ means that a maximum of $H_{F0}^*(\bar{\mathbf{q}})$ at $\bar{\mathbf{q}} = \mathbf{0}$ implies a maximum of $H_{R0}^*(\bar{\mathbf{q}})$ at the same point; since a maximum of $H_0(\bar{\mathbf{q}})$ implies instability (see [4–6]), the proof of the theorem is complete.

It should be pointed out that the theorem makes no use of linearization; thus the result is valid for arbitrary nonlinear systems, even if the linear terms in the equations of motion vanish.

4 Illustrative Examples

It is not always readily recognized that the stability behavior of free and restrained systems is different in general. This is probably due to the fact that in the simple case of a heavy symmetric gyro with the spin angle as an ignorable coordinate, the equations of motion of the free and the restrained system are exactly the same. In what follows three mainly tutorial examples are used to elucidate the differences between free and restrained systems. A fourth example more important for technical applications is mentioned in Section 5.

Motion in a Central Force Field. The simplest example has one ignorable and one essential coordinate. Although such a system is not very general, since no gyroscopic forces exist in the reduced system (after elimination of the cyclic coordinates) it may very well serve to illustrate the different behavior of the free and the restrained systems.

Consider the planar motion of a point P of mass m in an inverse square central force field. Using polar coordinates r, ϕ,

$$L(r, \dot{r}, \dot{\phi}) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\phi}^2) + \frac{\gamma m}{r}, \tag{22}$$

where the constant $\gamma > 0$ defines the strength of the force field. Clearly ϕ is ignorable and one has

$$\frac{\partial L}{\partial \dot{\phi}} = mr^2\dot{\phi} = c, \tag{23}$$

yielding the Routhian

$$R(r, \dot{r}) = \frac{m}{2}\dot{r}^2 - W_F(r), \quad W_F(r) = -\frac{\gamma m}{r} + \frac{c^2}{2mr^2}. \tag{24}$$

From (24) one obtains

$$W_F'(r) = \frac{\gamma m}{r^2} - \frac{c^2}{mr^3}, \quad W_F''(r) = -\frac{2\gamma m}{r^3} + \frac{3c^2}{mr^4}. \tag{25}$$

The steady motions are obtained by setting W_F' equal to zero; this gives

$$r_0 = \frac{c^2}{\gamma m^2}, \tag{26}$$

so that the steady motions are

$$r(t) \equiv r_0 = \frac{c^2}{\gamma m^2}, \quad \dot{\phi}(t) \equiv \dot{\phi}_0 = \frac{c}{mr^2}. \tag{27}$$

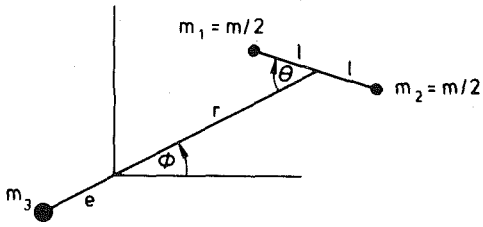


Fig. 1 Motion of a dumbbell satellite about a planet

It follows that $W_F''(r_0) = \gamma m/r_0^3 > 0$ so that the dynamic potential has a minimum and the steady motion of the free system is stable in $r, \dot{r}, \dot{\phi}$ as is well known. The equations of motion of the free system are obtained from (24) and (23)

$$\ddot{r} + \frac{\gamma}{r^2} - \frac{c^2}{m^2 r^3} = 0, \quad \dot{\phi} = \frac{c}{mr^2}, \tag{28}$$

where the constant c is fixed by the initial conditions. Now consider the Lagrangian L_R of the restrained system obtained from (22) by setting $\dot{\phi}(t) \equiv \dot{\phi}_0$ in L

$$L_R(r, \dot{r}) = \frac{m}{2}\dot{r}^2 - W_R(r), \quad W_R(r) = -\frac{\gamma m}{r} - \frac{m}{2}\dot{\phi}_0^2 r^2. \tag{29}$$

Physically the restrained system could possibly be realized by means of additional forces due to jets which would maintain artificially constant the angular speed $\dot{\phi}$.

Here W_R' vanishes for the same value of r_0 and $W_R''(r_0) = -3\gamma m/r_0^3 < 0$. The dynamic potential of the restrained system has a maximum at $r = r_0$. Since there are no gyroscopic forces in the restrained system $H_{R0}^* = W_R$, the maximum of W_R implies instability. The equations of motion of the restrained system are

$$\ddot{r} + \frac{\gamma}{r^2} - \dot{\phi}_0^2 r = 0, \quad \dot{\phi} = \dot{\phi}_0. \tag{30}$$

It is readily seen that the differential equations for $r(t)$ given in (28) and (30) are quite different. While the steady motion of the free system is stable, the steady motion of the restrained system is unstable in the present case. The example shows very well the different behavior of the free and the restrained system.

Dumbbell Satellite. The next slightly more general case has one ignorable and two essential coordinates. Here the reduced system—after elimination of the ignorable coordinate—has two degrees of freedom, so that gyroscopic forces may be present. A dumbbell moving in a central inverse square force field would be such a system. In order to make the example more interesting we will, however, consider a dumbbell satellite moving not in a central force field but under the action of the gravitational forces of a planet of finite mass which is not fixed in space. We shall see that the behavior of such a satellite is to some extent different from its behavior in a central force field.

Consider the planar motion of a dumbbell satellite about a planet, described by means of the generalized coordinates, r, θ, ϕ as shown in Fig. 1. The center of mass of the complete system is assumed to be at rest, so that the distance e of the planet (m_3) to the center of mass is given by $e = \nu r$, $\nu = m/m_3$.

The Lagrangian of this (free) system is then given by

$$L(r, \theta, \dot{r}, \dot{\theta}, \dot{\phi}) = \frac{m}{2}[(1+\nu)(\dot{r}^2 + r^2\dot{\phi}^2) + l^2(\dot{\theta} - \dot{\phi})^2] - U(r, \theta) \tag{31}$$

with the potential energy

$$U(r, \theta) = -\frac{m}{2}\mu\left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right) \tag{32}$$

where

$$\rho_{1,2} = \sqrt{l^2 + (1+\nu)^2 r^2 \mp 2(1+\nu)\, lr \cos\theta}, \tag{33}$$

with $\mu = Gm_3$, G being the general gravitational constant. The angle ϕ is obviously ignorable and with

$$\frac{\partial L}{\partial \dot{\phi}} = m\{[(1+\nu)r^2 + l^2]\dot{\phi} - l^2\dot{\theta}\} = c \tag{34}$$

the Routhian is obtained as

$$R(r, \theta, \dot{r}, \dot{\theta}) = \frac{m}{2}\left[(1+\nu)\dot{r}^2 + \frac{(1+\nu)r^2}{(1+\nu)r^2 + l^2} l^2\dot{\theta}^2\right] - \frac{cl^2}{(1+\nu)r^2 + l^2}\dot{\theta} - W_F \tag{35}$$

$$W_F(r, \theta) = \frac{1}{2}\frac{c^2}{m[(1+\nu)r^2 + l^2]} - \frac{m}{2}\mu\left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right). \tag{36}$$

The steady motions are of course such that the center of mass of the dumbbell describes a circle and the dumbbell axis is in the "spoke" or "arrow" position. Let us consider here only the arrow position, in which the dumbbell axis is tangential to the orbit of its center of mass. In this case some intermediate calculations give

$$\theta_0 = \pi/2, \quad r \equiv r_0$$
$$\dot{\phi}_0^2 = \omega_0^2 \frac{1+\nu}{[(1+\nu)^2 + \kappa_0^2]^{3/2}}. \tag{37}$$

Here $\omega_0^2 = \mu/r_0^3$ is the square of the angular orbital speed which would be obtained for the circular motion of a concentrated mass about a fixed center of attraction, and $\kappa_0 = l/r_0$.

For each value of l and r_0 a different steady motion is obtained. If the dynamic potential $W_F(r, \theta)$ is developed in power series about the steady motion one obtains

$$\overline{W}_F(\bar{r}, \bar{\theta}) = W_F\left(r_0 + \bar{r}, \frac{\pi}{2} + \bar{\theta}\right) - W_F\left(r_0, \frac{\pi}{2}\right)$$
$$= a\bar{r}^2 + b\bar{\theta}^2 + \mathcal{R} \tag{38}$$

where $\bar{r} = r - r_0$, $\bar{\theta} = \theta - \pi/2$ and $\mathcal{R}$ represents the terms of higher order in $\bar{r}, \bar{\theta}$. Hereby a, b are given by

$$a = \frac{\mu m(1+\nu)^3 r_0^2[(1+\nu)^2 r_0^2 + (1-3\nu)l^2]}{2[(1+\nu)^2 r_0^2 + l^2]^{5/2}[(1+\nu)r_0^2 l^2]}, \tag{39}$$

$$b = -\frac{3}{2} m\mu \frac{1}{\rho_0^5}(1+\nu)^2 l^2 r_0^2 < 0, \tag{40}$$

with

$$\rho_0 = \sqrt{l^2 + (1+\nu)^2 r_0^2}. \tag{41}$$

Clearly $\overline{W}_F$ can be negative-definite or indefinite, depending on the sign of the expression in brackets in the numerator of (39). It will be negative definite for

$$\kappa_0^2 > \frac{(1+\nu)^2}{3\nu - 1} > 0, \text{ indefinite}$$

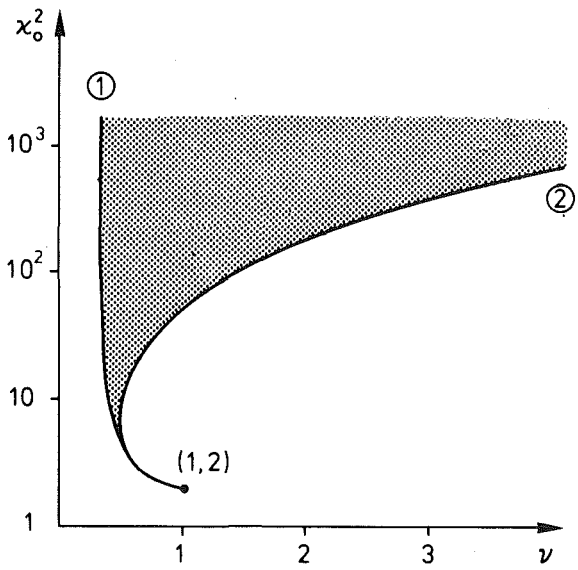
for $\kappa_0^2 < (1+\nu)^2/(3\nu - 1)$ or $\nu < 1/3$ (see Fig. 2). For all values of ν and κ_0^2 in realistic problems, $\overline{W}_F$ will therefore be indefinite and from the Thomson-Tait theorem we know that the corresponding steady motions are certainly unstable, the number of unstable roots being odd [7, p. 161]. On the other hand it may be instructive to examine the whole parameter plane, even for values of the parameters which are unrealistic from the viewpoint of practicality. We know that gyroscopic stabilization can be possible in the case of negative definite $\overline{W}_F$. The linearized equations of motion obtained from (35) are of the type

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}\begin{Bmatrix} \bar{r} \\ \bar{\theta} \end{Bmatrix} + \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix}\begin{Bmatrix} \bar{r} \\ \bar{\theta} \end{Bmatrix} + \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}\begin{Bmatrix} \bar{r} \\ \bar{\theta} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{42}$$

where $\alpha_1, \alpha_2, \beta, k_1$, and k_2 are functions of the given parameters.

If the stability analysis for (42) is done in the usual way it turns out that gyroscopic stabilization does indeed occur. The linearly stable region is shown as the shaded area in Fig. 2. This is to a certain extent surprising, since it is well known that the same steady motion is always unstable in the case of a dumbbell moving around a fixed center of attraction.

So far we have only examined the free system. Let us now try to



① $\varkappa_0^2 = \dfrac{(1+\nu)^2}{3\nu - 1}$

② $\varkappa_0^2 = (1+\nu)(17\nu - 4 \pm 12\sqrt{2\nu^2 - \nu})$

Fig. 2 Stability region for the arrow position of a dumbbell satellite (shaded area ≙ stability)

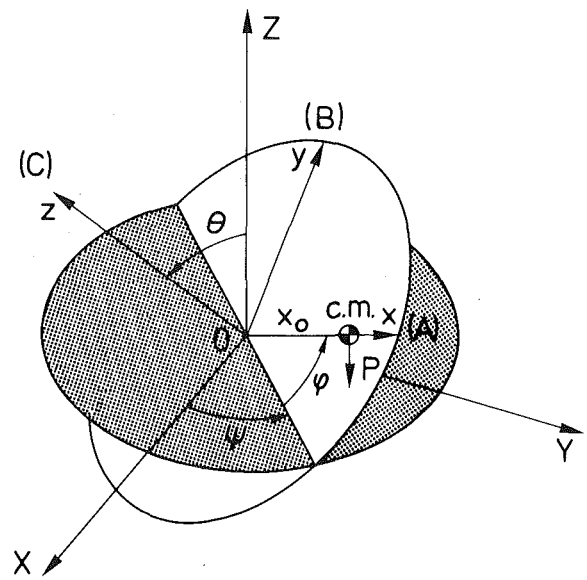


Fig. 3 Euler angles and location of the center of mass

apply the instability theorem and examine the stability behavior of the same steady motion for the case of the restrained system. For this the function H_{F0}^* is calculated from the modified Routhian

$$R^*(r, \theta, \dot{r}, \dot{\theta}) = R(r, \theta, \dot{r}, \dot{\theta}) + \frac{cl^2}{(1+\nu)r_0^2 + l^2}\dot{\theta}. \tag{43}$$

The function $H_{F0}^*(r, \theta)$ is obtained in the usual way, and it can be verified, that this function does not assume a maximum for the steady motions under consideration. The instability theorem in its coordinate-dependent form (see [6]) can therefore not be applied to this problem. The stability analysis of the steady motion for the restrained system yields in instability for all values of the positive parameters ν and κ_0^2.

The Heavy Gyrostat. In both examples previously discussed the original systems before the elimination of the ignorable coordinate had no gyroscopic terms, i.e., both the vectors **f** and **g** in (1) vanished identically. We now wish to give an example in which this is not the case. Also, in the two first examples the free system was "more stable" than the restrained system and this may suggest that this fact is always true; the stability theorem given in [1] and the instability theorem of the present paper support this thesis. In what follows we give an example which shows that this is not necessarily true in general.

To this purpose let us reconsider the heavy gyrostat already discussed in [1] with the Lagrangian

$$L = \frac{1}{2}A\omega_1^2 + \frac{1}{2}B\omega_2^2 + \frac{1}{2}C\omega_3^2 + k\omega_1 - Px_0 \sin\phi \sin\theta, \tag{44}$$

and

$$\begin{aligned} \omega_1 &= \dot{\psi}\sin\theta\sin\phi + \dot{\theta}\cos\phi, \\ \omega_2 &= \dot{\psi}\sin\theta\cos\phi - \dot{\theta}\sin\phi, \\ \omega_3 &= \dot{\phi} + \dot{\psi}\cos\theta. \end{aligned} \tag{45}$$

Here A, B, and C are the principal moments of inertia of the main body including the rotor mass, ω_1, ω_2, ω_3 are the components of the angular velocity of the main body along the principal axes, k is the constant relative moment of momentum of the rotor relative to the main body and acting along the x-axis, P is the weight, and x_0 gives the position of the center of mass on the principal axis corresponding to A (x-axis). The Euler angles ϕ, ψ, θ are shown in Fig. 3. Clearly the angle ψ is ignorable and the vectors **f** and **g** are

$$\mathbf{g} = \begin{Bmatrix} 0 \\ -k\cos\phi \end{Bmatrix}, \tag{46}$$

$$f = -k\sin\phi\sin\theta, \tag{47}$$

if we set $\bar{\mathbf{q}}^T = \{\phi, \theta\}$. It should be observed that here the free system already carries a constant speed rotor. By elimination of ψ one obtains the Routhian and

$$W_F(\phi, \theta) = \frac{1}{2}\frac{(c - k\sin\phi\sin\theta)^2}{C\cos^2\theta + (A\sin^2\phi + B\cos^2\phi)\sin^2\theta} + Px_0\sin\phi\sin\theta, \tag{48}$$

with

$$c = \partial L/\partial\dot{\psi}. \tag{49}$$

The gradient of W_F is given by

$$\frac{\partial W_F}{\partial\phi} = \frac{\sin\theta\cos\phi}{u^2}[Px_0u^2 - ku(c - k\sin\phi\sin\theta) - (A - B)\sin\phi\sin\theta(c - k\sin\phi\sin\theta)^2], \tag{50}$$

$$\frac{\partial W_F}{\partial\theta} = \frac{\cos\theta}{u^2}[Px_0u^2\sin\phi - ku(c - k\sin\phi\sin\theta)\sin\phi - (A\sin^2\phi + B\cos^2\phi - C)\sin\theta(c - k\sin\phi\sin\theta)^2] \tag{51}$$

with

$$u = C\cos^2\theta + (A\sin^2\phi + B\cos^2\phi)\sin^2\theta. \tag{52}$$

Equation (51) is fulfilled identically for all values of ϕ if we set $\theta_0 = \pi/2$. Since the case $\phi_0 = \pi/2$ was already dealt with in [1], we wish to discuss only the case $\phi_0 \neq \pi/2$ here. Setting (50) equal to zero we then obtain

$$Px_0 = k\frac{c - k\sin\phi_0}{A\sin^2\phi_0 + B\cos^2\phi_0} + (A - B)\left[\frac{c - k\sin\phi_0}{A\sin^2\phi_0 + B\cos^2\phi_0}\right]^2\sin\phi_0. \tag{53}$$

Using (49) with $\dot{\psi}_0 = \omega$ one obtains from (53)

$$\sin \phi_0 = \frac{Px_0 - k\omega}{\omega^2(A - B)} \tag{54}$$

Steady motions with $\theta_0 = \pi/2$, $\phi_0 \neq \pi/2$ are therefore possible for arbitrary values of ϕ_0 if the position of the center of mass is properly chosen, provided $A \neq B$. They correspond physically to regular precessions and their stability will be discussed in what follows. In the new variables $\bar{\phi} = \phi - \phi_0$, $\bar{\theta} = \theta - \pi/2$ we have from (48)

$$\overline{W}_F(\bar{\phi}, \bar{\theta}) = -\frac{\alpha}{2}\bar{\theta}^2 - \frac{\beta}{2}\bar{\phi}^2 + \mathcal{R} \tag{55}$$

where $\mathcal{R}$ represents the higher-order terms and

$$\alpha = \omega^2(C - B),$$
$$\beta = \left\{\omega^2 (A - B) - \frac{[k + 2\omega \sin \phi_0(A - B)]^2}{A \sin^2 \phi_0 + B \cos^2 \phi_0}\right\} \cos^2\phi_0. \tag{56}$$

The function $\overline{W}_F(\bar{\phi}, \bar{\theta})$ is positive-definite if $\alpha < 0$, $\beta < 0$ hold simultaneously, it is negative definite if both inequalities are reversed. If both inequalities are fulfilled the steady motion is certainly stable as follows from the Lagrange-Dirichlet theorem.

If only one of the inequalities $\alpha < 0$, $\beta < 0$ holds then $\overline{W}_f(\bar{\phi}, \bar{\theta})$ is indefinite and from the Thomson-Tait theorem on gyroscopic stabilization it follows that the steady motion is unstable, since there is an odd number of unstable degrees of freedom.

Let us now fix our attention for example on the case $\phi_0 = 0$, which corresponds to $Px_0 = k\omega$, and we assume $B < C$, $B < A$, $(k/\omega)^2 > (A - B)B$. This steady motion corresponds to a regular precession in which the rotor axis describes a horizontal plane through the point of suspension. W_F is indefinite and the steady motion is unstable for the free system.

The Lagrangian of the restrained system is easily obtained from (44) by setting $\dot{\psi} = \omega$ constant. The equations of motion of the restrained system linearized at $\phi_0 = 0$, $\theta_0 = \pi/2$ are then

$$\begin{aligned} C\ddot{\bar{\phi}} - (A - B + C)\,\omega\dot{\bar{\theta}} - (A - B)\omega^2\bar{\phi} &= 0, \\ A\ddot{\bar{\theta}} + (A - B + C)\,\omega\dot{\bar{\phi}} - (C - B)\,\omega^2\bar{\theta} &= 0. \end{aligned} \tag{57}$$

The characteristic equation is obtained with $\bar{\phi} = \bar{\bar{\phi}}e^{\lambda t}$, $\bar{\theta} = \bar{\bar{\theta}}e^{\lambda t}$ as

$$\begin{aligned} AC\lambda^4 + \lambda^2\omega^2\,[-A(A - B) - C(C - B) \\ + (A - B + C)^2] + \omega^4(A - B)(C - B) = 0 \end{aligned} \tag{58}$$

and it can easily be verified that all the four roots are purely imaginary under the conditions previously assumed. Therefore the steady motion $\theta_0 = \pi/2$, $\phi_0 = 0$ is (linearly) stable for the restrained system and unstable for the free system. This shows that the free system is not necessarily always more stable than the restrained system as one might conjecture.

5 Conclusion

In the present paper the instability theorem given in [1] is generalized.

The generally different behavior of the free and restrained systems is examined in three simple examples. The theorems obtained as well as the two first examples suggest that the free system may always be more stable than the restrained system. That this is not the case is shown by the third example, in which the precessional motion of a heavy gyrostat is studied. The same behavior is also found in gyrostat satellites [8] where sometimes the steady motion of the restrained system is stable and simultaneously that of the free system is unstable. Further work will be done on the gyrostat satellite with an inner rotor with controlled speed, of which the free and the restrained rotor are the two extreme cases.

Acknowledgment

The authors are indebted to S. Otterbein for checking the sometimes lengthy calculations.

References

1 Hagedorn, P., "On the Stability of Steady Motions in Free and Restrained Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 427–432; see also; Erratum, ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, p. 462.

2 Pascal, M., "Sur la recherche des mouvements stationnaires dans les systèmes ayant des variables cycliques," *Celestial Mechanics*, Vol. 12, 1975, pp. 337–358.

3 Stepanov, S., "Relation Between the Stability Conditions for Steady Motions for Three Types of Working Conditions," (in Russian), *Problems of Analytical Mechanics, Stability Theory and Control, Proceedings of the Second Chetaev Conference*, Vol. 2, Kazan, 1976.

4 Hagedorn, P., "Über die Instabilität konservativer Systeme mit gyroskopischen Kräften," *Archive for Rational Mechanics and Analysis*, Vol. 58, 1975, pp. 1–9.

5 Teschner, W., "Instabilität bei nichtlinearen konservativen Systemen mit gyroskopischen Kräften," Doctoral Dissertation, TH Darmstadt, 1977.

6 Teschner, W., "Zur Instabilität konservativer Systeme mit gyroskopischen Kräften", *ZAMM*, Vol. 57, 1977, pp. T 92–T 94.

7 Müller, P. C., *Stabilität und Matrizen*, Springer-Verlag, Berlin, 1977.

8 Otterbein, S. "On the Stability of Free and Restrained Gyrostats," (to appear).

E. A. Cohen, Jr.

Applied Mathematics Branch, Code R44,
Naval Surface Weapons Center,
White Oak Laboratory,
Silver Spring, Md. 20910

A Stochastic Model for Predicting the Velocity Distribution of Warhead Fragments[1]

It has been the practice of those who study the velocity of fragments from a warhead to average the velocity of the fragment over its path rather than to examine in greater detail the underlying stochastic process. In this paper we present a more careful analysis of the pertinent statistical process by taking into account the correlations that exist along the trajectory of the fragment. One assumes that the fragments generally tumble uniformly and that drag coefficient depends only on presented area. This seems to be a good approximation in the supersonic regime (from Mach 3 to Mach 6, for example). A brief discussion of the case when drag, lift, and gravity effects are all assumed to be present is given. Also, we present a new family of "multivariate uniform distributions" geared to the physics of the problem.

1 Introduction

The classical drag equation, useful in predicting the velocity of fragments in air, is given by

$$D = \tfrac{1}{2} C_D \rho A_p v^2 = -M_0 dv/dt, \tag{1}$$

where

D = drag force
C_D = drag coefficient
ρ = air density
A_p = presented area
v = velocity at time t
M_0 = mass of fragment

In a fragment trajectory analysis, such as that which we will be considering, the objects tumble in various ways as they proceed along their paths. It is too complicated to determine *a priori* just how they will tumble and just what, for that matter, the initial velocity will be. When a sample of fragments of identical size and shape is studied, it should be possible to procure a distribution of initial velocities and, through photography or other means, to obtain the presented area of the fragment from point to point along its path. Thus one could obtain sample correlations [5, p. 359]. We will assume that the velocity is supersonic (from roughly Mach 3 to Mach 6) and, therefore, consider that C_D is a function of presented area alone, ignoring the dependence on Mach number and Reynold's number [1, pp. 243–244]. We show, in the subsequent analysis, that we can then, treating (1) as a stochastic differential equation, model the process more systematically than has been done heretofore.

2 Stochastic Analysis for Prediction of Velocity (Constant Tumbling Rate Case)

Using distance r in place of time t as the running coordinate, differential equation (1) is easily converted into

$$dv/dr = -\tfrac{1}{2} C_D(A_p)\rho A_p v/M_0 \tag{2}$$

with stochastic initial condition

$$v = v_0. \tag{3}$$

Condition (3) means that we prescribe a distribution of initial velocities and, considering A_p as the random quantity (itself a function of time or, equivalently, distance), attempt to give a distribution of v as a function of time or distance. We think of (2), subject to (3), as a physical process each of whose realizations is a smooth representation. Therefore, when one solves (2), he obtains

$$v_r = v_0 \exp\left[-\rho \int_0^r C_D(A_p(s))A_p(s)ds/2M_0\right], \tag{4}$$

where v_r is the fragment velocity at distance r from its initial point. Thus v_r becomes a realization of a rather complicated random process. In a statistical context, (4) can be interpreted in a precise way [11, pp. 78–88].

Next let us expand (4) into an infinite series, obtaining

[1] This work was originally supported by the Systems Analysis Group, G10, of the Naval Surface Weapons Center.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, October, 1978; final revision, December, 1979.

$$v_r = v_0 \left[1 - \rho \int_0^r C_D(A_p(s)) A_p(s) ds / 2M_0 + \rho^2 \left(\int_0^r C_D(A_p(s)) A_p(s) ds \right)^2 / 8M_0^2 + \cdots \right]. \quad (5)$$

Now, in a bomb with a thin shell casing of small curvature, it is not unreasonable to assume that the fragments, which have relatively small content, are essentially planar. Also, depending on the shape of the fragment, it may either adopt a stable attitude or tumble through the air. The former is quite simply treated, for it is then seen that (4) has the solution

$$v_r = v_0 \exp(-kr), \quad (6)$$

where v_0 is a random variable and $k = \rho C_D(A_p) A_p / 2M_0$ is a function of the random variable A_p. The situation is even simpler, of course, when A_p ceases to be a random variable, i.e., when the fragment invariably adopts the same attitude. In that case, v_r becomes a linear function of the random initial speed v_0 alone. These cases are relatively easy to handle and, therefore, are not addressed in this paper. We are interested in the second case, wherein the fragment tumbles uniformly along its path and we want to predict statistically what the speed v_r will be as a function of the distance r from some initial position. For this purpose, we shall introduce a family of "multivariate uniform distributions" of sufficient generality for this problem. In the literature on multivariate analysis, there is very little mention of such distributions. One finds so-called multivariate Beta distributions and generalized multivariate Beta distributions discussed, but one notes, in every case, that the marginal distributions cannot be adjusted to be uniform for every random variable [6, pp. 231–238, 9, 10].

Such a family is indeed not too difficult to obtain. We first illustrate how to derive a bivariate uniform distribution F whose correlation coefficient [5, p. 86]

$$\rho_F = \text{cov}(X, Y) / \sigma(X)\sigma(Y) \quad (7)$$

can assume any value between -1 and 1. To do this, use will be made of two extreme distributions both of which are uniform distributions in two variates. The first will have a correlation coefficient of 1 and the second a correlation coefficient of -1. We make use of the fact that, for such distributions, the mass must be concentrated along a line [5, p. 87]. Also, we know, in the univariate case, that the uniform density is given by [5, p. 106]

$$f(A) = \begin{cases} 1/A_{\max}, & 0 \le A \le A_{\max} \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

This means that the cumulative distribution function for (8) is

$$F(A) = A/A_{\max}, \quad 0 \le A \le A_{\max}. \quad (9)$$

Let us suppose that $F_1(x, y)$ is to be a cumulative distribution function in two variates such that ρ_{F_1}, its correlation coefficient, is 1 and such that its marginal distribution function on either x or y is given by (9). Then, from (9), such will be true if we require that the measure of any rectangle in the (x, y)-plane be in direct proportion to the length of the line $y = x$ which it contains (see Fig. 1). Since it is necessary to normalize by the factor $2^{-1/2}$, this means that such a measure equals the length of the projection of the segment $y = x$ upon either axis.

For such a distribution, the cumulative function is well defined and is just

$$F_1(x, y) = P(X \le x, Y \le y) = \begin{cases} x/A_{\max}, & x \le y \\ y/A_{\max}, & y \le x. \end{cases} \quad (10)$$

F_1 is not an absolutely continuous function, i.e., it is not the integral of its second partial derivative, $\partial^2 F_1 / \partial x \partial y$, which, in fact, equals zero everywhere except on the line $y = x$, where it does not exist. From (10) it is clear that $F_1(x, A_{\max}) = x/A_{\max}$ and that $F_1(A_{\max}, y) = y/A_{\max}$, so that the marginal distribution functions are those of a uniform distribution. Therefore, (10) is a bivariate uniform distribution, and it can be shown, from the definition of the Riemann-Stieltjes integral [5, pp. 184–188] (generalized to two dimensions), that $\rho_{F_1}(x, y) = 1$.

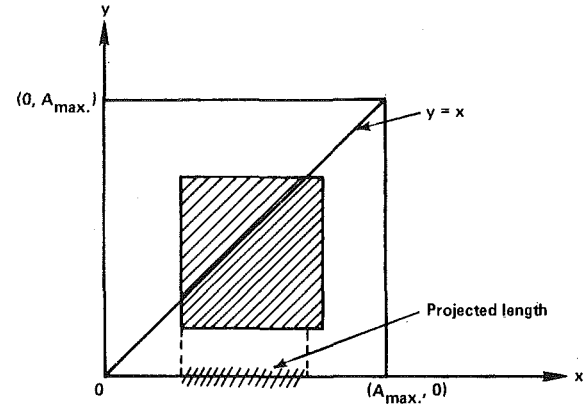


Fig. 1 Measure of line segment of $y = x$ when $\rho_{F_1} = 1$

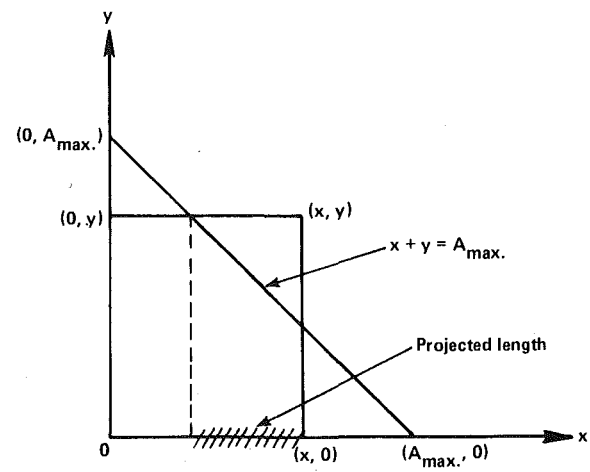


Fig. 2 Measure of line segment of $y = x$ when $\rho_{F_2} = -1$

Next we derive a distribution F_2 such that $\rho_{F_2} = -1$. For this purpose, let us consider the line $x + y = A_{\max}$, whose slope is -1. Referring to Fig. 2, we see that the measure of the rectangle with vertices $(0, 0)$, $(x, 0)$, (x, y), and $(0, y)$ is to be given by the length of the line segment of $y = A_{\max} - x$ contained within the rectangle, projected upon either axis and then normalized to 1. We find that the corresponding distribution function is

$$F_2(x, y) = \begin{cases} 0, & y \le A_{\max} - x \\ (x + y - A_{\max})/A_{\max}, & y \ge A_{\max} - x. \end{cases} \quad (11)$$

Again one sees that $F_2(x, A_{\max}) = x/A_{\max}$ and that $F_2(A_{\max}, y) = y/A_{\max}$. One can show, also, that $\rho_{F_2}(x, y) = -1$. Now let us form a new distribution function by mixing (10) and (11). We have

$$F(x, y) = \alpha F_1(x, y) + (1 - \alpha) F_2(x, y), \quad (12)$$

where $0 \le \alpha \le 1$. It is seen that (12) is likewise a cumulative uniform distribution, expressed as a convex combination of a distribution F_1 with $\rho_{F_1} = 1$ and a distribution F_2 with $\rho_{F_2} = -1$. In fact, one finds that

$$\rho_F(x, y) = 2\alpha - 1, \quad 0 \le \alpha \le 1, \quad (13)$$

since the marginal distributions of F_1 and F_2 are precisely the same. We have therefore discovered a one-parameter family of bivariate uniform distributions whose correlation coefficient can be adjusted according to the value of this parameter.

Let us now proceed to the trivariate case, i.e., a case in which three random variables X, Y, and Z are involved. In this situation, three pairwise correlations are to be considered, which represent three

constraints, and we would like to form again a basis of functions for a general class of trivariate uniform distributions. Since the convexity requirement constitutes a fourth constraint, we want a set of four functions. For this purpose, let $\bar{c} = (\rho_F(x, y), \rho_F(x, z), \rho_F(y, z))$ be a representative triple of correlation coefficients. Suppose that we can obtain four trivariate uniform functions of x, y, and z, $F_1(x, y, z)$, $F_2(x, y, z)$, $F_3(x, y, z)$, and $F_4(x, y, z)$, respectively, with the properties: $\bar{c}_{F_1} = (1, 1, 1)$, $\bar{c}_{F_2} = (-1, -1, 1)$, $\bar{c}_{F_3} = (-1, 1, -1)$, and $\bar{c}_{F_4} = (1, -1, -1)$. One then forms the convex combination

$$F(x, y, z) = \sum_{i=1}^{4} \alpha_i F_i(x, y, z), \quad \sum_{i=1}^{4} \alpha_i = 1, \quad \alpha_i \geq 0. \tag{14}$$

Once we assume that a fragment tumbles uniformly, it is clear that the correlations between presented areas at different points on its path must be a periodic function of the path length or separation. In fact, it is seen that a correlation coefficient, from a geometrical point of view, is simply the cosine of the angle between two given directions [8, p. 327]. Also, if we assume, as in the bivariate case, that the marginal distributions of the F_i's in (14) are all identical, then from (14) and definition (7), we see that

$$\rho_F(x, y) = \sum_{i=1}^{4} \alpha_i \rho_{F_i}(x, y), \tag{15}$$

with similar relationships for the pairs (x, z) and (y, z). Then, to procure a distribution function $F(x, y, z)$ with assigned correlation coefficients $\rho_F(x, y)$, $\rho_F(x, z)$, and $\rho_F(y, z)$ between presented areas at three given points of the fragment trajectory, we form the relations

$$\begin{aligned} \rho_F(x, y) &= \alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 = \cos(\omega_1 d_{12}) \\ \rho_F(x, z) &= \alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 = \cos(\omega_2 d_{13}) \\ \rho_F(y, z) &= \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = \cos(\omega_3 d_{23}) \\ &\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1, \quad \alpha_i \geq 0, \text{ every } i. \end{aligned} \tag{16}$$

Here ω_1, ω_2, and ω_3 are parameters to be determined, and d_{12}, d_{13}, and d_{23} are the distances between the three pairs of points mentioned above. (16) is a linear system in the α_i's which can be readily solved to give

$$\begin{aligned} \alpha_1 &= [1 + \cos(\omega_1 d_{12}) + \cos(\omega_3 d_{23}) + \cos(\omega_2 d_{13})]/4 \\ \alpha_2 &= [1 - \cos(\omega_1 d_{12}) + \cos(\omega_3 d_{23}) - \cos(\omega_2 d_{13})]/4 \\ \alpha_3 &= [1 - \cos(\omega_3 d_{23}) + \cos(\omega_2 d_{13}) - \cos(\omega_1 d_{12})]/4 \\ \alpha_4 &= [1 - \cos(\omega_3 d_{23}) - \cos(\omega_2 d_{13}) + \cos(\omega_1 d_{12})]/4. \end{aligned} \tag{17}$$

The critical question now arises as to whether or not the α_i's given by (17) can be expected to be nonnegative. Otherwise (14) may fail to be a distribution function, and it is possible that F would become negative. To show that this can happen, note that

$$E\left(\frac{X_1 - m_1}{\sigma_1} - \frac{X_2 - m_2}{\sigma_2} - \frac{X_3 - m_3}{\sigma_3}\right)^2 = 3 + 2(\rho_{23} - \rho_{12} - \rho_{13}) \geq 0, \tag{18}$$

where m_i is the expected value of X_i, σ_i^2 its variance, and ρ_{ij} the correlation coefficient between X_i and X_j, as defined by (7). Using (18), we see that $\alpha_2 \geq -\frac{1}{8}$, for example. It is clear, also, from (17), that we can get similar results for α_1, α_3, and α_4. In addition, note, from (17), that $\alpha_i + \alpha_j = (1 \pm \rho)/2$ when $i \neq j$, where ρ is one of the cosines. This means, of course, that at most one of the α's is negative. Since, as noted previously, the correlation coefficients must satisfy the law of cosines in spherical trigonometry, [2, p. 168], namely,

$$\cos\theta_3 = \cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2 \cos D, \tag{19}$$

we see that $\bar{c} = (-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$ is a feasible triple of correlations. For this set, $\alpha_1 = -1/8$, so that, indeed, the lower bound is attained. Although (14) may sometimes cease to be a distribution function in the classical sense of the word, there is no reason why we may not employ it. Also, one could use multivariate Gaussian processes as approximations to the multivariate uniform distributions. These processes would obviously have large variances in the cases where the marginals are uniform, and they could be used when the distributions are not multivariate uniform. It is an open question as to whether or not there exist more robust multivariate uniform distributions than the ones we are generating here.

We now form the distribution functions corresponding to F_1, F_2, F_3, and F_4, respectively. For F_1, we must build a c.d.f. corresponding to mass concentration along the line $x = y = z$. As in two dimensions, we define the measure of a typical volume element to be the normalized length of the segment of the line $x = y = z$ which it contains. We find that

$$F_1(x, y, z) = \begin{cases} x, & x \leq y \leq z, \quad x \leq z \leq y \\ y, & y \leq x \leq z, \quad y \leq z \leq x \\ z, & z \leq x \leq y, \quad z \leq y \leq x. \end{cases} \tag{20}$$

For F_2, we form the c.d.f. for mass concentration alone the line $y = z = A_{\max} - x$. This gives us

$$F_2(x, y, z) = \begin{cases} 0, & z \leq A_{\max} - x \\ 0, & x \leq A_{\max} - y \\ (x + y - A_{\max})/A_{\max}, & z \geq y \geq A_{\max} - x \\ (x + z - A_{\max})/A_{\max}, & y \geq z \geq A_{\max} - x. \end{cases} \tag{21}$$

The c.d.f. F_3 corresponds to the line $x = z = A_{\max} - y$ and is

$$F_3(x, y, z) = \begin{cases} 0, & z \leq A_{\max} - y \\ 0, & y \leq A_{\max} - x \\ (x + y - A_{\max})/A_{\max}, & z \geq x \geq A_{\max} - y \\ (y + z - A_{\max})/A_{\max}, & x \geq z \geq A_{\max} - y. \end{cases} \tag{22}$$

Finally, the c.d.f. F_4 is found to correspond to $x = y = A_{\max} - z$ and is given by

$$F_4(x, y, z) = \begin{cases} 0, & y \leq A_{\max} - z \\ 0, & z \leq A_{\max} - x \\ (x + z - A_{\max})/A_{\max}, & y \geq x \geq A_{\max} - z \\ (y + z - A_{\max})/A_{\max}, & x \geq y \geq A_{\max} - z. \end{cases} \tag{23}$$

A typical diagram, for the case of F_3, is shown in Fig. 3.

Now let us return to our discussion of the problem of assessment of fragment velocity. First of all, one needs a model for drag coefficient as a function of presented area. A simple model, using a first-order Taylor series development, is to assume that C_D is a linear function

$$C_D(A_p(s)) = a_1 A_p(s) + a_2 \tag{24}$$

and that a_1 and a_2 are shape-dependent parameters determined experimentally. Now, taking the expectation of both sides of (5) [6, p. 104], we have

$$\begin{aligned} \bar{v}_r = \bar{v}_0 \Big[1 &- \rho \int_0^r \overline{C_D A_p} ds/2M_0 + \rho^2 \int_0^r \int_0^r \overline{C_D(A_p(s))A_p(s)C_D(A_p(t))A_p(t)}\, dsdt/8M_0^2 \\ &- \rho^3 \int_0^r \int_0^r \int_0^r \overline{C_D(A_p(s))A_p(s)C_D(A_p(t))A_p(t)C_D(A_p(u))A_p(u)}\, dsdtdu/48M_0^3 + \ldots \Big], \end{aligned} \tag{25}$$

where a bar denotes expected value. Then, for instance, one finds, using (24), that

$$\begin{aligned} \overline{C_D(A_p(s))A_p(s)C_D(A_p(t))A_p(t)} &= a_1^2 \overline{A_p^2(s)A_2^2(t)} + a_1 a_2 \overline{A_p(s)A_p^2(t)} \\ &\quad + a_1 a_2 \overline{A_p^2(s)A_p(t)} + a_2^2 \overline{A_p(s)A_p(t)}. \end{aligned} \tag{26}$$

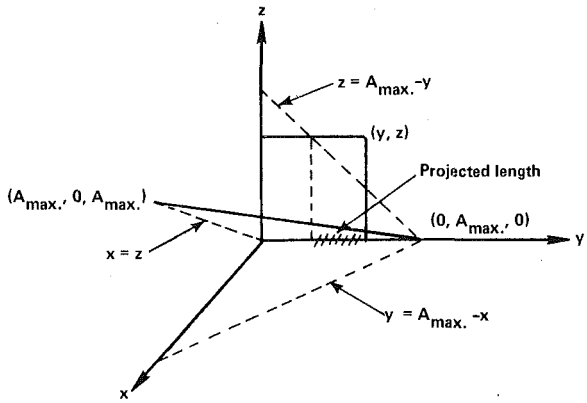


Fig. 3 Measure of line segment of $x = z = A_{max} - y$ for c.d.f. F_3

We are assuming in this paragraph that our stochastic process is stationary [11, p. 69], i.e., that our phenomena depend only on the separation, or distance, $|s - t|$. This is a simple consequence of tumbling with constant angular velocity, for, in that case, the distance between points is clearly all that matters in assessing the relevant statistical quantities. In particular, then, the second and third terms of (26) should be equal. One now computes a set of sample correlation coefficients $\rho_F{}^*(A_p(s_i), A_p(t_i))$ [5, p. 359] and assumes that ρ_F in (13) has the functional form $\cos(\omega(s_i - t_i))$. Then the parameter ω is to be determined by solving the nonlinear least-squares problem

$$\min_{\omega} \sum_{i=1}^{k} [\rho_F{}^*(A_p(s_i), A_p(t_i)) - \cos(\omega(s_i - t_i))]^2. \tag{27}$$

After finding a feasible ω-value, α and $1 - \alpha$ are expressed through the equations

$$\alpha = [1 + \cos(\omega d_12)]/2, \; 1 - \alpha = [1 - \cos(\omega d_{12})]/2,$$

where d_{12} is the separation between any two points on the trajectory. Going back to the form for F as given by (12), equation (26) may then be computed, appealing to the definition of the Riemann-Stieltjes integral [3, pp. 184–188], generalized to two dimensions. Now a very interesting, but simple, observation can be made. Using (14) and the four functional representations given by (20)–(23), and setting $z = A_{max}$, we find, as it should, that the two-dimensional marginal c.d.f. reduces to the case (12). We can do the same in turn with $y = A_{max}$, computing the marginal for x and z, and $x = A_{max}$, obtaining the marginal for y and z. The result is that, when the tumbling rate is constant, $\omega = \omega_1 = \omega_2 = \omega_3$, which is physically clearly the case. The frequency of the correlation coefficients should all be the same. This means that only one optimization problem need be solved, namely, that of (27). Now the third integrand in (25) is found to be

$$\begin{aligned}\overline{C_D(A_p(s))A_p(s)C_D(A_p(t))A_p(t)C_D(A_p(u))A_p(u)} \\ = a_1{}^3\overline{A_p{}^2(s)A_p{}^2(t)A_p{}^2(u)} + a_1{}^2a_2\overline{A_p{}^2(s)A_p{}^2(t)A_p(u)} \\ + 2a_1{}^2a_2\overline{A_p(s)A_p{}^2(t)A_p{}^2(u)} + 2a_1a_2{}^2\overline{A_p(s)A_p{}^2(t)A_p(u)} \\ + a_1a_2{}^2\overline{A_p(s)A_p(t)A_p{}^2(u)} + a_2{}^3\overline{A_p(s)A_p(t)A_p(u)}.\end{aligned} \tag{28}$$

Equation (28) can be computed using (14), together with (17). The integrands are seen to involve only linear combinations of cosines, so they are easily handled. One would proceed generally, developing quadrivariate uniform distributions and higher order multivariate uniform processes. Then, since (25) is an alternating series, the remainder after a given number of terms is always bounded in absolute value by the first term neglected, provided the terms are monotone decreasing [12, pp. 293–294]. Thus we keep employing our procedure, estimating the moments in our series, at the same time accounting for our truncation error. Now, from (4),

$$v_r{}^2 = v_0{}^2 \exp\left[-\rho \int_0^r C_D(A_p(s))A_p(s)ds/M_0\right], \tag{29}$$

which has the same form as (4). It is clear then that the variance of v_r can be similarly computed.

One special case arises when one decides only to retain the first two terms in (25), considering the remainder to be small. This would mean that we have

$$\bar{v}_r \cong \bar{v}_0[1 - \rho(\mu_1\mu_2 + \mu_{12})r/2M_0], \tag{30}$$

where $\mu_1 = E(C_D)$, $\mu_2 = E(A_p)$, and $\mu_{12} = \text{cov}(C_D, A_p)$. The right side of (30) is the linear part of

$$\bar{v}_o \exp[-\rho(\mu_1\mu_2 + \mu_{12})r/2M_0], \tag{31}$$

which bears some semblance to an expression often given in the literature on this subject except for the presence of the covariance term μ_{12}. If the same approximation is invoked for the mean of (29), we have

$$\text{var } v_r \cong \text{var}(v_0) \exp[-\rho(\mu_1\mu_2 + \mu_{12})r/M_0]. \tag{32}$$

It follows from (31) and (32) that the coefficient of variation [5, p. 72] is approximately constant

$$(\text{var } v_r)^{1/2}/\bar{v}_r \cong (\text{var } v_0)^{1/2}/\bar{v}_0. \tag{33}$$

3 Stochastic Analysis for Prediction of Velocity (Nonconstant Tumbling Rate Case)

In the general case, the situation is governed by a distribution of initial angular momenta. Also, due to drag, the rate of tumbling will vary along the trajectory. The analysis is still straightforward, but a bit more complicated. For one thing, the stochastic process is no longer stationary. The ω_i's in (17) now depend on the distance traversed. We preserve the form of the correlation function except that we introduce an analytic expression for the ω_i's in terms of distance. Also, we suppose that our observation pairs, in the bivariate discussion, are taken at $s = 0$ and $s = s_i$, where s is the distance along the path. Thus we take an initial reading of presented area, together with later readings. The ω_i's can be developed in a Taylor's series up to a convenient order term. However, we have observed, in Section 2, that the ω_i's should reduce to one quantity ω. Therefore, if, for example, we expand ω up to the second-order term, we have a quadratic expression for it, namely,

$$\omega(s) = b_0 + b_1s + b_2s^2. \tag{34}$$

The optimization problem to be solved is then

$$\min_{b_0,b_1,b_2} \sum_{i=1}^{k} [\rho_F{}^*(A_p(0), \quad A_p(s_i)) - \cos(\omega(s_i)s_i)]^2, \tag{35}$$

where $\omega(s)$ is given by (34). The analysis, in essential detail, remains the same as in Section 2. Once $\omega(s)$ is determined, one is forced to integrate cosines of polynomial functions of distance. This must be done numerically. Also, such parameters as the expected velocity and variance are computed, conditioned upon a given initial angular momentum. Therefore, the unconditional expected velocity is

$$\bar{v} = \int_D \bar{v}(\mathbf{M})dG(\mathbf{M}), \tag{36}$$

where $\mathbf{M}$ represents the initial angular momentum vector, D the domain of angular momenta, and G the cumulative distribution function. $\bar{v}(\mathbf{M})$ is expected velocity, given $\mathbf{M}$. Similarly, we have

$$\overline{v^2} = \int_D \overline{v^2}(\mathbf{M})dG(\mathbf{M}), \tag{37}$$

where $\overline{v^2}(\mathbf{M})$ is the expected value of the square of velocity for a given initial $\mathbf{M}$.

4 Extension of Analysis to Include Lift and Gravity Effects

It is important to note that any forces perpendicular to the direction of motion will not affect the magnitude of the velocity vector, but only its direction. Therefore, if gravity is neglected, our stochastic analysis remains unaffected, with distance r being replaced by arc length s. As soon as gravity is included, however, the analysis becomes more complicated [1, pp. 235–254], and we obtain a coupled pair of differ-

ential equations, as we shall presently show. Now the magnitude of the lift force is customarily given by [1, pp. 235–254]

$$L = \frac{1}{2} C_L \rho A_p v^2, \tag{38}$$

where C_L is the lift coefficient. Referring to Fig. 4, let j be a unit vector pointing upward and i a unit vector perpendicular to it, as indicated (we restrict attention here to planar motion). Let the origin of coordinates be at the initial position of the fragment, so that, at time $t = 0$, $x(0) = 0$ and $y(0) = 0$. The initial velocity components are assumed to be $v_x = v_{0x}$ and $v_y = v_{0y}$, respectively.

The vector equation of motion is then

$$-\frac{1}{2} C_D \rho A_p v^2 \cdot \frac{v_x i + v_y j}{v} + \frac{1}{2} C_L \rho A_p v^2 \cdot \frac{v_y i - v_x j}{v} - M_0 g j = M_0(\dot{v}_x i + \dot{v}_y j). \tag{39}$$

When (39) is expressed as a pair of differential equations, we have

$$M_0 \dot{v}_x = -\tfrac{1}{2} \rho A_p v (C_D v_x - C_L v_y)$$
$$M_0 \dot{v}_y = -\tfrac{1}{2} \rho A_p v (C_D v_y + C_L v_x) - M_0 g. \tag{40}$$

It is possible to solve (40) numerically for v_x and v_y, using Runge-Kutta methods [4]. Knowing the statistical distributions for C_D, C_L, and A_p, we would sample from these distributions. Then one could attempt to obtain distributions for v_x and v_y by Monte Carlo methods. Other than that, no further methodology is presently available for this general case.

In case the aerodynamic forces dominate the gravity force, one can analyze the pair $\mathbf{v} = (v_x, v_y)$ and produce reasonable approximations for the expected value of the vector and for the correlation matrix $E(\mathbf{v}\mathbf{v}^T)$. This we can do by appealing to the method of Picard for successive approximations [7, pp. 105–111]. It turns out that, if the integrand is Hölder continuous, then this method provides a sequence of approximants $\mathbf{v}_n$ which converge in the mean to the random velocity vector $\mathbf{v}$. It also follows that the expected value of $\mathbf{v}_n$ converges to the expected value of $\mathbf{v}$ [7, pp. 58–60]. The procedure is as follows: If we can reasonably neglect the gravity term in (40), then we can, as before, consider distance to be the independent variable rather than time, and the system (40) is converted into

$$\frac{dv_x}{ds} = -\rho |A_p| (C_D v_x - C_L v_y)/2M_0$$
$$\frac{dv_y}{ds} = -\rho |A_p| (C_D v_y + C_L v_x)/2M_0, \tag{41}$$

where we shall assume, for example, that

$$C_D(A_p) = a_0 + a_1 |A_p| \tag{42}$$

and

$$C_L(A_p) = a_2 + a_3 A_p, \tag{43}$$

with a_0, a_1, a_2, and a_3 to be determined experimentally. Note that we have now ascribed a sign to the presented area to conform with the fact that the sign of the lift depends on the sign of the angle of attack. The essence of our method is unchanged except that we now assume that $-A_{\max} \le A_p \le A_{\max}$ and that presented area is still a uniform random quantity. The system (41) can be written in matrix notation

$$\frac{d\mathbf{v}}{ds} = M(s)\mathbf{v}, \tag{44}$$

where

$$\mathbf{v} = \begin{pmatrix} v_x(s) \\ v_y(s) \end{pmatrix} \tag{45}$$

and

$$M(s) = -\frac{\rho |A_p|}{2M_0} \begin{pmatrix} C_D(A_p(s)) & -C_L(A_p(s)) \\ C_L(A_p(s)) & C_D(A_p(s)) \end{pmatrix}. \tag{46}$$

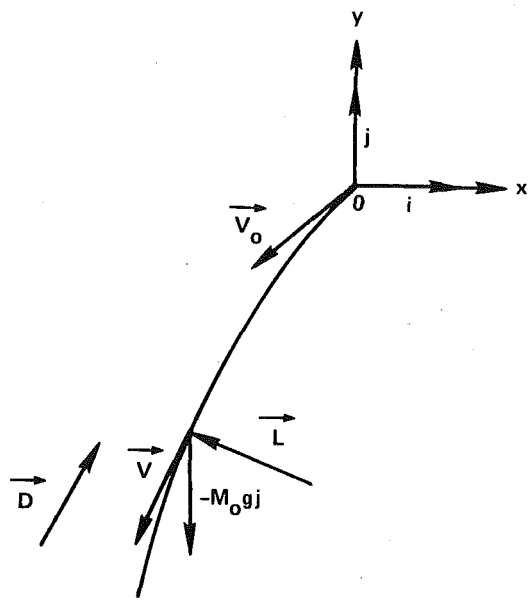


Fig. 4 Fragment trajectory history

(44) is then converted to integral form

$$\mathbf{v}(s) = \mathbf{v}_0 + \int_0^s M(x)\mathbf{v}(x)dx. \tag{47}$$

Letting $\mathbf{v}_1(x) = \mathbf{v}_0$, one then forms the sequence of approximants given by

$$\mathbf{v}_n(s) = \mathbf{v}_0 + \int_0^s M(x)\mathbf{v}_{n-1}(x)dx. \tag{48}$$

One can successively also compute the expected value of $\mathbf{v}_n(s)$ in the usual way by interchanging the operations of integration and expected value in (48) after applying the expected value operator. One can likewise compute the covariance matrix of $\mathbf{v}_n(s)$ by making use of $E(\mathbf{v}_n(s)\mathbf{v}_n{}^T(s))$ and the expected value of $\mathbf{v}_n(s)$. As mentioned before, the expected values of these approximants must converge to the desired parameters as n tends to infinity. We again can make use of the multivariate distributions which we have developed.

5 Summary

We have presented an algorithm for computing the first two statistics of the remaining speed v_r along a curved trajectory when the effect of gravity is ignored. The idea is to take account of as many moments as are necessary in order to appraise the relevant statistical parameters. The algorithm has the advantage of simplicity of implementation, since only one optimization problem need be solved and integrations over the trajectory involve only those of linear combinations of cosine functions. When gravity is included in the analysis, a coupled pair of differential equations is obtained, and it appears that the solution to the system must be obtained by Monte Carlo methods, together with Runge-Kutta procedures. In the special case where the gravity term can be neglected, we have provided a Picard approximation procedure for analyzing velocity.

6 Acknowledgments

The author would like to thank Dr. Thomas Phipps, Mr. William McDonald, and Mr. Ted Orlow, all of the Naval Surface Weapons Center, for guidance and inspiration. Their insights, especially as far as the underlying physics was concerned, were especially valuable.

References

1 Arthur, W., and Fenster, P. K., *Mechanics*, Holt, Rinehart, and Winston, Inc., 1969.

2 Ayres, F., Jr., *Schaum's Outline of Theory and Problems of Plane*

and Spherical Trigonometry, Schaum Publishing Co., New York, 1954.
3 Courant, R., *Differential and Integral Calculus,* Vol. II, Interscience Publishers, Inc., New York, 1961.
4 Ferguson, R. E. and Orlow, T. A., FNOL3, *A Computer Program to Solve Ordinary Differential Equations,* NOLTR 71-2, 1 Mar. 1971.
5 Fisz, M., *Probability Theory and Mathematical Statistics,* Wiley, New York, 1963.
6 Johnson, N. L. and Kotz, S., *Distributions in Statistics: Continuous Multivariate Distributions,* Wiley, New York, 1972.
7 Jazwinski, A. H., *Stochastic Processes and Filtering Theory,* Academic Press, New York and London, 1970.
8 Kendall, M. G. and Stuart, A., *The Advanced Theory of Statistics, Vol. 2,* Charles Griffin & Company Limited, London, 1961.
9 Mitra, S. K., "A Density-Free Approach to the Matrix Variate Beta Distribution," *Sankhyā,* Series A, Vol. 32, pp. 81–88.
10 Olkin, I., and Rubin, H., "Multivariate Beta Distributions and Independence Properties of the Wishart Distribution," *Annals of Math. Statistics,* Vol. 35, pp. 261–269.
11 Parzen, E., *Stochastic Processes,* Holden-Day, Inc., San Francisco, London; Amsterdam, 1964.
12 Widder, D. V., *Advanced Calculus,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1961.

K. C. Park
Staff Scientist.
Mem. ASME

C. A. Felippa
Staff Scientist.

Applied Mechanics Laboratory,
Lockheed Palo Alto Research Laboratory,
3251 Hanover Street,
Palo Alto, Calif. 94304

Partitioned Transient Analysis Procedures for Coupled-Field Problems: Accuracy Analysis

Partitioned solution procedures for direct time integration of second-order coupled-field systems are studied from the standpoint of accuracy. These procedures are derived by three formulation steps: implicit integration of coupled governing equations, partitioning of resulting algebraic systems and extrapolation on the right-hand partition. It is shown that the combined effect of partition, extrapolation, and computational paths governs the choice of stable extrapolators and preservation of rigid-body motions. Stable extrapolators for various computational paths are derived and implementation-extrapolator combinations which preserve constant-velocity and constant-acceleration rigid-body motions are identified. A spectral analysis shows that the primary error source introduced by a stable partition is frequency distortion. Finally, as a guide to practical applications, the advantages and shortcomings of five specific partitions are discussed.

Introduction

Dynamic problems in applied mechanics often involve two or more distinctive subsystems that are tightly coupled. The computerized analysis of such systems generally involves two distinct phases. A spatially discretized model is produced through finite-element, boundary-element or finite-difference techniques. The resulting equations of motion are then solved by direct time integration.

Often the response characteristics of the component (single-field) subsystems are markedly different. As examples of this situation we can cite fluid-structure, soil-structure, and certain structure-structure interaction problems. This difference may be due to material properties, different spatial discretization techniques, localized nonlinearities, boundary-layer effects; or be even induced by computational manipulations.

For coupled systems befitting the preceding description, it is natural to think of tailoring the time integration procedure to subsystem response features. An effective way of implementing this idea is to view each single-field subsystem as a computational entity, which is treated by an appropriate time integration program module ("analyzer"). The time-advancing process results from coordinating the execution of the analyzers in sequential or parallel fashion. Field coupling effects are incorporated through cyclic information transfers based upon appropriate extrapolation formulas. Solution procedures based on this approach were called *partitioned transient analysis procedures* (or simply *partitioned procedures*) in [1]. In that paper, a general partitioning procedure for treating second-order coupled-field equations of motion was presented.

Rational selection of a specific partitioning procedure among competing ones demands a fairly deep understanding of numerical stability and accuracy characteristics. In [1], a general stability analysis technique was developed to eliminate case-by-case considerations. This technique has provided a basis for uniform evaluation on grounds of stability, which is critical in the initial method-design stages.

Once satisfactory stability requirements are achieved, it remains to assess accuracy. Understanding of this subject has so far been based largely on numerical experiments [2–7]. This paper endeavors to fill the void by developing an accuracy assessment theory. More specifically, the following three points are addressed: (*a*) strong interdependency of accuracy, extrapolator formula, and computational sequence details; (*b*) preservation of rigid-body motions under partitioning; and (*c*) incremental effect of partitioning on numerical damping and frequency distortion measures. We view these three aspects as providing sufficient grounds for a comparative assessment of competing stable partitioned procedures.

Partitioned Integration Overview

This section provides a quick review of partitioned integration procedures for the class of mechanical systems considered in [1]. It is included to make this paper reasonably self-contained.

We consider dynamic systems governed by the second-order equations of motion

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{D}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{f} \tag{1}$$

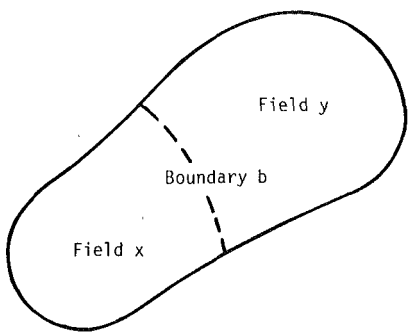
where $\mathbf{M}$, $\mathbf{D}$, and $\mathbf{K}$ are mass, damping and stiffness matrices, respectively, $\mathbf{u}$ is the state vector of generalized displacements and $\mathbf{f}$ the corresponding applied force vector; dot superscripts denote temporal differentiation.

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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, September, 1979; final revision, March, 1980.

Table 1 Examples of partitions for two-field problems

Node-by-Node	Element-by-Element	DOF-by-DOF	Staggered
$\underset{\sim}{K}_2 = \begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{K}_{yb} & \underset{\sim}{K}_{yy} \end{bmatrix}$	$\underset{\sim}{K}_2 = \begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{K}^y_{bb} & \underset{\sim}{K}_{by} \\ \underset{\sim}{0} & \underset{\sim}{K}_{yb} & \underset{\sim}{K}_{yy} \end{bmatrix}$	$\underset{\sim}{K}_2 = \begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{K}_{yy} \end{bmatrix}$	$\underset{\sim}{K}_2 = \begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{K}_{yb} & \underset{\sim}{0} \end{bmatrix}$



Coupled Field Stiffness-Matrix

$$\underset{\sim}{K} = \begin{bmatrix} \underset{\sim}{K}_{xx} & \underset{\sim}{K}_{xb} & \underset{\sim}{0} \\ \underset{\sim}{K}_{bx} & \underset{\sim}{K}^x_{bb} + \underset{\sim}{K}^y_{bb} & \underset{\sim}{K}_{by} \\ \underset{\sim}{0} & \underset{\sim}{K}_{yb} & \underset{\sim}{K}_{yy} \end{bmatrix}$$

$$\underset{\sim}{K} = \underset{\sim}{K}_1 + \underset{\sim}{K}_2$$

As a *first step* in the *partitioned treatment* of (1), the entire system is time-discretized by an m-step implicit integration formula of linear multistep type

$$\mathbf{w}_n = \delta \dot{\mathbf{w}}_n + \mathbf{h}^w_n$$

$$\mathbf{h}^w_n = \sum_{j=1}^{m} (\alpha_j \mathbf{w}_{n-j} - \delta \beta_j \dot{\mathbf{w}}_{n-j}) \tag{2}$$

In (2), $\mathbf{w}$ stands for both $\mathbf{u}$ and $\dot{\mathbf{u}}$, $\mathbf{w}_n \equiv \mathbf{w}(t_n)$ at sample times t_n; α_j and β_j are coefficients characterizing the integration formula used, δ is a generalized time stepsize and $\mathbf{h}^w_n$ is a historical vector. Insertion of (2) into (1), followed by various algebraic manipulations detailed in [8], yields the implicit difference equations:

$$\mathbf{E}\mathbf{u}_n = \mathbf{g}_n$$

$$\mathbf{E} = \mathbf{M} + \delta \mathbf{D} + \delta^2 \mathbf{K} \tag{3}$$

$$\mathbf{g}_n = \delta^2 \mathbf{f}_n + \mathbf{M}(\mathbf{h}^u_n + \delta \mathbf{h}^{\dot{u}}_n)$$

The *second step* is to partition the algebraic system (3a) as follows:

$$\mathbf{E}_1 \mathbf{u}_n = \mathbf{g}_n - \mathbf{E}_2 \mathbf{u}_n$$

$$\mathbf{E}_1 = \mathbf{M} + \delta \mathbf{D}_1 + \delta^2 \mathbf{K}_1$$

$$\mathbf{E}_2 = \delta \mathbf{D}_2 + \delta^2 \mathbf{K}_2 \tag{4}$$

$$\mathbf{D} = \mathbf{D}_1 + \mathbf{D}_2$$

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

Note that the mass matrix $\mathbf{M}$ is *not* partitioned, which is crucial to the success of these methods.

The *third and final step* is to select an extrapolator formula for the $\mathbf{u}_n$ appearing on the right-hand partition of (4a):

$$\mathbf{E}_1 \mathbf{u}_n = \mathbf{g}_n - \mathbf{E}_2 \mathbf{u}^p_n \tag{5}$$

where $\mathbf{u}^p$ denoted predicted value in terms of past computed solutions.

Specific partitions result from choosing "block patterns" for matrices $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{K}_1$, and $\mathbf{K}_2$ tailored to the problem under consideration (and perhaps to available software). To illustrate this point, Table 1 catalogs four potentially useful partitions of the stiffness matrix $\mathbf{K}$ for a coupled mechanical system consisting of two fields: x and y, which interact through a boundary b.

Effect of Computational Path and Extrapolator Selection

In reference [8] we discussed the organization of the time-advancing calculations and identified four "computational paths," which were labeled (0′), (0), (1), and (2). Paths differ according to the way in which auxiliary quantities such as velocities are calculated. In [1] it was shown that computational paths can have a significant effect on the stability of partitioned solution procedures, and that this effect is tied up with the selection of extrapolators. For examples, when the trapezoidal integration rule is used in conjunction with the last-solution extrapolator,

$$\mathbf{u}^p_n = \mathbf{u}_{n-1} \tag{6}$$

the solution remains stable if the velocity is computed by integrating the acceleration that is in turn obtained from the equations of motion. On the other hand, if the velocity is computed from the difference formula (2a), the solution becomes unstable. In terms of the nomenclature of [8], these two cases correspond to the ($C0'$) and the ($C1$)-computational path, respectively, as summarized in Table 2. [The letter prefix identifies the choice of auxiliary vector in reducing (1) to a first-order system: C = conventional, J = Jensen; this choice has no effect on stability].

Combinations of different extrapolators and computational paths will give rise in general to different characteristic equations, even if all other factors (integration formula, partition) remain the same. This has major implications as regards stability and accuracy. To study these effects, we introduce a fairly general class of extrapolators:

$$\mathbf{u}^p_n = \sum_{j=1}^{m} (\hat{\alpha}_j \mathbf{u}_{n-j} + \delta \hat{\beta}_j \dot{\mathbf{u}}_{n-j} + \delta^2 \hat{\gamma}_j \ddot{\mathbf{u}}_{n-j}) \tag{7}$$

where $\hat{\alpha}_j$, $\hat{\beta}_j$, and $\hat{\gamma}_j$ are extrapolation coefficients. This class is broader than that considered in [1], which did not account for historical derivative terms $\dot{\mathbf{u}}_{n-j}$, $\ddot{\mathbf{u}}_{n-j}$.

Table 2 Computational path-dependent formulas for partitioned procedures ($\mathbf{E}_1\mathbf{u}_n = \mathbf{g}_n - \mathbf{E}_2\mathbf{u}_n^p$)

Computational Path	Computational Sequences	Recommended Predictors, $\underline{u}_n^p$
(CO')	a $\underline{g}_n = (\underset{\sim}{M} + \delta \underset{\sim}{D})\, \underline{h}_n^u + \delta \underset{\sim}{M}\, \underline{h}_n^{\dot{u}} + \delta^2 \underline{f}_n$ b $\underline{u}_n = \underset{\sim}{E}_1^{-1} (\underline{g}_n - \underset{\sim}{E}_2\, \underline{u}_n^p)$ c $\underline{\dot{u}}_n = (\underline{u}_n - \underline{h}_n^u)/\delta$ d $\underline{\ddot{u}}_n = \underset{\sim}{M}^{-1} (\underline{f}_n - \underset{\sim}{D}\, \underline{\dot{u}}_n - \underset{\sim}{K}\, \underline{u}_n)$ e $\underline{\dot{u}}_n = \delta\, \underline{\ddot{u}}_n + \underline{h}_n^{\dot{u}}$	1) $\sum_{j=1}^{m} - \alpha_j\, \underline{u}_{n-j}$ 2) $\sum_{j=1}^{m} [-\alpha_j\, \underline{u}_{n-j} + \delta(\beta_j/\beta_o - \alpha_j)\, \underline{\dot{u}}_{n-j} + \delta^2 (\beta_j/\beta_o - \alpha_j)\underline{\ddot{u}}_{n=j}]^{\dagger}$
(C1)	a-d same as (CO'), e skipped	$\sum_{j=1}^{m} [-\alpha_j\, \underline{u}_{n-j} + \delta\, (\beta_j/\beta_o - \alpha_j)\, \underline{\dot{u}}_{n-j}$
(C2)	a-c same as (CO') d $\underline{\ddot{u}}_n = (\underline{\dot{u}}_n - \underline{h}_n^{\dot{u}})/\delta$ e skipped	$\sum_{j=1}^{m} [-\alpha_j\, \underline{u}_{n-j} + \delta\, (\beta_j/\beta_o - \alpha_j)\, \underline{\dot{u}}_{n-j} + \delta^2\, \beta_j/\beta_o\, \underline{\ddot{u}}_{n-j}]$
(JO)	a $\underline{g}_n = \underset{\sim}{M}\, \underline{h}_n^u + \delta\, \underline{h}_n^v + \alpha^2 \underline{f}_n$ b-c same as (CO') d $\underline{\dot{v}}_n = \underline{f}_n - \underset{\sim}{K}\, \underline{u}_n$ e $\underline{v}_n = \delta\, \underline{\dot{v}}_n + \underline{h}_n^v$	$\sum_{j=1}^{m} [-\alpha_j\, \underline{u}_{n-j} + \delta\, \beta_j/\beta_o\, \underline{\dot{u}}_{n-j}]$

† This extrapolator is recommended for implicit-explicit procedures when accuracy, including rigid body motions, becomes important. In this case the stability limit is somewhat reduced.

The characteristic matrix equations for the four computational sequences of Table 2 can be derived by seeking nontrivial solutions of the form

$$\mathbf{u}_k = \lambda^k \mathbf{u}_0 = \left(\frac{1+z}{1-z}\right)^k \mathbf{u}_0 \tag{8}$$

The resulting formulas are collected in Table 3. An undamped system ($\mathbf{D} = \mathbf{O}$) is assumed for simplicity; the damped case can be handled in a similar fashion; for derivation details, see Appendix *A* of [1].

Numerical stability requires that the characteristic equation roots z_i have no positive real parts, i.e., Re $(z_i) \leq 0$. This imposes in turn constraints on the selection of extrapolators for each path; recommended extrapolators are listed in Table 2. For the important case of the trapezoidal integration rule, stable extrapolators are listed in Table 4.

Remark. If partitioning is performed at the *differential* equation level as opposed to the difference equation level stressed in this paper, viz.,

$$\mathbf{M}\,\ddot{\mathbf{u}} + \mathbf{D}_1\dot{\mathbf{u}} + \mathbf{K}_1\mathbf{u} = \mathbf{f} - \mathbf{D}_2\dot{\mathbf{u}} - \mathbf{K}_2\mathbf{u} \tag{9}$$

the following characteristic equation results if $\mathbf{D} = \mathbf{O}$:

$$\det |\rho^2\mathbf{M} + \delta^2\sigma^2\mathbf{K}_1 + \delta^2(\sigma^2 e_0 + \rho\sigma e_1 + \rho^2 e_2)\mathbf{K}_2| = 0 \tag{10}$$

This is *independent* of the computational path. It can be proved that, when the trapezoidal rule is used, the differential partitioning procedure (9) is stable provided the extrapolator is chosen in the form

$$\mathbf{u}_n^{(p)} = \mathbf{u}_{n-1} + h\dot{\mathbf{u}}_{n-1} + \frac{h^2}{4}\ddot{\mathbf{u}}_{n-1} \tag{11}$$

in which $\dot{\mathbf{u}}$ and $\ddot{\mathbf{u}}$ are computed by differentiation formulas regardless of the implementation paths adopted. The above extrapolator for the differential partitioning procedure (9) is identical to that of path (*C*2) for the algebraic partitioning procedure (4) (see Table 4). It therefore appears as if partitioning at the difference equation level provides superior flexibility of implementation because we can then choose various computational paths while taking care of stability through an appropriate extrapolator. This flexibility is lost in differential partitioning.

Preservation of Rigid-Body Motions

The spatial convergence of nonconforming finite-element models has been traditionally probed with Irons' patch test [9], which among other things verifies whether rigid-body displacement fields are preserved. Similarly, when considering candidate partitioned integration procedures for structures experiencing large rigid-body motions, it is important to understand which combinations of partition, extrapolator and computational path preserve such motions.

The maximum order of accuracy attainable with *A*-stable integration formulas of the type (2) is two (Dahlquist's theorem). This means that constant-velocity and constant-acceleration rigid-body motions can be traced exactly if the system is treated as one entity. What happens if a partition is introduced? Then that formal order of accuracy is generally lost for solution components belonging to a partition boundary, unless special safeguards are taken. It turns out that the combined characteristics of the computational paths and the predictor forms play a decisive role for the preservation of rigid-body motions. This aspect is examined in detail for the important case of the trapezoidal rule.

Rigid-body motions $\mathbf{u}$ of a flexible structure satisfy

$$\mathbf{K}\mathbf{u} = \mathbf{0} \tag{12}$$

Introducing the partition (4*e*) yields

$$\mathbf{K}_1\mathbf{u} = -\mathbf{K}_2\mathbf{u} \tag{13}$$

Table 3 Computational path-dependent characteristic equations (D = O)

Computational Path	Characteristic Equation
(CO')	$\lvert \rho^2 \underset{\sim}{M} + \delta^2 \sigma^2 \underset{\sim}{K}_1 + \delta^2 (e_0 \rho + \sigma^2 - \rho \lambda^m) \underset{\sim}{K}_2 - \delta^4 (e_1 \sigma + e_2 \rho) \underset{\sim}{K}_2^{-1} \underset{\sim}{K} \rvert = 0$
(C1)	$\lvert \rho^2 \underset{\sim}{M} + \delta^2 \sigma^2 \underset{\sim}{K}_1 + \delta^2 (\sigma e_0 + \sigma^2 - \sigma \lambda^m + \rho e_1) \underset{\sim}{K}_2 - \delta^4 e_2 \underset{\sim}{K}_2 \underset{\sim}{M}^{-1} \underset{\sim}{K} \rvert = 0$
(C2)	$\lvert \rho^2 \underset{\sim}{M} + \delta^2 \sigma^2 \underset{\sim}{K}_1 + \delta^2 (\sigma^2 e_0 + \rho \sigma e_1 + \rho^2 e_0) \underset{\sim}{K}_2 \rvert = 0$
(JO)	$\lvert \lambda^m \rho^2 \underset{\sim}{M} + \delta^2 \lambda^m \sigma^2 \underset{\sim}{K}_1 + \delta^2 (\sigma^2 \lambda^m + e_0 \rho \sigma + e_1 \rho^2) \underset{\sim}{K}_2 - \delta^4 \sigma e_2 \underset{\sim}{K}_2 \underset{\sim}{M}^{-1} \underset{\sim}{K} \rvert = 0$

In the above equations, ρ, σ, e_1 and e_2 are given by:

$$\begin{cases} \rho(\lambda) = \sum_{j=0}^{m} \alpha_j \lambda^{m-j} \\ \sigma(\lambda) = \sum_{j=0}^{m} \beta_j \lambda^{m-j} \\ e_0 (e_1, e_2) = \sum_{j=1}^{m} \hat{\alpha}_j (\hat{\beta}_j, \hat{\gamma}_j) \lambda^{m-j} \end{cases}$$

$$\lambda = (1 + z)/(1 - z)$$

Table 4 Stable extrapolators for trapezoidal rule integration

Computational Path	Stable Extrapolator
(CO')	$\underline{u}_n^p = \begin{cases} \underline{u}_{n-1}{}^* \\ (\underline{u}_{n-1} + h\, \dot{\underline{u}}_{n-1} + \frac{h^2}{2} \ddot{\underline{u}}_{n-1})^\dagger \end{cases}$
(C1)	$\underline{u}_n^p = \underline{u}_{n-1} + h\, \dot{\underline{u}}_{n-1}$
(C2)	$\underline{u}_n^p = \underline{u}_{n-1} + h\, \dot{\underline{u}}_{n-1} + \frac{h^2}{4} \ddot{\underline{u}}_{n-1}$
(JO)	$\underline{u}_n^p = \underline{u}_{n-1} + \frac{h}{2} \dot{\underline{u}}_{n-1}$

* This extrapolator is extensively used in [1].

† Applicable strictly to implicit-explicit partitions with somewhat reduced stability limits, e.g., $\omega_{max} \cdot h \leq 1.57$ for DOF-by-DOF I - E partitions.

Integrator: $\underline{u}_n = \underline{u}_{n-1} + \delta (\dot{\underline{u}}_n + \dot{\underline{u}}_{n-1}), \ \delta = \frac{h}{2}$

Extrapolator Form: $\underline{u}_n^p = \underline{u}_{n-1} + \hat{\beta} \delta \dot{\underline{u}}_{n-1} + \hat{\gamma} \delta^2 \ddot{\underline{u}}_{n-1}$

Now equation (13) implies that

$$(\mathbf{M} + \delta^2 \mathbf{K}_1)\, \mathbf{u} = (\mathbf{M} - \delta^2 \mathbf{K}_2) \mathbf{u} \tag{14}$$

If computational path (0′) is followed, the following difference equation holds (cf. Table 2):

$$(\mathbf{M} + \delta^2 \mathbf{K}_1) \mathbf{u}_n = \delta^2 \mathbf{f}_n + \mathbf{M}(\mathbf{u}_{n-1} + 2\delta \dot{\mathbf{u}}_{n-1} + \delta^2 \ddot{\mathbf{u}}_{n-1}) - \delta^2 \mathbf{K}_2 \mathbf{u}_{n-1}^p \tag{15}$$

For constant-acceleration rigid-body motion, we must have

$$\mathbf{M} \ddot{\mathbf{u}}_k = \mathbf{f}_k, \quad k = 0, 1 \ldots n \tag{16}$$

Inserting (16) into (15) and accounting for (14), it follows that preservation of rigid-body motions demands that either $\mathbf{K}_2 \mathbf{u}_n \equiv \mathbf{0}$ or that the following predictor be used:

$$\mathbf{u}_n^p = \mathbf{u}_{n-1} + 2\delta \dot{\mathbf{u}}_{n-1} + 2\delta^2 \ddot{\mathbf{u}}_{n-1} \tag{17}$$

This is the second extrapolator formula given for path (0′) in Table 3.

Remark 1. If one insists only on exact preservation of constant-velocity rigid body motions, for which $\ddot{\underline{u}}_n \equiv \underline{0}$, the following predictors can be used:

$$\begin{aligned} \mathbf{u}_n^p &= \mathbf{u}_{n-1} + 2\delta \dot{\mathbf{u}}_{n-1} \\ \mathbf{u}_n^p &= \mathbf{u}_{n-1} + 2\delta \dot{\mathbf{u}}_{n-1} + \delta^2 \ddot{\mathbf{u}}_{n-1} \end{aligned} \tag{18}$$

Remark 2. Both the element-by-element and node-by-node partitions preserve any rigid-body motions because these two partitions satisfy the rigid-body conserving identity

$$\mathbf{K}_1 \mathbf{u}_n = \mathbf{K}_2 \mathbf{u}_n = \mathbf{O} \tag{19}$$

For these partitions, the last-solution extrapolator $\mathbf{u}_n^p = \mathbf{u}_{n-1}$ is optimally stable for path (0′) and preserves rigid-body motions. For other computational paths, the equivalent predictors listed in Table 2 retain these properties.

Remark 3. The last-solution extrapolator, although stable, distorts constant velocity and constant-acceleration rigid body motions for the DOF-by-DOF and staggered partitions. This effect is illustrated in Fig. 1 for a two-degree-of-freedom problem. This distortion can only be eliminated by iterating at each time step to convergence.

Remark 4. The accurate extrapolator (17) gives rise to numerical

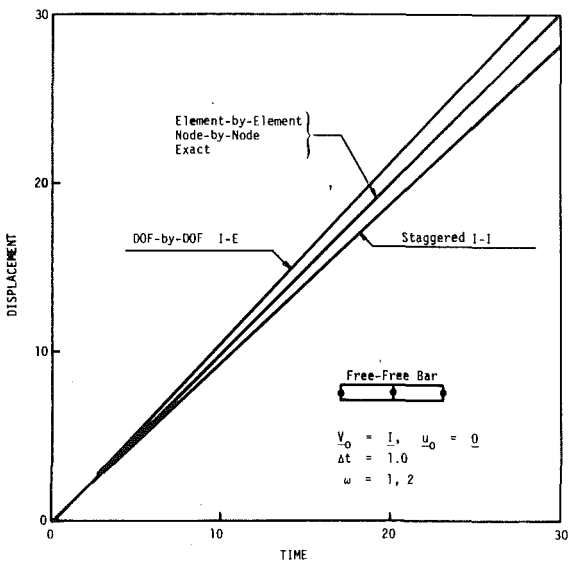


Fig. 1 Distortion of constant-velocity rigid-body response of free-free bar for the DOF-by-DOF and staggered partitions

instability for paths other than (0′). Thus stability and accuracy (as far as preservation of rigid-body motions is concerned) are not generally equivalent attributes for partitioned integration.

The main findings of this study are summarized in Table 5. Clearly path (0′) is the choice when accurate representation of constant-acceleration rigid-body motions (e.g., free-falling structures) is important, and a DOF-by-DOF partition is used. The price paid for this formulation is the need for additional calculations, such as solving for accelerations, in the advancing step [8]. If the accuracy requirement is relaxed to preservation of constant-velocity rigid-body motions, then either path (1) or (2) is acceptable for those partitions (because predictors (18*a*) and (18*b*) are stable for those paths, cf. Table 2), and the computational effort is accordingly reduced. Finally, path (0) distorts rigid-body motions other than constant-displacement for the DOF-by-DOF and staggered partitions.

So far we have focused our attention on the construction of stable and accurate extrapolators and on conditions for preserving rigid-body motions. The remainder of the paper will be devoted to the examination of the "incremental" effect of partition on algorithmic accuracy.

Effect of Partitioning on Algorithmic Accuracy

The conventional Fourier technique for assessing accuracy of integration formulas for second-order systems proceeds as follows. First, the homogeneous difference equations are decoupled by projection on normal coordinates. Second, the frequency distortion and numerical damping of an uncoupled difference equation subjected to harmonic input of frequency ω are presented as functions of the normalized sampling frequency ωh. This procedure is not immediately applicable to the partitioned difference equations (4*b*), however, because these are not generally diagonalized by the normal modes of the semidiscrete equations (1).

We shall instead use a limit differential equation approach to assess the frequency-dependent accuracy of various partitioned procedures, assuming that the solutions can be expanded in Taylor series up to the order of the integration formula. This approach has been extensively used for the evaluation of artificial viscosity [10,11]. It will be shown that the limit differential equations for the partitioned difference equations display a frequency distortion effect as primary error source due to partitioning.

When the trapezoidal formula is used in conjunction with path (*C*0′) and the previous solution vector $\mathbf{u}_{n-1}$ is used as the extrapolator, the following limit differential equation results for $\mathbf{D} = \mathbf{O}$:

$$\left(\mathbf{M} - \frac{2}{3}\delta^2\mathbf{K}_1 + \frac{1}{3}\delta^2\mathbf{K}_2\right)\ddot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{0}(\delta^3) \tag{19}$$

Note that the mass matrix is considerably modified whereas the stiffness matrix is not, and also the absence of perturbation terms associated with the velocity vector $\dot{\mathbf{u}}$. We therefore conclude that the primary "incremental" algorithmic error caused by partitioning is

Table 5 Test on rigid-body motion preservation (trapezoidal rule)

Procedure	Computational Paths			
	(C0′)	(C1)	(C2)	(J0)
Node-by-Node Implicit-Explicit	Any†	Any	Any	Any
Element-by-Element Implicit-Explicit	Any	Any	Any	Any
DOF-by-DOF Implicit-Explicit	A	V	V	D
Staggered Implicit-Implicit	A	V	V	D
Element-by-Element Implicit-Implicit	A	A	A	A

† Any = any rigid-body motion; A = preserves constant acceleration (free fall); V = preserves constant velocity; D = distorts all but constant displacement.

Remark: For all computational paths, only stable predictors are used (see Table 1 or 2).

manifested in the frequency distortion. The amount of distortion can be obtained from the limit characteristic equation,

$$\det\left|\left(\mathbf{M} - \frac{2}{3}\delta^2\mathbf{K}_1 + \frac{1}{3}\delta^2\mathbf{K}_2\right)\Omega^2 - \delta^2\mathbf{K}\right| = 0, \quad \Omega = \omega\delta \tag{20}$$

The two limit cases for (20) are: $\mathbf{K}_2 = \mathbf{0}$ (fully implicit integration) and $\mathbf{K}_2 = \mathbf{K}$ (fully explicit integration). These correspond in turn to the trapezoidal rule and central difference formula, respectively. It is well known that these two formulas introduce no numerical damping. Moreover, the central difference formula shortens the period whereas the trapezoidal rule elongates it by an amount roughly twice as big; this can be immediately deduced from (20). That the frequency distortion is also the primary algorithmic error due to partitioning can also be shown to be true for other, numerically damped, integration formulas. (The demonstration relies on the fact that nonzero odd-derivative residual terms in the Taylor series occur beyond the order of the integration formula truncation error.)

As noted previously, the matrix equation (19) is not generally diagonalizable by the natural modes of the original equations of motion. To assess the accuracy performance of various partitions we can exploit, however, the fact that the frequency error due to partitioning emanates from partition boundaries. This enables us to construct a two-degree-of-freedom model system from which the magnitude of this error as a function of the time stepsize can be appraised. The free-vibration equations of this system are

$$\begin{bmatrix} 1+\alpha & 0 \\ 0 & 1 \end{bmatrix}\ddot{\mathbf{u}} + \begin{bmatrix} 1+\dfrac{1}{\alpha} & -1 \\ -1 & 1 \end{bmatrix}\mathbf{u} = \mathbf{0} \tag{21}$$

with appropriate initial conditions. Equations (21) can be interpreted as axial equations of motion for a fixed-free bar discretized into two elements whose length ratio is $1{:}\alpha$.

For the foregoing two-DOF system the partition-caused artificial frequency distortion can be asymptotically estimated (for $\Omega \ll 1$) from equation (20). Alternatively, the distortion can be computed exactly from the characteristic equation for the partitioned difference equations (see Table 3). As the calculations by the two approaches have shown no discernible difference, we present the results obtained by the latter approach.

Figs. 2(a) and 2(b) show the frequency distortion of the low and high frequency components for the equal-length case $\alpha = 1$. For this case the node-by-node implicit-explicit partition is the most accurate, followed by the element-by-element, the DOF-by-DOF, and the staggered partition. As the element length ratio (roughly square of the frequency ratio) is decreased, however, an intermediate value of α is reached at which the element-by-element procedure becomes more accurate than the node-by-node procedure (see Fig. 3). If the length ratio is further decreased, the low (high) frequency error approaches that of fully explicit (implicit) formula, respectively, as evidenced by Fig. 4. The increase in the frequency ratio weakens the coupling effect between the two modes.

From the foregoing accuracy analysis of the model system it can be seen that accuracy of solution components far from the partition boundaries is largely controlled by that of the integration formulas used. For solution components at or near the partition boundaries, the element-by-element and the node-by-node partitions enjoy a somewhat higher accuracy than either the explicit or the implicit formula itself in the sense that the frequency distortions are less than those of the fully explicit (implicit) integration formulas (see Figs. 2 and 3). On the other hand, for the DOF-by-DOF and the staggered partitions the accuracy of the boundary solution components is inferior to that of the fully explicit (implicit) formulas.

Conclusions

The material covered herein complements an earlier paper [1], which was devoted to the general categorization of partitioned integration procedures. The present paper expands on stability and accuracy as properties strongly related to the choice of extrapolation

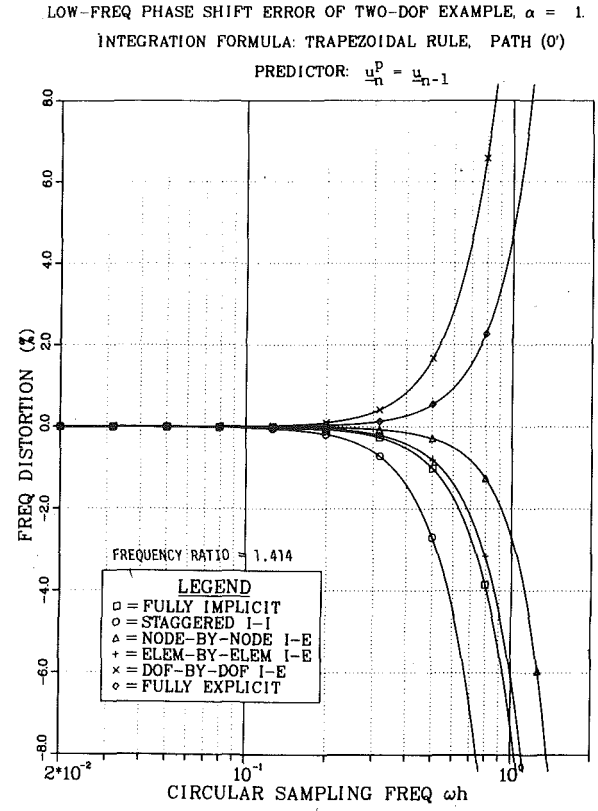


Fig. 2(a) Distortion of low-frequency root of model problem (21) as function of sampling frequency, for α = 1.0, trapezoidal integration rule with last-solution extrapolator, and specific partitions

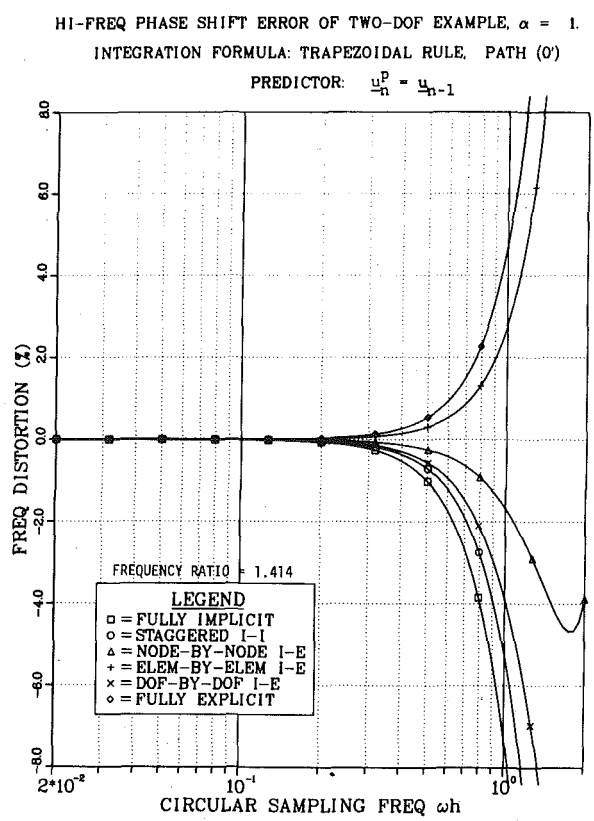


Fig. 2(b) Distortion of high-frequency root of model problem (21) as function of sampling frequency, for α = 1.0, trapezoidal integration rule with last-solution extrapolator, and specific partitions

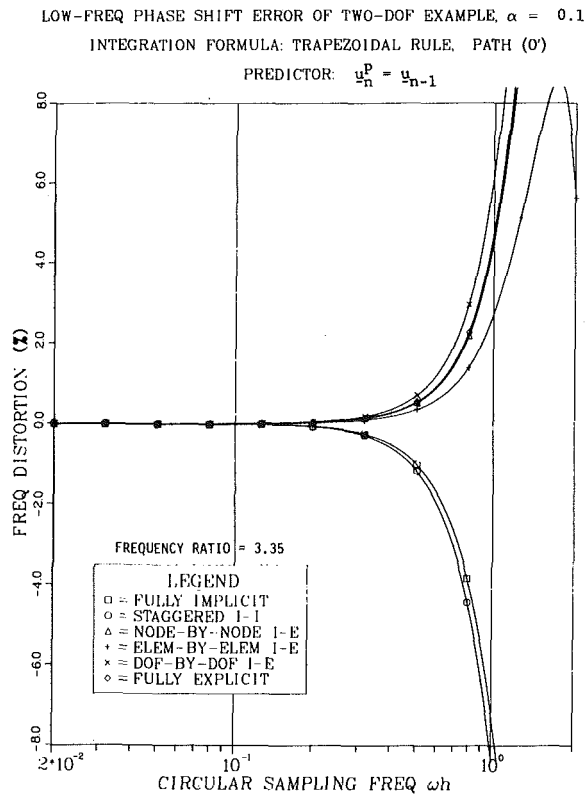


Fig. 3(*a*) Distortion of low-frequency root of model problem (21) as function of sampling frequency, for $\alpha = 0.1$, trapezoidal integration rule with last-solution extrapolator, and specific partitions

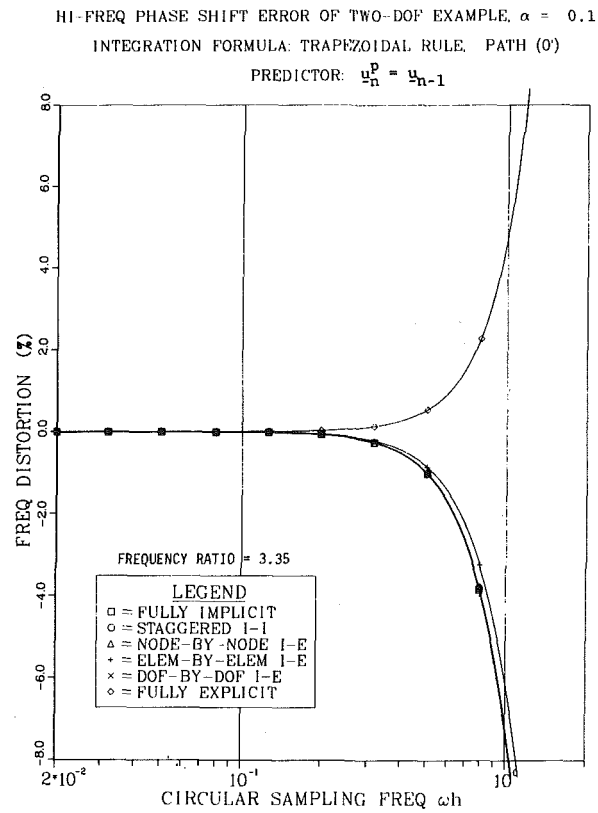


Fig. 3(*b*) Distortion of high-frequency root of model problem (21) as function of sampling frequency, for $\alpha = 0.1$, trapezoidal integration rule with last-solution extrapolator, and specific partitions

formula and computation path; two aspects that were only briefly commented upon in [1]. The information presented here may be used as an initial guide in method selection process. It should be noted, however, that the selection of a partitioned procedure over another may be based on nonalgorithmic considerations. Thus the availability of computational tools for treating certain subsystems of the overall problem (e.g., finite-element structural analyzers, boundary-integral analyzers for infinite domains, etc.) may naturally dictate the use of certain partitions.

We have summarized in Table 6 general algorithmic characteristics of five partitioned solution procedures. From an unconditional stability viewpoint the implicit-implicit partitions (staggered and element-by-element partitions) would be preferred. This would be the case if the envisioned integration stepsize is large compared to the shortest characteristic time of each subsystem, as usually is the case in tracing late-time responses.

Implicit-explicit partitions are desirable if subsystems display widely different response characteristics and stepsizes are comparable with shortest characteristic times of one or more subsystems. They offer the advantage of computational simplicity for the explicitly treated subsystems. From an applications viewpoint the element-by-element partition appears most natural for finite-element discretizations, but can become cumbersome for extensively connected partitions (e.g., volume-coupled finite-element models). The node-by-node partition appears to require additional implementation effort, as "boundary nodes" must be appropriately labeled. It becomes attractive, however, when the number of boundary unknowns is considerable. The DOF-by-DOF partition, although less accurate than the previous ones, may be advantageous when the explicit-implicit partition is to occur within each element or node point of a finite-element or finite-difference model. An example would be the computational separation of translational and rotational degrees of freedom in lumped-mass finite-element models.

We now summarize our main findings.

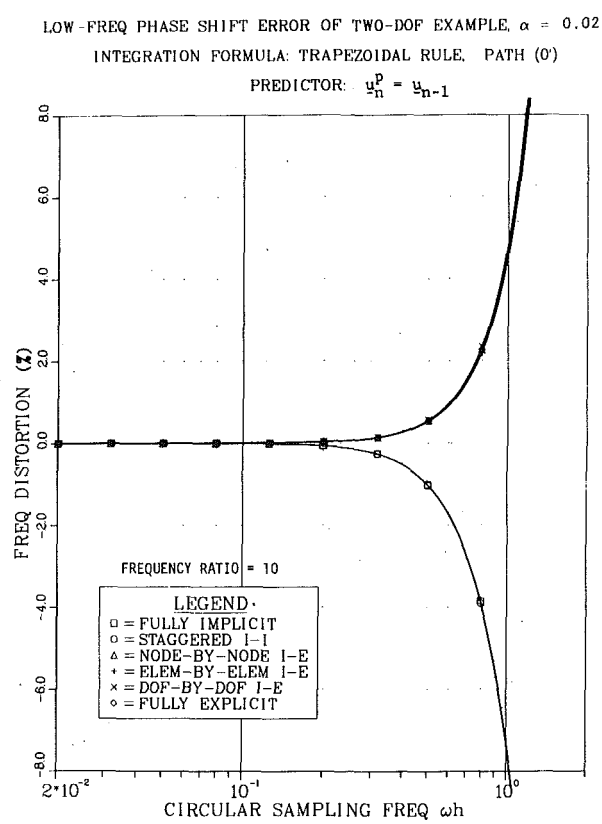


Fig. 4 Distortion of low-frequency root of model problem (21) as function of sampling frequency, for $\alpha = 0.02$, trapezoidal integration rule with last-solution extrapolator, and specific partitions

Table 6 Attributes of different partitioned procedures

Partitioned Procedure	Boundary Rigid-Body Motions	Stability† (See Ref. 1)	Remarks
Element-by-Element I-E	Exact	$2/\omega_{HL}$	o Natural partition for coupled finite element programs o Inefficient for problems with many boundary unknowns
Node-by-Node I-E	Exact	$2/\omega_{BM}$	o Efficient for problems with many boundary unknowns o Modularization impaired because of implementation difficulties within implicit integrator
DOF-by-DOF I-E	Exact only for Path (CO')	$2/\omega_{DOF}$	o Allows degree-by-degree implicit (explicit) integration within each finite element
Staggered I-I	Inexact	Stable	o Accuracy loss may be a problem o Attractive for coupling finite element and boundary integral codes
Element-by-Element I-I	Exact	Stable	o No loss of accuracy due to partitioning o Natural partition for coupled finite element program

ω_{HL}, ω_{BM}, ω_{DOF} designate the maximum frequency of the explicitly partitioned domains.

For a given integration formula, stability and accuracy of partitioned integration procedures are affected by the interaction of the computational path and selected extrapolator. The user is free to select partition-type, computational path, and extrapolator from many possible combinations. Such versatility is an important attribute of the present formulation: implicit time-discretization, algebraic partitioning, and extrapolation. It is not shared by the "differential partitioning" formulation: partitioning at differential equation level, extrapolation, and time discretization, which was used in [2–6].

Possible distortion of rigid-body motions due to partitioning can be prevented for most partitions if the computational path (CO') in conjunction with a stable extrapolator is adopted. This is not the case when the differential partitioning approach is used except in the case of element-by-element and node-by-node partitions.

The primary error due to partitioning is manifested in frequency distortion for solution components adjacent to a partition boundary. Partition-caused numerical damping is of secondary importance. In particular, for integration formulas which possess no numerical damping, such as the trapezoidal and the central difference formulas, partitioning introduces no additional numerical damping if the extrapolator is judiciously selected.

Acknowledgments

The authors wish to acknowledge support of this work by the Office of Naval Research under Contract N00014-74-C-0355 and by the Independent Research Program of Lockheed Missiles and Space Company, Inc.

References

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

2 Belytschko, T., and Mullen, R., "Mesh Partitions of Explicit-Implicit Time Integration," *Formulations and Computational Algorithms in Finite Element Analysis*, ed., Bathe, J., et al., MIT Press, 1977, pp. 673–690.

3 Belytschko, T., and Mullen, R., "Stability of Explicit-Implicit Mesh Partitions in Time Integration," *International Journal Num. Meth. In Engr.*, Vol. 12, 1978, pp. 1575–1586.

4 Park, K. C., Felippa, C. A., and DeRuntz, J. A., "Stabilization of Staggered Solution Procedures for Fluid-Structure Interaction Analysis," in *Computational Methods for Fluid-Structure Interaction Problems*, eds, Belytschko, T., and Geers, T. L., ASME Applied Mechanics Symposia Series, AMD-Vol. 26, 1977, pp. 94–124.

5 Hughes, T. J. R., and Liu, W. K., "Implicit-Explicit Finite Elements in Transient Analysis: Stability Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 371–374.

6 Hughes, T. J. R., and Liu, W. K., "Implicit-Explicit Finite Elements in Transient Analysis: Implementation and Numerical Examples," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 375–378.

7 Hughes, T. J. R., Private Communication, Mar. 1979.

8 Felippa, C. A., and Park, K. C., "Computational Aspects of Time Integration Procedures in Structural Dynamics: Implementation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 595–602.

9 Irons, B. M., and Razzaque, A., "Experience With the Patch Test for Convergence of Finite Elements," *Proceedings of the Symposium on the Mathematical Foundation of the Finite-Element Method With Applications to Partial Differential Operators*, Aziz, A. K., ed., Academic Press, New York, 1972, pp. 557–587.

10 von Neumann, J., and Richtmeyer, R. D., "A Method for the Numerical Calculations of Hydrodynamical Shocks," *Journal of Applied Physics*, Vol. 21, 1950, pp. 232–257.

11 Roache, P. J., "On Artificial Viscosity," *Journal of Comp. Physics*, Vol. 10, 1972, pp. 169–184.

D. J. Inman[1]
Assistant Professor,
Department of Mechanical Engineering,
State University of New York
at Buffalo,
Amherst, N. Y. 14260
Assoc. Mem. ASME

A. N. Andry, Jr.
Associate Professor,
Department of Mechanical Engineering,
Michigan State University,
E. Lansing, Mich. 48824

Some Results on the Nature of Eigenvalues of Discrete Damped Linear Systems

An analysis of the conditions under which the modes of a damped linear lumped parameter system are either all critically damped, overdamped or underdamped is presented. These conditions are derived from the definiteness of certain combinations of the coefficient matrices. Some results concerning the completeness of the eigenvectors are stated. The conditions are compared to previous results and their usefulness is illustrated by numerical examples.

Introduction

In the case of a single-degree-of-freedom damped linear system described by the scalar equation

$$m\ddot{x} + c\dot{x} + kx = 0$$

where m, c, and k are the mass, viscous damping constant and spring stiffness, respectively, it is well known that the nature of the solution is determined by the damping ratio

$$\zeta = \frac{1}{2}\left(\frac{c}{m}\right)\left(\frac{m}{k}\right)^{1/2}.$$

This ratio is then used to characterize critical damping ($\zeta = 1$), overdamping ($\zeta > 1$), and underdamping ($\zeta < 1$). Thus the nature of the solution is known by examination of the coefficients c/m and k/m without solving the differential equation. The intent of this paper is to provide similar criteria for multidegree-of-freedom systems described by the matrix differential equation

$$M\ddot{\mathbf{x}} + C\dot{\mathbf{x}} + K\mathbf{x} = 0 \tag{1}$$

where M, C, and K are $n \times n$ symmetric matrices and $\mathbf{x}$ is an n-dimensional vector. It is assumed that M and K are positive-definite and C is positive-semidefinite.

In previous work Duffin [1] defined an overdamped system in terms of a function of the quadratic forms of the coefficient matrices. More recently, Nicholson [2] defined an underdamped system in terms of the eigenvalues of the mass, damping, and stiffness matrices. Müller [3] responded to Nicholson's work and defined an underdamped system in terms similar to Duffin's and derived a sufficient condition in terms of the definiteness of the coefficient matrices.

The conditions of Duffin and Nicholson require substantial calculations to check. What is offered here are uniform definitions of the four possible classifications of viscous damping and some conditions involving the coefficient matrices which require less computation. Also, a comparison is made between the results stated here and those of Duffin and Nicholson. It is shown that the result here is equivalent to that of Müller's in the special case when the damped system is diagonalized by the undamped modal matrix.

Definitions

For simplicity let us consider the transformation $\mathbf{x} = M^{-1/2}\mathbf{y}$, where $M^{1/2}$ denotes the positive-definite square root of the positive-definite matrix M. Then equation (1) becomes

$$\ddot{\mathbf{y}} + \tilde{C}\dot{\mathbf{y}} + \tilde{K}\mathbf{y} = 0 \tag{2}$$

where $\tilde{C} = M^{-1/2}CM^{-1/2}$ is positive-semidefinite and $\tilde{K} = M^{-1/2}KM^{-1/2}$ is positive-definite. Note that $\tilde{C}$ and $\tilde{K}$ reflect the geometry of the system as well as the values of the system parameters. As usual the eigenvalues of (2) are taken to be the $2n$ roots of the polynomial equation

$$|\lambda^2 I + \lambda\tilde{C} + \tilde{K}| = 0 \tag{3}$$

where $|\cdot|$ denotes the determinant. More precisely they are the $2n$ latent roots of the λ-matrix [4]

$$D_2(\lambda) = \lambda^2 I + \lambda\tilde{C} + \tilde{K}.$$

We will further assume that the system is asymptotically stable in the sense that all motions exponentially decay to zero.

[1] Formerly, Instructor, Department of Mechanical Engineering, Michigan State University, E. Lansing, Mich. 48824.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, December, 1979; final revision, March, 1980. Paper No. 80-WA/APM-19.

Following the definitions stated for the single-degree-of-freedom system we choose to define the various types of viscous damping in terms of the critical damping matrix defined and denoted by $C_c = 2\tilde{K}^{1/2}$ as follows:

Critical Damping: The system described by equation (2) is critically damped if $\tilde{C} = C_c$.

Overdamping: The system described by equation (2) is overdamped if $\tilde{C} - C_c$ is positive-definitive.

Underdamping: The system described by equation (2) is underdamped if $C_c - \tilde{C}$ is positive-definite.

Mixed Damping: The system described by equation (2) is said to have mixed damping if $\tilde{C} - C_c$ is indefinite.

Results

With these definitions we will show that, based on the definiteness of $\tilde{C} - C_c$, statements can be made about the nature of the eigenvalues (and hence the solutions) which are direct analogies to the sign of ζ-1 in the one-degree-of-freedom case.

Critically Damped Systems. We will first show that if equation (2) is critically damped then there are at most n distinct negative real roots of equation (3) and no complex roots. Hence each of the modes of the damped system behaves in a critically damped manner.

To illustrate this let $\tilde{C} = 2\tilde{K}^{1/2} = C_c$ and relabel $\mathbf{y}$ as $\mathbf{x}$ in equation (2) and obtain an expression of the form

$$\ddot{\mathbf{x}} + 2\tilde{K}^{1/2}\dot{\mathbf{x}} + \tilde{K}\mathbf{x} = 0. \tag{4}$$

Let S be the orthogonal undamped modal matrix and let $\mathbf{x} = S\mathbf{y}$. Then (4) becomes

$$\ddot{\mathbf{y}} + 2S^T\tilde{K}^{1/2}S\dot{\mathbf{y}} + S^T\tilde{K}S\mathbf{y} = 0. \tag{5}$$

Now $S^T\tilde{K}S = \Lambda$ is a diagonal matrix with positive entries along the diagonal. From matrix theory [5] we have

$$S^T\tilde{K}^{1/2}S = \Lambda^{1/2}, \tag{6}$$

which is also a diagonal matrix. Equation (5) now becomes

$$\ddot{\mathbf{y}} + 2\Lambda^{1/2}\dot{\mathbf{y}} + \Lambda\mathbf{y} = 0, \tag{7}$$

which is a diagonal system. The ith equation of (7) is

$$\ddot{y}_i + (2\Lambda^{1/2})_{ii}\dot{y}_i + (\Lambda)_{ii}y_i = 0, \tag{8}$$

where A_{ij} denotes the i-jth element of the matrix A. Hence the discriminant of the characteristic equation for equation (8) is

$$(2\Lambda^{1/2}{}_{ii})^2 - 4\Lambda_{ii} = 0 \tag{9}$$

Therefore each root of (3) is a repeated negative real number. Since similarity transformations such as S preserve eigenvalues we have that (2) has at most n negative real eigenvalues and no complex eigenvalues.

We also note here that since $C_c = 2\tilde{K}^{1/2}$ is diagonalized by S, a necessary condition for (2) to be critically damped is for the damping matrix $\tilde{C}$ to be diagonalized by the undamped modal matrix, that is for $\tilde{C}\tilde{K} = \tilde{K}\tilde{C}$ [6].

Underdamped Systems. Next we will show that if equation (2) is underdamped the eigenvalues of (2) are all complex and appear in conjugate pairs with negative real parts. This corresponds to all modes of the system oscillating in damped harmonic motion.

To show that this is the case we proceed from the definition noting that from [4, p. 10] if the matrix A is real, symmetric, and positive-definite then $\bar{\mathbf{q}}^T A\mathbf{q} > 0$ for all nonzero complex vectors $\mathbf{q}$. Hence the definition of underdamping demands that

$$\bar{\mathbf{x}}^T(2\tilde{K}^{1/2} - \tilde{C})\mathbf{x} > 0$$

for all nonzero complex vectors $\mathbf{x}$. This yields

$$2\bar{\mathbf{x}}^T\tilde{K}^{1/2}\mathbf{x} > \bar{\mathbf{x}}^T\tilde{C}\mathbf{x}. \tag{10}$$

The matrix $\tilde{C}$ is taken to be positive-semidefinite so squaring (10) yields

$$4(\bar{\mathbf{x}}^T\tilde{K}^{1/2}\mathbf{x})^2 > (\bar{\mathbf{x}}^T\tilde{C}\mathbf{x})^2 \tag{11}$$

The Cauchy-Schwarz inequality with norm and inner product defined via the usual scalar product of two vectors is

$$(\bar{\mathbf{q}}^T\mathbf{r})^2 \le \bar{\mathbf{q}}^T\mathbf{q}\,\bar{\mathbf{r}}^T\mathbf{r}.$$

With $\mathbf{q} = \mathbf{x}$ and $\mathbf{r} = \tilde{K}^{1/2}\mathbf{x}$ we obtain, since $\tilde{K}^{1/2}$ is real symmetric,

$$(\bar{\mathbf{x}}^T\tilde{K}^{1/2}\mathbf{x})^2 \le \bar{\mathbf{x}}^T\mathbf{x}\,\bar{\mathbf{x}}^T\tilde{K}\mathbf{x}. \tag{12}$$

Combining inequalities (11) and (12) yields

$$(\bar{\mathbf{x}}^T\tilde{C}\mathbf{x})^2 - 4\bar{\mathbf{x}}^T\mathbf{x}\,\bar{\mathbf{x}}^T\tilde{K}\mathbf{x} < 0. \tag{13}$$

If we multiply $D_2(\lambda)$ on the right by the eigenvector $\mathbf{x}$ and the left by $\bar{\mathbf{x}}^T$ we obtain

$$\lambda^2\bar{\mathbf{x}}^T\mathbf{x} + \lambda\bar{\mathbf{x}}^T\tilde{C}\mathbf{x} + \bar{\mathbf{x}}^T\tilde{K}\mathbf{x} = 0, \tag{14}$$

which is a scalar equation in λ with solution given by

$$2\bar{\mathbf{x}}^T\mathbf{x}\lambda = -\bar{\mathbf{x}}^T\tilde{C}\mathbf{x} \pm [(\bar{\mathbf{x}}^T\tilde{C}\mathbf{x})^2 - 4\bar{\mathbf{x}}^T\mathbf{x}\bar{\mathbf{x}}^T\tilde{K}\mathbf{x}]^{1/2}. \tag{15}$$

Inequality (13) implies that the discriminant in (15) is negative. All the quadratic forms in (15) are real numbers, hence all the eigenvalues have nonzero imaginary parts and appear in conjugate pairs. This also shows, of course, that all the eigenvectors of an underdamped system are complex. The real part of λ is negative or zero via the definiteness condition on $\tilde{C}$. Zero is excluded as a possibility by the assumption of asymptotic stability. For a convenient criterion for asymptotic stability, see Walker and Schmitendorf [7].

Overdamped Systems. If equation (2) is overdamped then the eigenvalues of equation (2) are all negative real numbers, and none of the modes oscillate.

To see this result we note that if $\tilde{C} - 2\tilde{K}^{1/2}$ is positive-definite then there exists a positive-definite matrix P and positive scalar ϵ_0 such that

$$\tilde{C} = 2\tilde{K}^{1/2} + \epsilon_0 P \tag{16}$$

Motivated by the discriminant in equation (15) we define the scalar form

$$D = (\bar{\mathbf{y}}^T\tilde{C}\mathbf{y})^2 - 4\bar{\mathbf{y}}^T\mathbf{y}\,\bar{\mathbf{y}}^T\tilde{K}\mathbf{y}$$

for all nonzero complex vectors $\mathbf{y}$. Substituting (16) into this form we define the scalar function of the scalar variable ϵ_0, by

$$D(\epsilon_0) = [\bar{\mathbf{y}}^T(2\tilde{K}^{1/2} + \epsilon_0 P)\mathbf{y}]^2 - 4\bar{\mathbf{y}}^T\mathbf{y}\,\bar{\mathbf{y}}^T\tilde{K}\mathbf{y}$$

for all nonzero complex vectors $\mathbf{y}$. For ϵ such that $0 < \epsilon < \epsilon_0$, and for a fixed D, define $D(\epsilon)$ by

$$D(\epsilon) = 4(\bar{\mathbf{y}}^T\tilde{K}^{1/2}\mathbf{y})^2 + 4\epsilon\bar{\mathbf{y}}^T\tilde{K}^{1/2}\mathbf{y}\,\bar{\mathbf{y}}^T P\mathbf{y} + \epsilon^2(\bar{\mathbf{y}}^T P\mathbf{y})^2 - 4\bar{\mathbf{y}}^T\mathbf{y}\,\bar{\mathbf{y}}^T\tilde{K}\mathbf{y}, \tag{17}$$

and note that $D(\epsilon_0) > D(\epsilon)$. Differentiating (17) with respect to ϵ for a fixed P yields

$$\frac{d}{d\epsilon}(D(\epsilon)) = 4\bar{\mathbf{y}}^T\tilde{K}^{1/2}\mathbf{y}\,\bar{\mathbf{y}}^T P\mathbf{y} + 2\epsilon\,(\bar{\mathbf{y}}^T P\mathbf{y})^2. \tag{18}$$

Now note that if $\tilde{C} - 2\tilde{K}^{1/2}$ is positive-definite then $D'(\epsilon) > 0$ for all nonzero complex vectors $\mathbf{y}$, since $\tilde{K}^{1/2}$ and P are positive-definite and $\epsilon > 0$. In particular then $D'(\epsilon) > 0$ for all eigenvectors of (2).

Consider next, $D(\epsilon)$ defined on the set of all eigenvectors of (2) which we denote by $\mathbf{x}$. When $\epsilon = 0$, $\tilde{C} = 2\tilde{K}^{1/2}$ and from [6] the damped modal vectors are the eigenvectors of $\tilde{K}$. Thus

$$\begin{aligned} D(0) &= 4[(\bar{\mathbf{x}}^T\tilde{K}^{1/2}\mathbf{x})^2 - \bar{\mathbf{x}}^T\mathbf{x}\,\bar{\mathbf{x}}^T\tilde{K}\mathbf{x}] \\ &= 4[(\lambda^{1/2}\bar{\mathbf{x}}^T\mathbf{x})^2 - \bar{\mathbf{x}}^T\mathbf{x}(\lambda\bar{\mathbf{x}}^T\mathbf{x})] \\ &= 4[\lambda(\bar{\mathbf{x}}^T\mathbf{x})^2 - \lambda(\bar{\mathbf{x}}^T\mathbf{x})^2] = 0 \end{aligned}$$

where λ is the eigenvalue associated with the eigenvector $\mathbf{x}$. We have $D'(\epsilon) > 0$ for all $\epsilon > 0$, therefore $D(\epsilon) > 0$ for all $\epsilon > 0$ on the set of eigenvectors of (2). But, ϵ_0 may be arbitrarily large so that the discri-

minant of (15) is always positive. Hence, the eigenvalues of (2) must be real as must the eigenvectors for an overdamped system. The sign of the eigenvalues follows from the assumption that $\tilde{C}$ is positive-definite. We see that for the overdamped case none of the modes oscillate.

Mixed Damped Systems. For the special case that $\tilde{C}$ is diagonalized by the undamped modal matrix ($\tilde{C}\tilde{K} = \tilde{K}\tilde{C}$) we will show that equation (2) exhibits mixed damping if and only if there is at least one real root and at least one complex conjugate pair of roots of equation (3). One of the modes of the damped system will oscillate and at least one will not.

This result follows from the diagonalization of equation (2). Since $\tilde{C}\tilde{K} = \tilde{K}\tilde{C}$, there exists a transformation U such that $U^TU = I$ and $\Lambda_c = U^T\tilde{C}U$ and $\Lambda_k = U^T\tilde{K}U$ are both diagonal. If we let $\mathbf{y} = U\mathbf{x}$ in equation (2) we have the expression

$$\ddot{\mathbf{x}} + \Lambda_c\dot{\mathbf{x}} + \Lambda_k\mathbf{x} = 0. \tag{19}$$

Recalling equation (6), we have

$$\tilde{C} - C_c \sim \Lambda_c - 2\Lambda_k{}^{1/2}, \tag{20}$$

where the tilde denotes similarity equivalence [5]. The right-hand side of expression (20) is diagonal. If it is indefinite we must have at least one value of i such that

$$(\Lambda_c - 2\Lambda_k{}^{1/2})_{ii} \geq 0, \tag{21}$$

and at least one other value of i such that

$$(\Lambda_c - 2\Lambda_k{}^{1/2})_{ii} < 0. \tag{22}$$

The ith pair of eigenvalues of (19) are found from

$$\lambda_i = -\frac{(\Lambda_c)_{ii}}{2} \pm \frac{[(\Lambda_c)_{ii}{}^2 - 4(\Lambda_k)_{ii}]^{1/2}}{2}$$

Clearly if (21) holds λ is real and if (22) holds λ is complex with negative real part. Hence, at least one mode of (3) will oscillate and at least one mode will not. The proof in the opposite direction is straightforward.

Discussion of Results

Except for the mixed damping case the foregoing relationships between C_c and $\tilde{C}$ are, in general, only sufficient to determine the nature of the resulting solutions. However, all of the relationships just stated become both necessary and sufficient for the special case that (2) is diagonalized by the undamped modal matrix.

Duffin's definition states that an overdamped system is one such that

$$(\mathbf{x}^T\tilde{C}\mathbf{x})^2 > 4\mathbf{x}^T\mathbf{x}\mathbf{x}^T\tilde{K}\mathbf{x}$$

for all real $\mathbf{x}$. This inequality holds on the set of eigenvectors of (2) if $\tilde{C} - C_c$ is positive-definite. In this sense Duffin's definition, though harder to check, is equivalent to the definition of overdamping stated here. Lancaster [4] proves that, if a system is overdamped, the λ-matrix defining the system must be of simple structure, meaning that the eigenvectors form a complete set, and hence, span an n-dimensional vector space (theorems 4.4 and 7.3 of [4]). By following the proof of Lancaster it can easily be shown that if $\tilde{C} - C_c$ is positive-definite then the eigenvectors of (2) are complete. Hence, in terms of the definition offered here, a system which is overdamped has a complete set of eigenvectors associated with it regardless of the multiplicity of its eigenvalues.

Nicholson's definition of underdamping [2] states that the system is underdamped if all the modes of (2) are underdamped. He then states that a sufficient condition for (2) to be underdamped is for

$$c_1 \leq 2k_m{}^{1/2}$$

where c_1 is the largest eigenvalue of the matrix $\tilde{C}$ and k_m is the smallest eigenvalue of the matrix $\tilde{K}$. This requires substantial calculation to check because it involves finding the eigenvalues of both $\tilde{C}$ and $\tilde{K}$. Müller [3] improves this result and extends it by showing

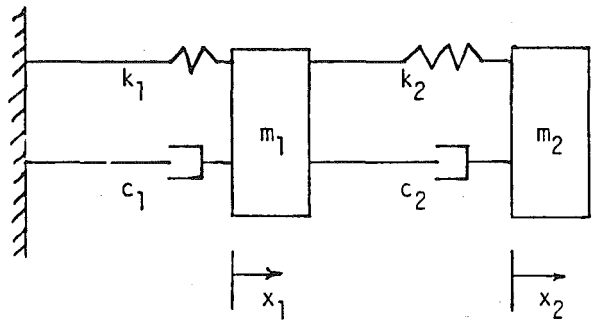


Fig. 1

that a sufficient condition for (2) to be underdamped is for $4\tilde{K} - \tilde{C}^2$ to be positive-definite.

It can easily be shown that Müller's condition is equivalent to the one stated here if $\tilde{C}$ and $\tilde{K}$ commute. This is seen by assuming that $2\tilde{K}^{1/2} - \tilde{C}$ is positive-definite. Since $2\tilde{K}^{1/2}$ and $\tilde{C}$ are both positive-definite $2\tilde{K}^{1/2} + \tilde{C}$ is also positive-definite. Using two well-known results (a) that the product of two positive-definite matrices is positive-definite if and only if they commute; (b) if $\tilde{C}$ and $\tilde{K}$ commute, then so do $\tilde{C}$ and $\tilde{K}^{1/2}$) we have

$$\begin{aligned}(2\tilde{K}^{1/2} - \tilde{C})\,(2\tilde{K}^{1/2} + \tilde{C}) &= 4\tilde{K} - 2\tilde{C}\tilde{K}^{1/2} + 2\tilde{K}^{1/2}\tilde{C} - \tilde{C}^2\\ &= (2\tilde{K}^{1/2} + \tilde{C})\,(2\tilde{K}^{1/2} - \tilde{C})\\ &= 4\tilde{K} - \tilde{C}^2;\end{aligned}$$

which is positive-definite. Recalling that if A and B are positive-definite matrices, then $A < B$ implies $A^{1/2} < B^{1/2}$ [8], we see that Müller's result and our result are equivalent if $\tilde{C}\tilde{K} = \tilde{K}\tilde{C}$.

Examples

Several two-dimensional examples serve to illustrate the validity of the foregoing results.

The definiteness of $\tilde{C} - C_c$ can easily be checked by examining the determinant of each of its minors. The square root of $\tilde{K}$ can be found by finding the eigenvectors and eigenvalues of $\tilde{K}$ and using (6), Newton's method or a generalization of Newton's method given in [9]. For an easier first check one can look at the definiteness of $\tilde{C}^2 - 4\tilde{K}$, since $\tilde{C}^2 - 4\tilde{K} > 0$ implies $\tilde{C} - C_c > 0$, $4\tilde{K} - \tilde{C}^2 > 0$ implies $C_c - \tilde{C} > 0$ and $\tilde{C}^2 = 4\tilde{K}$ implies $\tilde{C} = C_c$. However, if $\tilde{C}^2 - 4\tilde{K}$ is indefinite $C_c - \tilde{C}$ should still be checked since it yields stronger results than those based on the definiteness of $\tilde{C}^2 - 4\tilde{K}$.

Critical Damping. Consider the system described by equation (2) with

$$\tilde{K} = \begin{bmatrix} 2.5 & 1.25 \\ 1.25 & 1.25 \end{bmatrix}$$

which is positive-definite. Then

$$C_c = 2\tilde{K}^{1/2} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

is also positive-definite. The associated eigenvalue problem is then

$$\left\{\begin{bmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{bmatrix} + \begin{bmatrix} 3\lambda & \\ \lambda & 2\lambda \end{bmatrix} + \begin{bmatrix} 2.5 & 1.25 \\ 1.25 & 1.25 \end{bmatrix}\right\} x = 0$$

The characteristic equation is

$$\lambda^4 + 5\lambda^3 + 8.75\lambda^2 + 6.25\lambda + 1.5625 = 0,$$

which has the following roots:

$$\lambda_{1,2} = -0.690983005,$$
$$\lambda_{3,4} = -1.809016994.$$

Thus there are at most $n = 2$ negative real roots as was predicted.

For the last two examples consider the physical example illustrated

in Fig. 1. Here, x_1 and x_2 are the displacements from equilibrium of masses m_1 and m_2 which are taken to be unity for simplicity.

The equations of motion for this system are

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\ddot{\mathbf{x}} + \begin{bmatrix}c_1 + c_2 & -c_2\\-c_2 & c_2\end{bmatrix}\mathbf{x} + \begin{bmatrix}k_1 + k_2 & -k_2\\-k_2 & k_2\end{bmatrix}\mathbf{x} = 0$$

The most interesting case for this problem is one in which the C matrix is not diagonalized by the undamped modal matrix transformation. Hence we will choose parameters such that $\tilde{C}\tilde{K} \neq \tilde{K}\tilde{C}$.

Overdamping. Suppose $c_1 = 5$, $c_2 = 4$, $k_1 = 2$, and $k_2 = 1$ then

$$\tilde{C} = \begin{bmatrix}9 & -4\\-4 & 4\end{bmatrix} \quad \text{and} \quad \tilde{K} = \begin{bmatrix}3 & -1\\-1 & 1\end{bmatrix}$$

so that

$$\tilde{C}^2 - 4\tilde{K} = \begin{bmatrix}85 & -48\\-48 & 28\end{bmatrix}$$

which is positive-definite so that $\tilde{C} - 2\tilde{K}^{1/2}$ is positive-definite and the system is overdamped. The characteristic equation is

$$\lambda^4 + 13\lambda^2 + 24\lambda^2 + 13\lambda + 2 = 0$$

with roots

$$\lambda_1 = -0.266$$

$$\lambda_2 = -0.532$$

$$\lambda_3 = -1.294$$

$$\lambda_4 = -10.907$$

which are all negative real numbers in agreement with our results.

Underdamping. Next, let $c_1 = 2$, $c_2 = 1$, $k_1 = 4$, and $k_2 = 1$ so that

$$\tilde{C} = \begin{bmatrix}3 & -1\\-1 & 1\end{bmatrix} \quad \text{and} \quad \tilde{K} = \begin{bmatrix}5 & -1\\-1 & 1\end{bmatrix}.$$

Then

$$4\tilde{K} - \tilde{C}^2 = \begin{bmatrix}10 & 0\\0 & 6\end{bmatrix}$$

which is positive-definite, so that $2\tilde{K}^{1/2} - \tilde{C}$ is positive-definite and the system is underdamped. The characteristic equation is

$$\lambda^4 + 4\lambda^3 + 8\lambda^2 + 6\lambda + 4 = 0$$

which has roots

$$\lambda_{1,2} = -0.337 \pm 0.8326j$$

$$\lambda_{3,4} = -1.66 \pm 1.481j$$

where $j = \sqrt{-1}$. All the roots are complex conjugate pairs with negative real parts as was predicted.

References

1 Duffin, R. J., "A Minimax Theory for Overdamped Networks," *Journal of Rational Mechanics and Analysis,* Vol. 4, 1955, pp. 221–233.

2 Nicholson, D. W., "Eigenvalue Bounds for Damped Linear Systems," *Mechanics Research Communications,* Vol. 5, No. 3, 1978, pp. 147–152.

3 Müller, P. C., "Oscillatory Damped Linear Systems," *Mechanics Research Communications,* Vol. 6, No. 2, 1979, pp. 81–85.

4 Lancaster, P., *Lambda-Matrices and Vibrating Systems,* Pergamon Press, 1966.

5 Bellman, R., *Introduction to Matrix Analysis,* 2nd ed., McGraw-Hill, New York, 1970, p. 95.

6 Caughey, T. K., and O'Kelly, M. E. J., "Classical Normal Modes in Damped Linear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, 1965, pp. 583–588.

7 Walker, J. A., and Schmitendorf, W. E., "A Simple Test for Asymptotic Stability in Partially Dissipative Symmetric Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, 1973, pp. 1120–1121.

8 Bellman, R., "Some Inequalities for the Square Root of a Positive-Definite Matrix," *Linear Algebra and Its Applications,* Vol. 1, 1968, pp. 321–324.

9 Hoskins, W. D., and Walton, D. J., "A Faster More Stable Method for Computing the pth Roots of Positive-Definite Matrices," *Linear Algebra and Its Applications,* Vol. 26, 1979, pp. 139–163.

C. S. Hsu

Professor of Applied Mechanics,
Department of Mechanical Engineering,
University of California,
Berkeley, Calif. 94720
Fellow ASME

A Theory of Cell-to-Cell Mapping Dynamical Systems

The method of point-to-point mappings has been receiving increasing attention in recent years. In this paper we discuss instead dynamical systems governed by cell-to-cell mappings. The justifications of considering such mappings come from the unavoidable accuracy limitations of both physical measurements and numerical evaluation. Because of these limitations one is not really able to treat a state variable as a continuum of points but rather only as a collection of very small intervals. The introduction of the idea of cell-to-cell mappings has led to an algorithm which is found to be potentially a very powerful tool for global analysis of dynamical systems. In this paper an introductory theory of cell-to-cell mappings is offered. The theory provides a basis for the algorithm presented in [14]. In the first half of the paper we discuss the analysis of cell-to-cell mappings in their own right. In the second half the cell-to-cell mappings which are obtained from point-to-point mappings by discretization are examined in order to see what properties of the point mapping systems are preserved in the discretization process.

1 Introduction

Consider a dynamical system governed by

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t, \mathbf{x}(t)) \tag{1}$$

where $\mathbf{x}$ is a real-valued N-vector and $\mathbf{F}$ is a real-valued vector function. If the system is periodic so that $\mathbf{F}$ is explicitly periodic in t, then one may in principle integrate the equation over one period to relate the state of the system at the end of one period to the state at the end of the next period. Viewed in this manner, the governing equation for the system takes on the form

$$\mathbf{x}(n+1) = \mathbf{G}(\mathbf{x}(n)) \tag{2}$$

A point $\mathbf{x}(n)$ in the state space is mapped by $\mathbf{G}$ after one period into a point $\mathbf{x}(n+1)$. Such a point-to-point mapping dynamical system is called a point map or a Poincaré map in the mathematical literature. In recent years this method of point mapping has been receiving increasing attention as an attractive tool for treating nonlinear dynamical systems. The general method dates back to Poincaré [1] and Birkhoff [2]. In the past 20 years or so it has received a great deal of mathematical development; see, for instance, [3–6]. Recently this theory has been applied in [7–11] for studying certain strongly nonlinear mechanical systems under periodic parametric excitations. By this approach various interesting bifurcation phenomena and other nonlinear features can be studied in a very effective manner. It is also strongly believed that since the point-to-point mapping method is extremely well suited for computer adaptation, there could be a much greater development of the methods in the coming years, particularly for the purpose of studying nonlinear systems.

Point mapping dynamical systems have some peculiar features of their own. One is the possible existence of a cascade of bifurcations from periodic motions into periodic motions of ever higher periods leading to a seemingly chaotic motion [6]. Then there is the matter of the existence of homoclinic and heteroclinic points which make the numerical determination of the global behavior extremely difficult [12]. Another remarkable feature is that a motion which seems to be chaotic at one scale would exhibit a definite structure when a finer scale is used [13]. Features like these are most interesting and challenging but also vexing.

Here, one is tempted to take a somewhat more simple-minded but perhaps also more realistic point of view. This is concerned with the question of how fine a scale one is allowed to use in specifying a state variable. From the point of view of physical measurement, there is a limit of accuracy of measurement, say h. Two values of a state variable differing by less than h cannot be differentiated and for practical purposes they have to be treated as same. Also, from the point of view of computation, one is limited by the numerical precision to incur roundoffs. From both points of view one cannot really hope to deal with a true *continuum* of a state variable, but rather is forced to deal with only a large but discrete set of values for each of the state variables. This leads to the idea of considering a state variable not as a continuum of points but as a collection of cells. This consideration provides basically the motivation for the study of the present paper.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, May, 1980; final revision, June, 1980. Paper No. 80-WA/APM-26.

2 Cell-to-Cell Mapping Dynamical Systems

Let the coordinate axis of a state variable x_i ($i = 1, 2, \ldots, N$) be divided into a large number of intervals with an interval size h_i. The interval Z_i of the x_i-axis is defined to be one which contains all x_i satisfying

$$(Z_i - \tfrac{1}{2})h_i \leqq x_i < (Z_i + \tfrac{1}{2})h_i \tag{3}$$

Here, by definition Z_i is an integer. A N-tuple Z_i, $i = 1, 2, \ldots, N$, is then called a *cell vector* of the state space and is denoted by $\mathbf{Z}$.[1] A point $\mathbf{x}(x_i, i = 1, 2, \ldots, N)$ belongs to a cell $\mathbf{Z}(Z_i = 1, 2, \ldots, N)$ if x_i belongs to Z_i for all i. Each cell is now considered as an entity and the state space is regarded as a collection of cells. Viewing the system in this manner and using an appropriate rule to identify $\mathbf{x}(n)$ and $\mathbf{x}(n + 1)$ with corresponding cells $\mathbf{Z}(n)$ and $\mathbf{Z}(n + 1)$, one can associate to the point-to-point mapping (2) a cell-to-cell mapping $\mathbf{C}$

$$\mathbf{Z}(n + 1) = \mathbf{C}(\mathbf{Z}(n)), \quad Z_i(n + 1) = C_i(\mathbf{Z}(n)) \tag{4}$$

where it is implied that $\mathbf{C}$ maps a set of integers to a set of integers. For convenience, we refer to (4) as a *cell-to-cell mapping dynamical system* and refer to $\mathbf{C}$ as a *cell-to-cell mapping*, or simply a *cell mapping* or a *cell map*.

Such a reformulation from a point mapping to a cell mapping, while interesting by itself, will be merely an academic exercise if it does not lead to any practical advantages. In a companion paper [14] we shall demonstrate the utility of cell-to-cell mappings by offering a computational algorithm which can be used in a very effective way to study the global behavior of nonlinear systems. That algorithm is however based upon the properties of cell-to-cell mappings. Therefore, we shall first carry out a theoretical study of cell-to-cell mappings in this paper in order to provide a basis for [14] and other developments.

We first describe in Section 3 periodic motions and periodic cells of cell mappings, bifurcation phenomena, and domains of attraction of periodic cells. The development given in this section is valid for systems of any dimension. In Section 4 we study one-dimensional systems in order to gain a better appreciation of the properties of cell mapping dynamical systems. A similar study of two-dimensional systems is presented in Section 5. Cell mappings are studied in their own right in Sections 4–5. In Sections 6–8 we study the relation between cell mappings and point mappings assuming that the cell mappings are in fact derived from the point mappings by some appropriate process of discretization. Here our main purpose is to see what properties are preserved, what properties are lost, and what new properties are introduced in going from a point mapping to a cell mapping.

3 Periodic Motions, Bifurcation, and Domains of Attraction

In this section we discuss some basic concepts of cell mapping systems. First, let us identify cells with points of a Euclidean space with integer-valued Cartesian coordinates. Let $\mathbf{e}_i$ be a unit vector in the direction of the Z_i-axis. It can also be regarded as an N-tuple of integers {0,0, . .0,1,0, . .0} with unit 1 at the ith position. A cell vector $\mathbf{Z}$ can can then be written as

$$\mathbf{Z} = \sum_{i=i}^{N} Z_i \mathbf{e}_i \tag{5}$$

A cell $\mathbf{Z}'$ is said to be a *contiguous cell* to $\mathbf{Z}$ in the Z_i-direction if

$$\mathbf{Z}' = \mathbf{Z} \pm \mathbf{e}_i. \tag{6}$$

The mapping of contiguous cells may be characterized by the local increments of the mapping $\mathbf{C}$. The *forward increment vectors* of $\mathbf{C}$ at $\mathbf{Z}$ are defined as

$$\Delta_j \mathbf{C}(\mathbf{Z}) = \mathbf{C}(\mathbf{Z} + \mathbf{e}_j) - \mathbf{C}(\mathbf{Z}). \tag{7}$$

The *backward increment vectors* of $\mathbf{C}$ at $\mathbf{Z}$ are defined as

$$\nabla_j \mathbf{C}(\mathbf{Z}) = \mathbf{C}(\mathbf{Z}) - \mathbf{C}(\mathbf{Z} - \mathbf{e}_j). \tag{8}$$

Consistency of course requires

$$\Delta_j \mathbf{C}(\mathbf{Z}) = \nabla_j \mathbf{C}(\mathbf{Z} + \mathbf{e}_j). \tag{9}$$

$$\Delta_j \mathbf{C}(\mathbf{Z}) + \Delta_k \mathbf{C}(\mathbf{Z} + \mathbf{e}_j) = \Delta_k \mathbf{C}(\mathbf{Z}) + \Delta_j \mathbf{C}(\mathbf{Z} + \mathbf{e}_k). \tag{10}$$

The components of the increment vectors at $\mathbf{Z}$ will be written as $\Delta_{ij}(\mathbf{Z}, \mathbf{C})$ and $\nabla_{ij}(\mathbf{Z}, \mathbf{C})$ respectively. Hence,

$$\Delta_{ij}(\mathbf{Z}, \mathbf{C}) = C_i(\mathbf{Z} + \mathbf{e}_j) - C_i(\mathbf{Z}), \tag{11}$$

$$\nabla_{ij}(\mathbf{Z}, \mathbf{C}) = C_i(\mathbf{Z}) - C_i(\mathbf{Z} - \mathbf{e}_j). \tag{12}$$

Equilibrium Cell. A cell $\mathbf{Z}^*$ which satisfies

$$\mathbf{Z}^* = \mathbf{C}(\mathbf{Z}^*) \tag{13}$$

is said to be an *equilibrium cell* of the system. An equilibrium cell is said to be *isolated* if none of its contiguous cells is an equilibrium cell. One notes that $\mathbf{Z}^* + \mathbf{e}_J$ is an equilibrium cell if and only if

$$\mathbf{C}(\mathbf{Z}^* + \mathbf{e}_J) = \mathbf{Z}^* + \mathbf{e}_J = \mathbf{C}(\mathbf{Z}^*) + \mathbf{e}_J \tag{14}$$

This implies

$$\Delta_J \mathbf{C}(\mathbf{Z}^*) = \mathbf{e}_J \quad \text{or} \quad \Delta_{iJ}(\mathbf{Z}^*, \mathbf{C}) = \delta_{iJ} \tag{15}$$

Similarly, $Z^* - \mathbf{e}_J$ is an equilibrium cell if and only if

$$\nabla_J \mathbf{C}(\mathbf{Z}^*) = \mathbf{e}_J \quad \text{or} \quad \nabla_{iJ}(\mathbf{Z}^*, \mathbf{C}) = \delta_{iJ} \tag{16}$$

Interpretation of (15) and (16) is simple. If and only if the increment vector $\Delta_J \mathbf{C}(\mathbf{Z}^*)$ is equal to the unit vector $\mathbf{e}_J$, then Z^* has a contiguous equilibrium cell at $\mathbf{Z}^* + \mathbf{e}_J$. It has a contiguous equilibrium cell at $\mathbf{Z}^* - \mathbf{e}_J$ if and only if the backward increment vector $\nabla_J \mathbf{C}(\mathbf{Z}^*)$ is equal to $\mathbf{e}_J$. The conditions for $\mathbf{Z}^*$ to be isolated are therefore

$$\Delta_J \mathbf{C}(\mathbf{Z}^*) \neq \mathbf{e}_J \text{ and } \nabla_J \mathbf{C}(\mathbf{Z}^*) \neq \mathbf{e}_J \text{ for all } J = 1, 2, \ldots, N. \tag{17}$$

If $\Delta_{ij}(Z^*, C)$ happens to be a unit matrix, then all the forward contiguous cells of $\mathbf{Z}^*$ are equilibrium cells. If $\nabla_{ij}(Z^*, C)$ happens to be a unit matrix, then all the backward contiguous cells of $\mathbf{Z}$ are equilibrium cells.

Core of Equilibrium Cells. Often one finds equilibrium cells which are contiguous. The largest collection of such contiguous equilibrium cells[2] is said to form a *core of equilibrium cells*. The *size of the core* is defined to be equal to the number of cells in the core.

Periodic Motions. Let $\mathbf{C}^m$ denote the cell mapping $\mathbf{C}$ applied m times with $\mathbf{C}^0$ understood to be the identity mapping. A sequence of K distinct cells $\mathbf{Z}^*(j)$, $j = 1, 2, \ldots, K$, which satisfy

$$\mathbf{Z}^*(m + 1) = \mathbf{C}^m(\mathbf{Z}^*(1)), \quad m = 1, 2, \ldots, K - 1$$

$$\mathbf{Z}^*(1) = \mathbf{C}^K(\mathbf{Z}^*(1)) \tag{18}$$

is said to constitute a *periodic motion of period K* of the cell mapping $\mathbf{C}$. For ease of reference we call such a motion a *P-K motion* and each of its elements $\mathbf{Z}^*(j)$ a *periodic cell of period K* or simply a *P-K cell*. According to this definition an equilibrium cell is of course a $P - 1$ cell.

The increment vectors of a cell mapping $\mathbf{C}$ may be generalized for the mapping $\mathbf{C}^m$. We define forward and backward increment vectors of $\mathbf{C}^m$ at $\mathbf{Z}$ as

$$\Delta_j \mathbf{C}^m(\mathbf{Z}) = \mathbf{C}^m(\mathbf{Z} + \mathbf{e}_j) - \mathbf{C}^m(\mathbf{Z}) \tag{19}$$

$$\nabla_j \mathbf{C}^m(\mathbf{Z}) = \mathbf{C}^m(\mathbf{Z}) - \mathbf{C}^m(\mathbf{Z} - \mathbf{e}_j) \tag{20}$$

Their corresponding components are

$$\Delta_{ij}(\mathbf{Z}, \mathbf{C}^m) = (\mathbf{C}^m(\mathbf{Z} + \mathbf{e}_j))_i - (\mathbf{C}^m(\mathbf{Z}))_i \tag{21}$$

[1] Since Z is often used to denote the set of integers, we shall use whenever possible $\mathbf{Z}$ and Z_i to denote the cell vector and its components.

[2] Sometimes, one finds equilibrium cells which are near to each other but not exactly contiguous. In such cases we refer to them, in a general term, as a cluster of equilibrium cells.

$$\nabla_{ij}(\mathbf{Z}, \mathbf{C}^m) = (\mathbf{C}^m(\mathbf{Z}))_i - (\mathbf{C}^m(\mathbf{Z} - \mathbf{e}_j))_i \tag{22}$$

Now we can study the condition for two contiguous cells to be periodic cells of equal or different periods. Let $\mathbf{Z}^*$ be a $P - K_1$ cell and $(\mathbf{Z}^* + \mathbf{e}_J)$ be a $P - K_2$ cell. Let L be the least common multiple of K_1 and K_2.

$$L = k_1K_1 = k_2K_2 \tag{23}$$

We can then evaluate $\Delta_J\mathbf{C}^L(\mathbf{Z}^*)$

$$\Delta_J\mathbf{C}^L(\mathbf{Z}^*) = \mathbf{C}^{k_2K_2}(\mathbf{Z}^* + \mathbf{e}_J) - \mathbf{C}^{k_1K_1}(\mathbf{Z}^*) = \mathbf{Z}^* + \mathbf{e}_J - \mathbf{Z}^*$$

leading to

$$\Delta_J\mathbf{C}^L(\mathbf{Z}^*) = \mathbf{e}_J \tag{24}$$

Thus (24) is the necessary condition for $\mathbf{Z}^*$ to be a $P - K_1$ cell and $(\mathbf{Z}^* + \mathbf{e}_J)$ to be a $P - K_2$ cell. In the sufficiency direction one can show that if (24) is true and $\mathbf{Z}^*$ is a $P - K_1$ cell, then $\mathbf{Z}^* + \mathbf{e}_J$ is a $P - K_2$ cell where K_2 is a positive integer factor of L. The condition (24) may also be written as

$$\nabla_J\mathbf{C}^L(\mathbf{Z}^* + \mathbf{e}_J) = \mathbf{e}_J \tag{25}$$

We can also use the increment vectors to indicate the separation between two periodic motions. Let $\mathbf{Z}^*(j)$, $j = 1, 2, \ldots, K_1$ be the $P - K_1$ cells of a $P - K_1$ motion. Let $\mathbf{Z}^{**}(j)$, $j = 1, 2, \ldots, K_2$ be the $P - K_2$ cells of a $P - K_2$ motion. $\mathbf{Z}^*(j)$ satisfies (18) with $K = K_1$ and $\mathbf{Z}^{**}(j)$ satisfies (18) with $K = K_2$. We have of course

$$\mathbf{C}^{K_1+m}(\mathbf{Z}^*(1)) = \mathbf{C}^m(\mathbf{Z}^*(1)) \tag{26a}$$

$$\mathbf{C}^{K_2+m}(\mathbf{Z}^{**}(1)) = \mathbf{C}^m(\mathbf{Z}^{**}(1)) \tag{26b}$$

Again let L be the least common multiple of K_1 and K_2. Assuming now that $Z^*(1)$ and $Z^{**}(1)$ are contiguous so that

$$\mathbf{Z}^{**}(1) - \mathbf{Z}^*(1) = \mathbf{e}_J \tag{27a}$$

Then one can show that

$$\mathbf{Z}^{**}(j) - \mathbf{Z}^*(j) = \Delta_J\mathbf{C}^{j-1}(\mathbf{Z}^*(1)), \quad j = 1, 2, \ldots L \tag{27b}$$

$$\mathbf{Z}^{**}(1 + L) - \mathbf{Z}^*(1 + L) = \Delta_J\mathbf{C}^L(\mathbf{Z}^*(1)) = \mathbf{e}_J \tag{27c}$$

The concept of an isolated equilibrium can also be generalized for $P - K$ cells. A $P - K$ cell $\mathbf{Z}^*$ is said to be an *isolated $P - K$ cell* if none of its contiguous cells is a $P - K$ cell. By conditions (24) and (25) a $P - K$ cell $\mathbf{Z}^*$ is isolated if and only if

$$\Delta_J\mathbf{C}^K(\mathbf{Z}^*) \neq \mathbf{e}_J, \quad \nabla_J\mathbf{C}^K(\mathbf{Z}^*) \neq \mathbf{e}_J \quad \text{for all} \quad J = 1, 2, \ldots N \tag{28}$$

A $P - K$ cell $\mathbf{Z}^*$ is said to be an *isolated periodic cell* if none of its contiguous cells is a periodic cell, of any period. It requires that

$$\Delta_J\mathbf{C}^{kK}(\mathbf{Z}^*) \neq \mathbf{e}_J \quad \text{and} \quad \nabla_J\mathbf{C}^{kK}(\mathbf{Z}^*) \neq \mathbf{e}_J$$
$$\text{for all} \quad J = 1, 2, \ldots, N \quad \text{and all} \quad k = 1, 2, \ldots \tag{29}$$

The concept of a core of equilibrium cells can also be generalized. A collection of contiguous $P - K$ cells with all the contiguous $P - K$ cells included is said to form a *core of $P - K$ cells*. The size of the core is the number of the cells in the core. Similarly, a collection of contiguous periodic cells of all periods with all the contiguous periodic cells included is said to form a *core of periodic cells*. Again, the size of the core is defined to be the number of the cells in the core.

We can also introduce a notion which describes one-step mapping local behavior of a $P - K$ cell. Let $\mathbf{Z}^*$ be such a $P - K$ cell. Let the norm of a cell vector $\mathbf{Z}$ be defined as

$$\|\mathbf{Z}\| = \sum_i |Z_i|. \tag{30}$$

If

$$\|\Delta_J\mathbf{C}^K(\mathbf{Z}^*)\| > 1$$

then $\mathbf{Z}^*$ is said to be forward-repulsive in the J-direction. Similarly,

$\|\Delta_J\mathbf{C}^K(\mathbf{Z}^*)\| = 1$ forward-neutral in the J-direction
$\|\Delta_J\mathbf{C}^K(\mathbf{Z}^*)\| = 0$ forward-attracting in the J-direction
$\|\nabla_J\mathbf{C}^K(\mathbf{Z}^*)\| > 1$ backward-repulsive in the J-direction
$\|\nabla_J\mathbf{C}^K(\mathbf{Z}^*)\| = 1$ backward-neutral in the J-direction
$\|\nabla_J\mathbf{C}^K(\mathbf{Z}^*)\| = 0$ backward-attracting in the J=direction

Bifurcation. The discussion of contiguous periodic cells is connected with bifurcation phenomena. Consider a cell mapping system which depends upon a parameter α

$$\mathbf{Z}(n + 1) = \mathbf{C}(\mathbf{Z}(n), \alpha) \tag{31}$$

where α is a real-valued parameter. $\mathbf{C}$ can take on only integer values. Thus, as α varies, each component of $\mathbf{C}$ either remains unchanged or suffers a jump of integer number of units. We say $\mathbf{C}$ is a *contiguous function* of α if as α varies the jumps C_i can take are only $+1$ or -1. In the following discussion we assume $\mathbf{C}$ to be a contiguous function of α.

Let $\mathbf{Z}^*(\alpha)$ be a $P - 1$ cell which may depend upon α. Assume that for $\alpha < \alpha_1$

$$\Delta_j\mathbf{C}(\mathbf{Z}^*(\alpha)) \neq \mathbf{e}_j, \quad \nabla_j\mathbf{C}(\mathbf{Z}^*(\alpha)) \neq \mathbf{e}_j \quad \text{for all} \quad j = 1, 2, \ldots, N. \tag{32}$$

Then according to (17), $\mathbf{Z}^*(\alpha)$ is an isolated $P - 1$ cell. If as α increases across α_1, one of the $\Delta_j\mathbf{C}(\mathbf{Z}^*(\alpha))$ and $\nabla_j\mathbf{C}(\mathbf{Z}^*(\alpha))$, say $\Delta_J\mathbf{C}(\mathbf{Z}^*(\alpha))$ (or $\nabla_J\mathbf{C}(\mathbf{Z}^*(\alpha))$), becomes $\mathbf{e}_J$, then the contiguous cell at $\mathbf{Z}^* + \mathbf{e}_J$ (or $\mathbf{Z}^* - \mathbf{e}_J$) becomes now a new $P - 1$ cell. We say that at $\alpha = \alpha_1$ a bifurcation from a $P - 1$ cell to a $P - 1$ cell has taken place at $\mathbf{Z}^*$ in the J-direction. At a still higher value of α, say α_2, another bifurcation from $P - 1$ to $P - 1$ might take place at $\mathbf{Z}^*(\alpha)$ in another direction. If, as α increases further, one of the increment vectors, say $\Delta_J\mathbf{C}(\mathbf{Z}^*(\alpha))$ (or $\nabla_J\mathbf{C}(\mathbf{Z}^*(\alpha))$), changes from $\mathbf{e}_J$ to a vector different from $\mathbf{e}_J$, then we have the phenomenon of disappearance of a contiguous $P - 1$ cell.

Consider next, the more general case. Let $\mathbf{Z}^*(\alpha)$ be a $P - K_1$ cell and L be a multiple of K_1. If as α increases across α_1, $\Delta_J\mathbf{C}^L(\mathbf{Z}^*(\alpha))$ (or $\nabla_J\mathbf{C}^L(\mathbf{Z}^*(\alpha))$) changes from a vector different from $\mathbf{e}_J$ to $\mathbf{e}_J$, then a new $P - K_2$ cell at $\mathbf{Z}^*(\alpha) + \mathbf{e}_J$ (or $\mathbf{Z}^*(\alpha) - \mathbf{e}_J$) comes into being at $\alpha = \alpha_1$ where K_2 is a positive integer factor of L. Here, we say that a bifurcation from $P - K_1$ to $P - K_2$ has taken place at $\alpha = \alpha_1$.

It might be appropriate to remark here that the forward and backward increments play a similar role in stability and bifurcation analysis as the Jacobian $\mathbf{DG}(\mathbf{x})$ does for a point mapping system, [10].

Domains of Attraction. Finally, we define the domains of attraction for cell mapping systems. A cell $\mathbf{Z}$ is said to be "*r-steps removed from a $P - K$ motion*" if r is the minimum positive integer such that $\mathbf{C}^r(\mathbf{Z}) = \mathbf{Z}^*(j)$ where $\mathbf{Z}^*(j)$ is one of the $P - K$ cells of that $P - K$ motion. In other words, $\mathbf{Z}$ is mapped in r steps into one of the $P - K$ cells of the $P - K$ motion and any further mapping will lock the evolution with this $P - K$ motion.

The set of all cells which are r steps or less removed from a $P - K$ solution is called the "*r-step domain of attraction*" for that $P - K$ motion. The *total domain of attraction* (or simply the *domain of attraction*) of a $P - K$ motion is its r-step domain of attraction with $r \to \infty$.

4 One-Dimensional Cell-Mapping Dynamical Systems

We shall discuss in this section a sample one-dimensional cell mapping system in order to illustrate the concepts introduced in Section 3. The mapping C is shown graphically in Fig. 1(a). Here we note that $Z^* = -3, -2, 6$ are $P - 1$ cells. $Z^* = 6$ is an isolated $P - 1$ cell while $Z^* = -3$ and $Z^* = -2$ are not. As by (15) $\Delta C(6) \neq 1$, $\nabla C(6) \neq 1$, $\Delta C(-3) = 1$, and $\nabla C(-2) = 1$. Here we drop the subscript J because we are dealing with a one-dimensional case. The two cells at $Z^* = -3$ and $Z^* = -2$ form a core of 2 $P - 1$ cells. There are $P - 2$ cells at $Z^* = 4$, 5, 7, and 9. The pair $Z^* = 4$ and $Z^* = 9$ form one $P - 2$ motion while $Z^* = 5$ and $Z^* = 7$ form the other $P - 2$ motion. There are $P -$

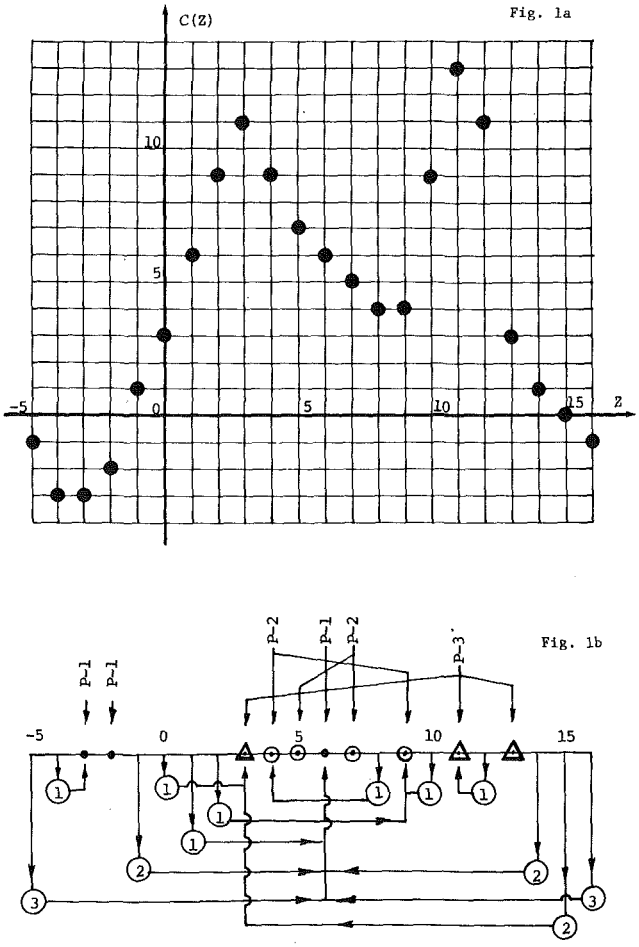


Fig. 1 An illustrative example of one-dimensional cell-to-cell mappings

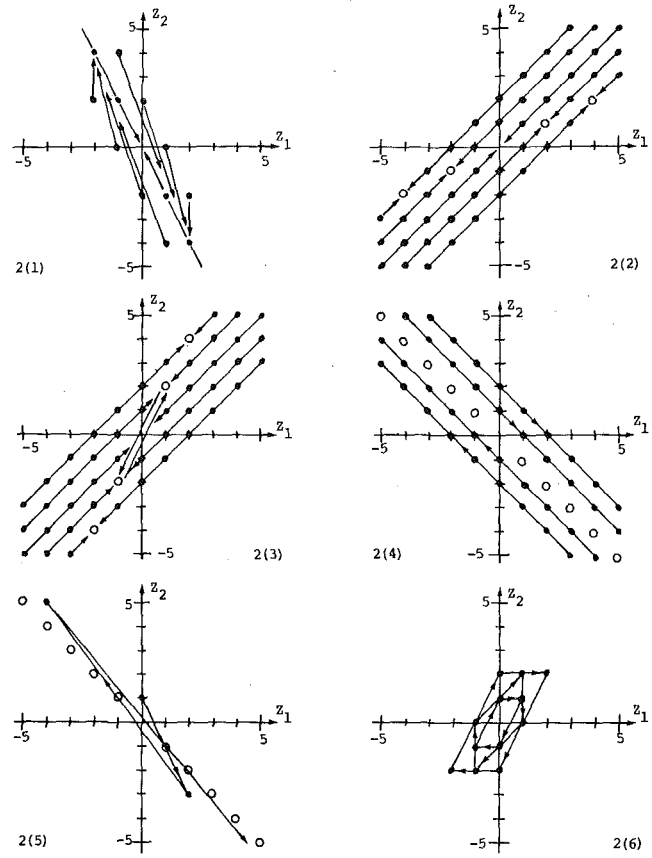


Fig. 2(1)–2(6) Various patterns of trajectories around the origin for two-dimensional linear cell-to-cell mappings

3 cells at $Z^* = 3$, $Z^* = 11$ and $Z^* = 13$ to make up a $P - 3$ motion. $Z^* = 4$ and $Z^* = 5$ form a core of $2\,P - 2$ cells. $Z^* = 3, 4, 5, 6$, and 7 form a core of 5 periodic cells.

Within the domain $-5 \le Z \le 16$ as indicated in Fig. 1, one-step domain of attraction for $Z^* = -3$ consists of only one cell at $Z^* = -4$. The one-step domain of attraction for the $P - 1$ cell at $Z^* = -2$ is empty. The one-step domain of attraction for the $P - 1$ cell at $Z^* = 6$ consists of one cell $Z = 1$. The two-step domain consists of $Z = -1, 1, 14$. The 3-step domain $-5, -1, 1, 14, 16$. The $P - 2$ motion ($Z^* = 4$ and 9) has $Z = 2, 8, 10$ as its one-step domain of attraction. The $P - 2$ motion ($Z^* = 5$ and 7) has no domains of attraction of any steps. The $P - 3$ motion ($Z = 3, 11$ and 13) has $Z = 0$ and 12 as its one-step domain of attraction, and $Z = 15$ added for its 2-step domain of attraction. These results are shown schematically in Fig. 1(b) where each circled number indicates the number of steps it takes to map a certain nonperiodic cell into a periodic cell. In this paragraph, when we list the cells in the domains of attraction, we have not included the periodic cells themselves.

5 Two-Dimensional Linear Cell-to-Cell Mappings

Next, we examine the properties of two-dimensional systems in order to gain further insight to the cell-to-cell mappings. We shall confine our investigation to linear systems and we are particularly interested in comparing the properties of cell mapping with those of point mapping systems.

Consider a linear two-dimensional cell mapping system

$$\mathbf{Z}(n+1) = \mathbf{H}\mathbf{Z}(n) \tag{33}$$

or, in component form,

$$\begin{aligned} Z_1(n+1) &= H_{11}Z_1(n) + H_{12}Z_2(n) \\ Z_2(n+1) &= H_{21}Z_1(n) + H_{22}Z_2(n) \end{aligned} \tag{34}$$

where H_{11}, H_{12}, H_{21}, and H_{22} are all integers. Let

$$\begin{aligned} A &= \text{trace}\,\mathbf{H} = H_{11} + H_{22} \\ B &= \det \mathbf{H} = H_{11}H_{22} - H_{12}H_{21} \end{aligned} \tag{35}$$

$\mathbf{Z}^* = \mathbf{0}$ is an equilibrium or a $P - 1$ cell. An analysis can be carried out to study the trajectories around the equilibrium cell at $\mathbf{Z}^* = \mathbf{0}$. Here a trajectory from $\mathbf{Z}$ means the sequence of cell vectors $\mathbf{C}^k(\mathbf{Z})$, $k = 1, 2, \ldots$. The development is very similar to that given in [10] for two-dimensional point mapping systems, except that now A and B are necessarily integers. The general nature of the trajectories around $\mathbf{Z}^* = \mathbf{0}$, hence the character of $\mathbf{Z}^* = \mathbf{0}$, is entirely determined by $\mathbf{H}$ and in particular by A and B. We describe the character of the trajectories around $\mathbf{Z}^* = \mathbf{0}$ by studying various typical cases according to the values of A and B. For comparison the reader may wish to refer to Fig. 6 of [10] which classifies the $P - 1$ points for point mapping systems.

1 $A = 0$ and $B = 0$. Both eigenvalues of $\mathbf{H}$ are zero. Also by Cayley Hamilton theorem $\mathbf{H}^2 = \mathbf{0}$. This means starting with *any* cell, 2 steps of mapping will take this cell to the $P - 1$ cell at the origin. Thus the whole cell space is the 2-step domain of attraction of $\mathbf{Z}^* = \mathbf{0}$. Thus $\mathbf{Z}^* = \mathbf{0}$ is indeed a very strong attracting cell. For an example, take

$$\mathbf{H} = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}$$

The evolution pattern is shown in Fig. 2(1). All cells at $\mathbf{Z} = (a, -2a)$ are mapped into $\mathbf{Z}^* = \mathbf{0}$ in one step. All cells at $\mathbf{Z} = (b, a - 2b)$ with different values of b are mapped into $\mathbf{Z} = (a, -2a)$ in one step and into $\mathbf{Z}^* = \mathbf{0}$ in the second step of mapping.

We note here that for point mapping systems the case $A = 0$ and $B = 0$ implies an asymptotically stable $P = 1$ point at the origin.

2 $A = 1$ and $B = 0$. The eigenvalues of $\mathbf{H}$ are 0 and 1. The pattern of evolution is as follows. Besides the $P - 1$ cell at $\mathbf{Z}^* = \mathbf{0}$ there are $P - 1$ cells at $\mathbf{Z}^*$ with components Z_1^* and Z_2^* which meet the condition

$$Z_1^*: Z_2^* = H_{12}: -(H_{11} - 1)$$

Any other points are mapped into one of these $P - 1$ cells in one step. Thus this case is characterized by (*i*) the existence of infinite number of $P - 1$ cells and (*ii*) each one of these cells has a one-step domain of attraction and the collection of these one-step domains of attraction exhausts the whole cell space. For an example, consider

$$\mathbf{H} = \begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix}$$

The evolution is shown in Fig. 2(2). All cells at $\mathbf{Z}^* = (2a, a)$ are $P - 1$ cells and all cells at $\mathbf{Z} = (2a + b, a + b)$, $b = \pm 1, \pm 2, \ldots$, are mapped into the $P - 1$ cell at $\mathbf{Z}^* = (2a, a)$ in one step.

For the point mapping system the case of $A = 1$ and $B = 0$ corresponds to a borderline case between an asymptotically stable node and a saddle point.

For other typical cases we list below 11 additional ones:

3 $A = -1$ and $B = 0$. Example: $\mathbf{H} = \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix}$. Trajectories: Fig. 2(3).

4 $A = 0$ and $B = -1$. All cells are either $P - 1$ or $P - 2$.

5 $A = 2$ and $B = 1$. Example: $\mathbf{H} = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$. Trajectories: Fig. 2(4).

6 $A = -2$ and $B = 1$. Example: $\mathbf{H} = \begin{pmatrix} 1 & 2 \\ -2 & -3 \end{pmatrix}$. Trajectories: Fig. 2(5).

7 $A = 0$ and $B = 1$. Example: $\mathbf{H} = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}$. Trajectories: Fig. 2(6).

8 $A = -1$ and $B = 1$. Example: $\mathbf{H} = \begin{pmatrix} 1 & 1 \\ -3 & -2 \end{pmatrix}$. Trajectories: Fig. 2(7).

9 $A = 1$ and $B = 1$. Example: $\mathbf{H} = \begin{pmatrix} 2 & 1 \\ -3 & -1 \end{pmatrix}$. Trajectories: Fig. 2(8).

10 $A = 2$ and $B = 4$. Example: $\mathbf{H} = \begin{pmatrix} 1 & -3 \\ 1 & 1 \end{pmatrix}$. Trajectories: Fig. 2(9).

11 $A = 3$ and $B = 1$. Example: $\mathbf{H} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. Trajectories: Fig. 2(10).

12 $A = 1$ and $B = -1$. Example: $\mathbf{H} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Trajectories: Fig. 2(11).

13 $A = 5$ and $B = 5$. Example: $\mathbf{H} = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$. Trajectories: Fig. 2(12).

Here we have displayed the qualitative behavior of the trajectories for some typical cases. When compared with point mapping systems with the same values of A and B the qualitative behavior is more or less preserved. However, because a large number of patterns is possible, it is difficult to devise a set of convenient names for classification. We shall be content to let the pattern tell the character of the system. One special feature for cell mapping systems is worthy of note. While for linear point mapping systems the center may be surrounded by periodic motions of any period or aperiodic motion depending upon the specific value of A in the range of $-2 \leqq A \leqq 2$, for cell mapping systems only $P - 1$, $P - 2$, $P - 3$, $P - 4$, and $P - 6$ motions are possible.

6 Cell Mapping as a Discretization of a Point Mapping

In this section we consider cell mappings as resulted from discretizing point mappings according to certain specified rules. Consider a point mapping system as represented by (2). For discretizing this system we first divide the state space $\mathbf{x}$ into a collection of $\mathbf{Z}$ cells according to (3). Within this cell framework the associated cell mapping for a point mapping $\mathbf{G}$ is defined in the following manner. For a given $\mathbf{Z}(n)$ one finds the center point $\mathbf{x}^{(d)}(n)$ of the cell $\mathbf{Z}(n)$. Evidently the components of $\mathbf{x}^{(d)}(n)$ are given by

$$x_i^{(d)}(n) = h_i Z_i(n) \tag{36}$$

Next, one evaluates

$$\mathbf{x}^{(d)}(n + 1) = \mathbf{G}(\mathbf{x}^{(d)}(n)), \text{ or } x_i^{(d)}(n + 1) = G_i(\mathbf{x}^{(d)}(n)) \tag{37}$$

The cell in which $\mathbf{x}^{(d)}(n + 1)$ lies is then taken to be $\mathbf{Z}(n + 1)$. This process of discretization defines a cell mapping $\mathbf{C}$ in the form of (4). In component form we have

$$Z_i(n + 1) = \mathrm{C}_i(\mathbf{Z}(n)) = \mathrm{Int}\left\{\frac{1}{h_i} G_i(\mathbf{x}^{(d)}(n)) + \frac{1}{2}\right\} \tag{38}$$

where $\mathbf{x}^{(d)}(n)$ is related to $\mathbf{Z}(n)$ through (36) and Int (y) denotes the largest integer, positive or negative, which is less than or equal to y.

Once the associated cell mapping has been defined, it is of paramount interest to find out what local properties of a point mapping system near its periodic points are carried over to the associated cell mapping system and what properties are not. For this purpose we study in the next two sections the discretization of one-dimensional and two-dimensional linear point mapping systems. Many new concepts will be introduced in the study; it is therefore advantageous to begin with one-dimensional systems.

7 Discretization of One-Dimensional Linear Point Mapping Systems

Let the linear point mapping be given by

$$y(n + 1) = a + by(n) \tag{39}$$

It has a $P - 1$ point[3] at $y^* = a/(1 - b)$. Let h denote the cell size for discretization. Let

$$Y^* = \mathrm{Int}\left(\frac{y^*}{h} + \frac{1}{2}\right) \tag{40}$$

so that y^* lies in the Y^* cell of the y-axis. Moreover, let us write

[3] For appropriate terminology used here for point mappings, the reader is referred to [10].

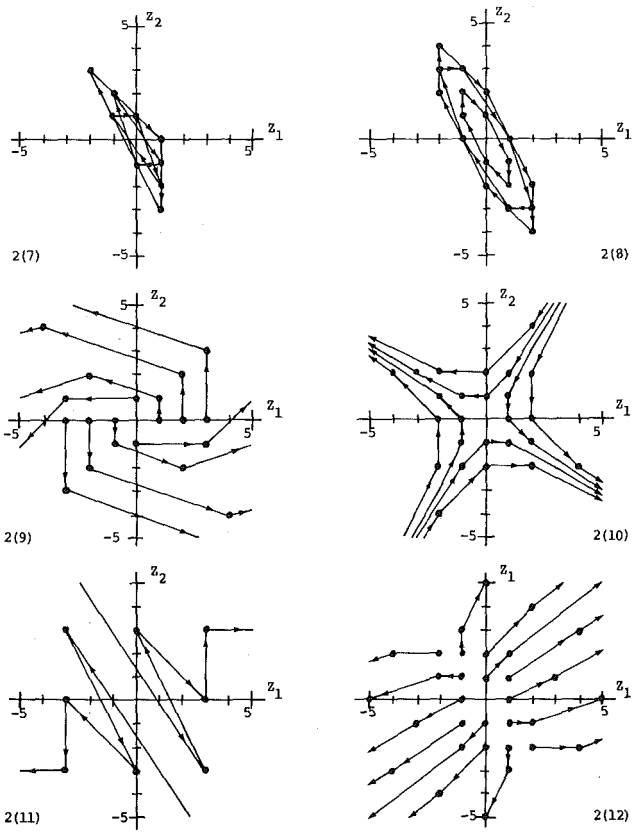


Fig. 2(7)–2(12) Various patterns of trajectories around the origin for two-dimensional linear cell-to-cell mappings

$$y^* = hY^* + h\delta, \quad -\frac{1}{2} \leqq \delta < \frac{1}{2} \tag{41}$$

so that $h\delta$ indicates the location of y^* away from the center of the Y^* cell; see Fig. 3. Next, let us introduce a new variable x with its origin of its coordinate axis located at the center of the Y^* cell; i.e.,

$$\begin{aligned} x(n+1) &= y(n+1) - hY^* \\ x(n) &= y(n) - hY^* \end{aligned} \tag{42}$$

In terms of $x(n)$, the point mapping is given by

$$x(n+1) = G(x(n)) = (1-b)\delta h + bx(n) \tag{43}$$

This is the point mapping to which we shall apply discretization. It is a mapping with a $P-1$ point at $x^* = \delta h$ which is locally stable or unstable depending upon whether $|b| \leqq 1$ or $|b| > 1$; see [10]. For this G we construct its associated cell mapping according to (38).

$$Z(n+1) = CZ(n)) = \text{Int}\{(1-b)\delta + bZ(n) + \tfrac{1}{2}\} \tag{44}$$

It is important to note here that while the original point mapping (43) is linear, the associated cell mapping C after discretization is in general not a linear one.

The first question we ask is whether the cell $Z = 0$ is always a $P-1$ cell? The answer is "not always." $Z = 0$ is a $P-1$ cell if and only if $C(0) = 0$, i.e.,

$$0 \leqq (1-b)\delta + \tfrac{1}{2} < 1 \tag{45}$$

The possibility of $Z = 0$ not being a $P-1$ cell can be readily seen in Fig. 3. If x^* is not located at the center of the cell (i.e., $\delta \neq 0$) and if the line $y(n+1) = a + by(n)$ has a very large slope, positive or negative, then the image of the point $x(n) = 0$ under G could very well give a value of $x(n+1)$ outside the range from $-h/2$ to $h/2$, and hence outside the cell $Z = 0$. Equation (45) may also be written as

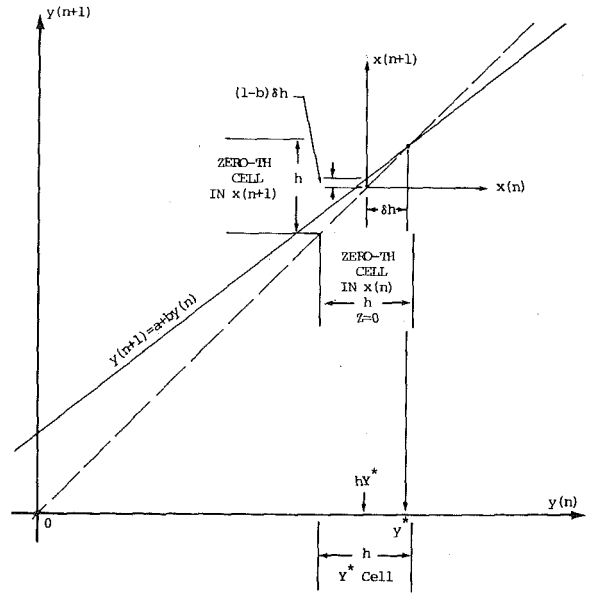


Fig. 3 Discretization of a one-dimensional point-to-point mapping

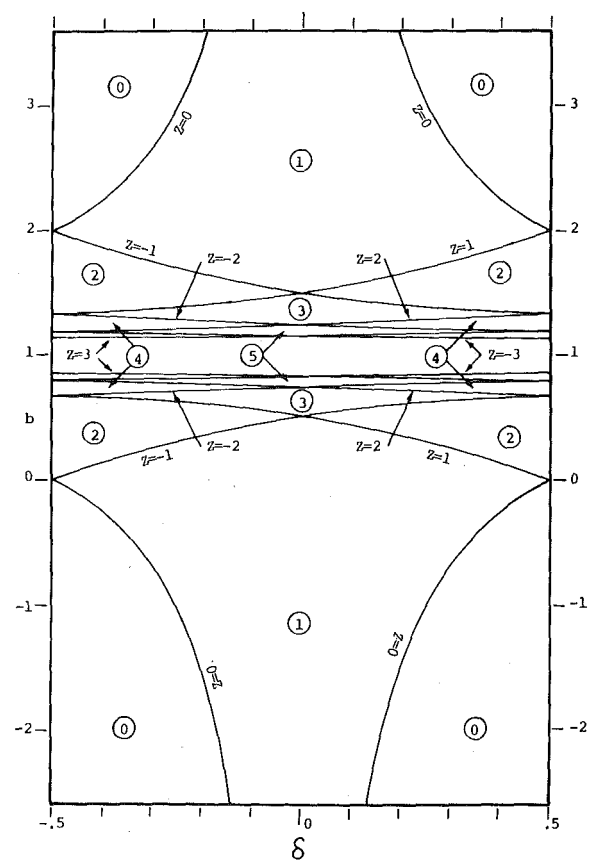


Fig. 4 Dependence of existence of various P-1 cells on δ and b

$$\begin{aligned} \frac{1}{\delta}\left(\frac{1}{2} + \delta\right) &\leqq b < \frac{1}{\delta}\left(-\frac{1}{2} + \delta\right) \quad \text{for} \quad \delta < 0 \\ \frac{1}{\delta}\left(-\frac{1}{2} + \delta\right) &< b \leqq \frac{1}{\delta}\left(\frac{1}{2} + \delta\right) \quad \text{for} \quad \delta > 0 \end{aligned} \tag{46}$$

In the δ-b parameter plane of Fig. 4 the region where $Z = 0$ is a $P-1$ cell is the central region between the four curves labeled $Z = 0$.

In general, in the neighborhood of $Z = 0$, a cell at $Z = M, M \neq 0$ will be a $P - 1$ cell if and only if

$$M \leqq (1 - b)\delta + bM + \tfrac{1}{2} < M + 1 \tag{47}$$

or

$$1 + \frac{1}{2(\delta - M)} \leqq b < 1 - \frac{1}{2(\delta - M)} \quad \text{for} \quad M > 0$$

$$1 - \frac{1}{2(\delta - M)} < b \leqq 1 + \frac{1}{2(\delta - M)} \quad \text{for} \quad M < 0 \tag{48}$$

In Fig. 4 the area between a pair of curves labeled $Z = M$ is the region where $Z = M$ is a $P - 1$ cell. One sees clearly that as b approaches 1 there will be more and more $P - 1$ cells surrounding $Z = 0$ to form a core of $P - 1$ cells. In Fig. 4 the number in circle for each region indicates the number of cells in the core of $P - 1$ cells for that region. The precise number of cells in this core of $P - 1$ cells, to be denoted by $S^{(1)}$, is given by

$$S^{(1)} = \left|\mathrm{Int}\left(|\delta| - \frac{1}{2(1-b)}\right)\right| + \left|\mathrm{Int}\left(|\delta| + \frac{1}{2(1-b)}\right)\right| \tag{49}$$

When b is nearly equal to one, $S^{(1)}$ will be rather large, an approximate formula for $S^{(1)}$ is

$$S^{(1)} \approx 1 + 2\,\mathrm{Int}\left(\frac{1}{2|b-1|}\right) \quad \text{for} \quad |b - 1| \ll 1 \tag{50}$$

We can also consider the conditions on b and δ in order for a cell at $Z = M_1$ to be mapped into a cell at $Z = M_2$. By (44), the conditions are

$$M_2 \leqq (1 - b)\delta + bM_1 + \tfrac{1}{2} < M_2 + 1 \tag{51}$$

Some of the results are shown in the various figures of Fig. 5. In each figure the different regions are shown in the δ-b parameter plane to indicate to which M_2 cell the cell M_1 is mapped. Figs. 5(1)–5(4) are, respectively, for $M_1 = 0, 1, 2, -1$. The area where $M_2 = M_1$ indicates the region where the cell at $Z = M_1$ is a $P - 1$ cell.

Next, consider possible $P - 2$ cells. Let $Z = M_1$ and $Z = M_2$ be the $P - 2$ cells of a $P - 2$ motion. They must satisfy

$$\begin{aligned} M_2 &\leqq (1 - b)\delta + bM_1 + \tfrac{1}{2} < M_2 + 1 \\ M_1 &\leqq (1 - b)\delta + bM_2 + \tfrac{1}{2} < M_1 + 1 \end{aligned} \tag{52}$$

Adding bM_2 to the first of (52) and bM_1 to the second, we can change (52) to

$$\begin{aligned} (1 + b)M_2 &\leqq (1 - b)\delta + b(M_1 + M_2) + \tfrac{1}{2} < (1 + b)M_2 + 1 \\ (1 + b)M_1 &\leqq (1 - b)\delta + b(M_1 + M_2) + \tfrac{1}{2} < (1 + b)M_1 + 1 \end{aligned} \tag{53}$$

When $b > 0$ or when $b < -2$ the interval $\{(1 + b)M_2, (1 + b)M_2 + 1\}$ and the interval $\{(1 + b)M_1, (1 + b)M_1 + 1\}$ cannot overlap. Therefore (53) cannot be satisfied simultaneously; hence, there can be no $P - 2$ cells. We need only to confine our attention to $-2 < b < 0$ for possible existence of $P - 2$ cells.

By examining Fig. 5(1) one notes easily that there is a region where the cell $Z = 0$ is mapped into the cell $Z = 1$. By Fig. 5(2) there is a region where the cell $Z = 1$ is mapped into the cell at $Z = 0$. These two regions overlap. The overlapped part is obviously the region where a $P - 2$ motion involving the cells $Z = 0$ and $Z = 1$ exists. In Fig. 6 this region is shown as the triangular shaped area $P_1P_4P_5$. Similarly, there are regions for other $P - 2$ motions. They are listed as follows:

$P - 2$ Motion: $\{Z = 0, Z = 1\}$—Area $P_1P_4P_5$
$P - 2$ Motion: $\{Z = -1, Z = 2\}$—Area $P_2P_3P_5$

$P - 2$ Motion: $\{Z = -1, Z = 1\}$—Area $P_5P_9P_{10}P_6$
$P - 2$ Motion: $\{Z = -2, Z = 2\}$—Area $P_5P_8P_{10}P_7$

$P - 2$ Motion: $\{Z = -1, Z = 0\}$—Area $P_{10}P_{14}P_{11}$
$P - 2$ Motion: $\{Z = -2, Z = 1\}$—Area $P_{10}P_{13}P_{12}$

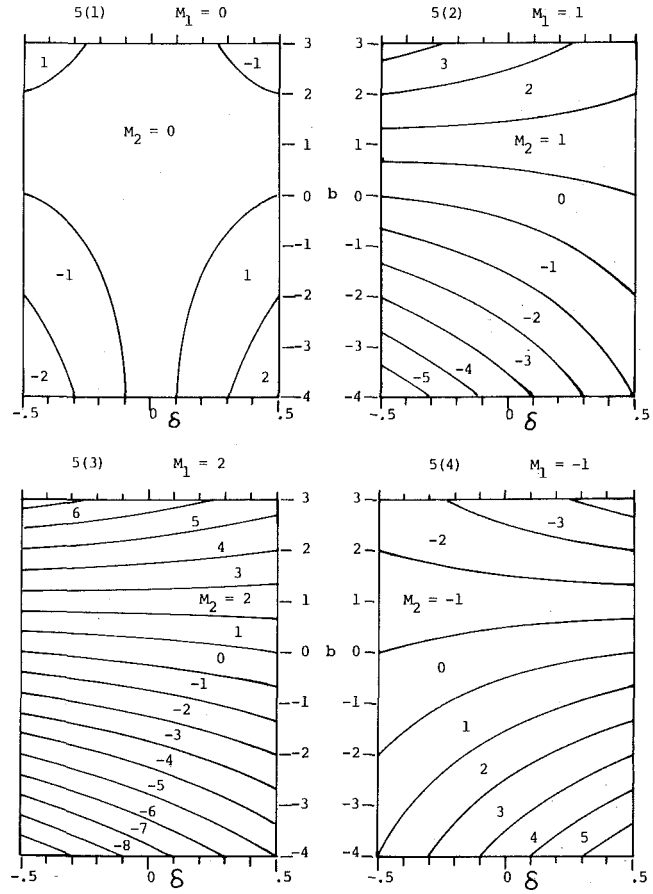


Fig. 5 Conditions on δ and b for a cell $Z = M_1$ to be mapped into $Z = M_2$

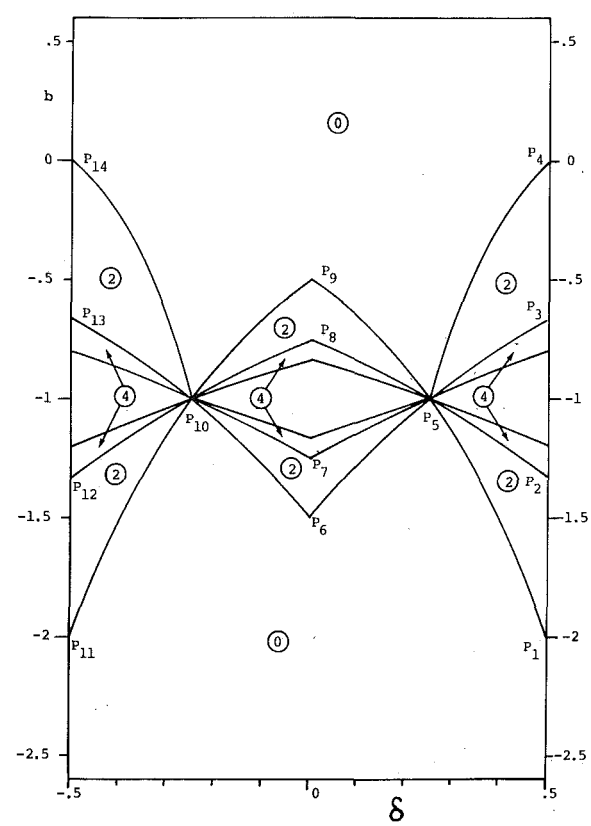


Fig. 6 Dependence of existence of various P-2 cells on δ and b

The dotted lines in the foregoing list denote other $P-2$ motions involving cells further away from $Z = 0$. The pattern is clear. The number of possible $P-2$ cells increases without limit as $b \to -1$.

The number of $P-2$ cells, denoted by $S^{(2)}$, may be computed as follows:

(*i*) For $0 < b$ or $b < -2$:

$$S^{(2)} = 0. \tag{54}$$

(*ii*) For $-1/4 < |\delta| < 1/4$:

$$b < 1 - \frac{1}{2|\delta|}, \quad S^{(2)} = 0$$

$$1 - \frac{1}{2|\delta|} < b < -1, \quad S^{(2)} = 2 \times \mathrm{Int}\left(\frac{(1-b)|\delta| - 0.5}{1+b}\right)$$

$$-1 < b < 0, \quad S^{(2)} = 2 \times \mathrm{Int}\left(\frac{0.5 - (1-b)|\delta|}{1+b}\right) \tag{55}$$

(*iii*) For $1/4 < |\delta| < 1/2$:

$$1 - \frac{1}{2|\delta|} < b < 0, \quad S^{(2)} = 0$$

$$-1 < b < 1 - \frac{1}{2|\delta|}, \quad S^{(2)} = 2 \times \mathrm{Int}\left(\frac{0.5 + (1-b)|\delta| + b}{1+b}\right)$$

$$b < -1, \quad S^{(2)} = 2 \times \mathrm{Int}\left(\frac{0.5 - (1-b)|\delta|}{1+b}\right) \tag{56}$$

A serarch has shown that there are no other $P-K$ cells with $K > 2$. Therefore, the $P-1$ and $P-2$ cells together form a core of periodic cells. The size of the core is the sum of $S^{(1)}$ and $S^{(2)}$. Let $S = S^{(1)} + S^{(2)}$. The core size approaches infinity as $|b| \to 1$. This is to be expected because a point mapping system with $b = 1$ has every point as a $P-1$ point and a point mapping system with $b = -1$ has the origin as a $P-1$ point and every other point as a $P-2$ point. One also notes here that there are values of b and δ such that no periodic cells exist or the core size is zero. However this could happen only if $|b| > 1$, corresponding to an unstable $P-1$ point at $x = 0$ for the point mapping system.

Let us denote the size of the core of periodic cells by $S(b, \delta)$ to indicate its dependence on b and δ. For $b > 0$ the variation is slight. For a given $b > 0$, $S(b,\delta)$ changes by no more than one unit in the whole range of δ. For $b < 0$ the variation of $S(b,\delta)$ with δ is more complicated. However, for a given b, whether positive or negative as long as $|b| \neq 1$, there exists a finite number $S_{\max}(b)$.

$$S_{\max}(b) = \max_{-0.5 \leq \delta \leq 0.5} S(b,\delta) \tag{57}$$

The existence of a finite maximum possible core size $S_{\max}(b)$ which is independent of δ and independent of the cell size h leads to a most important result. As is seen here, a $P-1$ point of a linear point mapping system is replaced by a core of $P-1$ and/or $P-2$ cells. In other words pseudo periodic cells are introduced by the process of discretization. The existence of $S_{\max}(b)$ implies however that if we reduce the cell size h to get a finer and finer discretization one can expect that the "physical size" of the core $[h \cdot S(b, \delta)]$ is bounded from above by $[h \cdot S_{max}(b)]$. Therefore, *as h approaches to zero, the physical size of the core of periodic cells also approaches to zero.*

The next question is whether the stability character of the $P-1$ point at $x^* = 0$ of the point mapping system (43) is preserved. If preserved, in what sense.

(*i*) $b > 1$. Let $Z = M$ be a $P-1$ cell at the right edge of the core of $P-1$ cells, i.e.,

$$C(M) = \mathrm{Int}\{(1-b)\delta + bM + \tfrac{1}{2}\} = M \tag{58}$$

but

$$C(M+1) = \mathrm{Int}\{(1-b)\delta + b(M+1) + \tfrac{1}{2}\} \neq M+1 \tag{59}$$

Because of (58) and $b > 1$, $\mathrm{Int}\{(1-b)\delta + b(M+1) + \frac{1}{2}\}$ has to be $\geqq M + 1$. However, $C(M+1) \neq M+1$, therefore, $C(M+1) > M+1$. Thus the cell $Z = M+1$ just outside the core is mapped further away from the core. One can show a similar result for the cell just outside the left edge of the core. The core is therefore a repulsive one for $b > 1$.

(*ii*) $0 < b < 1$. By (58) and the condition $0 < b < 1$, $\mathrm{Int}\{(1-b)\delta + b(M+1) + \frac{1}{2}\}$ has to be either M or $M+1$. However, $C(M+1) \neq M+1$, therefore $C(M+1) = M$. Thus the cell $Z = M+1$ is mapped into the core in one step. Similarly the cell outside the left edge of the core is mapped into the core in one step. The core is an attracting one for $0 < b < 1$.

(*iii*) $-1 < b < 0$. Consider first the case where $Z = 0$ is the only $P-1$ cell and there are no $P-2$ cells.

$$C(0) = \mathrm{Int}\{(1-b)\delta + \tfrac{1}{2}\} = 0 \tag{60}$$

One can evaluate $C(1)$ and find $C(1)$ has to be either -1 or 0. Similarly, $C(-1)$ has to be either 0 or $+1$. There are four possibilities.

(*a*) $C(1) = 0$ and $C(-1) = 0$. This says that the cells just outside the core are mapped into the core in one step.

(*b*) $C(1) = 0$ and $C(-1) = 1$. Again the cells just outside the core are mapped into the core in one step or in two steps.

(*c*) $C(1) = -1$ and $C(-1) = 0$. Same as in (*b*).

(*d*) $C(1) = -1$ and $C(-1) = 1$. This case is ruled out as we assume here nonexistence of $P-2$ cells. Therefore, in this case the core is an attracting one.

Consider next the case of a core consisting of $P-1$ and $P-2$ cells. Let $Z = M_1$ and $Z = M_2$ be the two edge cells of the core with $M_2 > M_1$. Then

$$\begin{aligned} M_2 &= C(M_1) = \mathrm{Int}\{(1-b)\delta + bM_1 + \tfrac{1}{2}\} \\ M_1 &= C(M_2) = \mathrm{Int}\{(1-b)\delta + bM_2 + \tfrac{1}{2}\} \end{aligned} \tag{61}$$

Because of (61) and $-1 < b < 0$, $C(M_1 - 1)$ has to be either M_2 or $M_2 + 1$ while $C(M_2 + 1)$ has to be either M_1 or $M_1 - 1$. Again, by considering the four possibilities and recognizing that $Z = M_1 - 1$ and $Z = M_2 + 1$ are not $P-2$ cells, one can show that the core is an attracting one.

(*iv*) $b < -1$. In a similar manner we can show that the core is a repulsive one for $b < -1$.

Summary of the Main Results. By going from a linear point mapping system to its associated cell mapping system, the $P-1$ point of the point mapping is replaced by a core of $P-1$ and $P-2$ cells. The number of the cells in the core is bounded for $|b| \neq 1$ and the bound depends only upon b and is independent of the cell size h. The stability character of the linear point mapping is preserved. An asymptotically stable $P-1$ point of the point mapping is replaced by an attracting core of periodic cells for the cell mapping system and an unstable $P-1$ point of the point mapping is replaced by a repulsive core of the cell mapping system. Moreover, as the cell size h approaches zero the physical size of the core also approaches zero.

8 Discretization of Two-Dimensional Linear Point Mapping Systems

Next, we consider two-dimensional linear point mapping systems and its associated cell mapping systems. Consider

$$\begin{pmatrix} x_1(n+1) \\ x_2(n+1) \end{pmatrix} = \begin{pmatrix} 1 - h_{11}, & -h_{12} \\ -h_{21}, & 1 - h_{22} \end{pmatrix} \begin{pmatrix} \delta_1 h_1 \\ \delta_2 h_2 \end{pmatrix} + \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix} \tag{62}$$

where $|\delta_1| \leqq 1/2, |\delta_2| \leqq 1/2$. This linear point mapping system has a $P-1$ point at a point $(\delta_1 h_1, \delta_2 h_2)$. Since we are only interested in the connection between the qualitative behavior of this system and that of its associated cell mapping, we shall assume that the $\mathbf{x}$-coordinate system is already such that matrix $\mathbf{H}$ is in its canonical forms; see p. 267 of [10]. Of the three canonical forms we shall discuss only one case here for which $\mathbf{H}$ in its canonical form is diagonal. For simplicity we shall also take $h_1 = h_2 = h$. The point mapping is then given by

$$x_1(n+1) = (1 - b_1)\delta_1 h + b_1 x_1(n) \tag{63a}$$

$$x_2(n+1) = (1 - b_2)\delta_2 h + b_2 x_2(n) \tag{63b}$$

Here, we have written h_{11} and h_{22} as b_1 and b_2, respectively. The two

equations are uncoupled. The discussion given in Section 7 for one-dimensional linear point mapping systems is then applicable to each. The associated cell mapping[4] is given by

$$Z_1(n+1) = C_1(Z_1(n)) = \text{Int}\{(1-b_1)\delta_1 + b_1 Z_1(n) + \tfrac{1}{2}\} \quad (64a)$$

$$Z_2(n+1) = C_2(Z_2(n)) = \text{Int}\{(1-b_2)\delta_2 + b_2 Z_2(n) + \tfrac{1}{2}\} \quad (64b)$$

A $P-1$ cell for C_1 and a $P-1$ cell for C_2 leads to a two-dimensional $P-1$ cell for $\mathbf{C}$. A $P-1$ cell for C_1 and a $P-2$ cell for C_2, a $P-2$ cell for C_1 and a $P-1$ cell for C_2, and a $P-2$ cell for C_1 and a $P-2$ cell for C_2 will all lead to a two-dimensional $P-2$ cell for $\mathbf{C}$. If S_1 is the number of cells in the core of periodic cells for (64a) and S_2 the number for (64b), then the two-dimensional cell mapping will have a core size equal to $S_1 \times S_2$. Also, the physical size of the two-dimensional case approaches zero as h approaches zero.

Let us now define a core of periodic cells to be attracting if it is attracting in both directions, and to be repulsive if there is one direction in which the core is repulsive. With these definitions at hand, it is then easy to show that if $|b_1| < 1$ and $|b_2| < 1$, then the two-dimensional core of periodic cells is an attracting one. If $|b_1| > 1$ or $|b_2| > 1$, then the core is a repulsive one. Thus the stability character of the point mapping systems is preserved in the associated cell mapping systems.

When the $\mathbf{H}$ has complex and conjugate eigenvalues, the canonical form may be put in the form

$$\mathbf{H} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \quad (65)$$

The associated cell mapping for this case is very interesting but is also very complex. Therefore this case will not be discussed in this introductory paper. The results will be given in a separate paper where a very detailed study of two-dimensional systems will be presented. It suffices to say here that again there are pseudo periodic cells introduced by the discretization into a cell mapping. The pseudo periodic cells increase in number as $\alpha^2 + \beta^2 \to 1$ but for a given set of α and β, $\alpha^2 + \beta^2 \neq 1$, there is an upper bound on the number of pseudo periodic cells. Also the stability character of the point mapping system is preserved in the cell mapping systems.

[4] Again, we emphasize that the associated cell mapping system to a linear point mapping system is in general not linear.

9 A Concluding Remark

It might be appropriate to remark here in concluding the paper that it is intended only to be an introduction to the theory of cell mappings. Many important theoretical questions remain to be studied and applications other than that given in [14] are to be explored.

Acknowledgments

The results presented here were obtained in the course of research supported by a grant from the National Science Foundation.

References

1 Poincaré, H., "Sur les courbes définies par les équations différentielles," *J. Math. Pure Appl.* Series (3), Vol. 7, 1881, pp. 275–422.

2 Birkhoff, G. D., "Surface Transformations and Their Dynamical Applications," *Acta Math.*, Vol. 43, 1920, pp. 1–119.

3 Arnol'd, V. D., "Small Denominators and Problems of Stability of Motion in Classical and Celestial Mechanics," *Russian Mathematical Survey,* Vol. 18, 1963, pp. 85–191.

4 Smale, S., "Differentiable Dynamical Systems," *Bulletin of the American Mathematical Society,* Vol. 73, 1967, pp. 747–817.

5 Bernussou, J., *Point Mapping Stability,* Pergamon Press, New York, 1977.

6 Marsden, J. E., and McCracken, M., *The Hopf Bifurcation and Its Applications,* Springer-Verlag, New York, 1976.

7 Hsu, C. S., Yee, H. C., and Cheng, W. H., "Determination of Global Regions of Asymptotic Stability for Difference Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 147–153.

8 Hsu, C. S., Yee, H. C., and Cheng, W. H., "Steady-State Response of a Nonlinear System Under Impulsive Periodic Parametric Excitation," *Journal of Sound and Vibration,* Vol. 50, 1977, pp. 95–116.

9 Hsu, C. S., "Nonlinear Behavior of Multibody Systems Under Impulsive Parametric Excitation," *Dynamics of Multibody Systems,* ed., Magnus, K., Springer-Verlag, 1978, pp. 63–74.

10 Hsu, C. S., "On Nonlinear Parametric Excitation Problems," *Advances in Applied Mechanics,* Vol. 17, 1977, pp. 245–301.

11 Hsu, C. S., "A Theory of Index for Point Mapping Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 185–190.

12 Troger, H., "On Point Mappings for Mechanical Systems Possessing Homoclinic and Heteroclinic Points," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 468–469.

13 Henon, M., "Numerical Study of Quadratic Area-Preserving Mappings," *Quarterly of Applied Mathematics,* Vol. 27, 1969, pp. 291–312.

14 Hsu, C. S., and Guttalu, R. S., "An Unravelling Algorithm for Global Analysis of Dynamical Systems: An Application of Cell-to-Cell Mappings," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 940–948.

C. S. Hsu
Professor of Applied Mechanics.
Fellow ASME

R. S. Guttalu
Research Assistant.

Department of Mechanical Engineering,
University of California,
Berkeley, Calif. 94720

An Unravelling Algorithm for Global Analysis of Dynamical Systems: An Application of Cell-to-Cell Mappings

A new method is offered here for global analysis of nonlinear dynamical systems. It is based upon the idea of constructing the associated cell-to-cell mappings for dynamical systems governed by point mappings or governed by ordinary differential equations. The method uses an algorithm which allows us to determine in a very effective manner the equilibrium states, periodic motions and their domains of attraction when they are asymptotically stable. The theoretic base and the detail of the method are discussed in the paper and the great potential of the method is demonstrated by several examples of application.

1 Introduction

For dynamical systems subjected to periodic excitation, it is possible in principle to reformulate the governing equations in the form of point-to-point mappings of Poincaré.

$$\mathbf{x}(n+1) = \mathbf{G}(\mathbf{x}(n)) \tag{1}$$

In recent years Poincaré maps have seen increasing applications in many fields of engineering and science [1, 2]. They have been used in [3–5] to study certain nonlinear mechanical systems under parametric excitation.

In analyzing a strongly nonlinear system, one is often interested in the global pattern of its motions in addition to the local behavior around its equilibrium states and periodic motions. However, for most nonlinear systems and for nonlinear point mapping systems (1) in particular, even seemingly very simple ones can have extremely complicated global behavior and they are in general very difficult to analyse. Moreover, there does not seem to be a general and standard method available for such global analyses. The most effective way at the present time is probably still the straightforward numerical evaluation to compute the motions and then to infer some global properties of the system from the numerical results. This procedure is however, very time-consuming and inefficient unless one is only interested in a very limited scope of the global picture. It is for this reason that any new method which could facilitate the study of the global behavior of strongly nonlinear systems would be most welcome.

In this paper we present such a method of global analysis. The method will be discussed first in the context of point-to-point mapping systems. It is based upon the idea of *cell-to-cell mapping*, a theory of which has been presented in [6]. Through a process of discretization a point mapping system is replaced by a cell-to-cell mapping system

$$\mathbf{Z}(n+1) = \mathbf{C}(\mathbf{Z}(n)) \tag{2}$$

Where $\mathbf{C}$ is the mapping and the components of $\mathbf{Z}$ take on only integer values. Using this idea of cell to cell mapping, an algorithm has been devised which allows us to determine in one run of "computing" all the periodic cells and all the domains of attraction of the periodic cells in the region of the state space which is of interest to us. Since this method of global analysis and the theory of cell-to-cell mappings are both new subjects of development, many novel concepts and new terminology are necessary. The reader is referred to [6] for any terms which might be unfamiliar to him and which are not explained in this paper.

In Section 2 we discuss briefly the cell-to-cell mappings and explain a set of terminology which is used repeatedly in the paper. In Section 3 the basic idea of the algorithm of global analysis is discussed first and specific methods of implementing the algorithm will be presented. We then illustrate the method by applying it in Section 4 to a simple two-dimensional point mapping system for which we have a reasonably complete understanding of its global behavior. In Section 5 we apply the method to the point mapping of a mechanical system with a single degree of freedom which is known to have very complicated global behavior.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Chicago, Ill., November 16–21, 1980, 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 March 1, 1981. Readers who need more time to prepare a Discussion should request an extension from the Editorial Department. Manuscript received by ASME Applied Mechanics Division, May 1980; final revision, June, 1980. Paper No. 80-WA/APM-27.

The application of the algorithm need not be restricted to point mapping and cell mapping systems. In Section 6 we show how the application may be extended to dynamical systems governed by nonlinear ordinary differential equations.

Being intended as an introduction to a new method, the paper does not pretend to present all aspects of the method in the most elegant and definitive form. Many further developments await. We do hope to show here that the method is very effective and has merit. Moreover, it is particularly tailored for computer usage and, therefore, its utility could be expected to increase with the advancement of computer development.

2 Cell-to-Cell Mappings

An N-dimensional state space is divided into a large number of cells. Each cell is identified by an N-tuple of integers. The evolution of the dynamical system is then recast in the form of a cell-to-cell mapping. Such a cell-to-cell mapping is denoted by (2) where $\mathbf{Z}$ is the N-dimensional cell vector whose components can take on only integer values. For the philosophical and scientific basis of such a reformulation the reader is referred to the discussion given in the introductory section of [6].

Periodic Motions and Periodic Cells. Let $\mathbf{C}^m$ denote the cell mapping $\mathbf{C}$ applied m times with $\mathbf{C}^0$ understood to be the identity mapping. A sequence of K distinct cells $\mathbf{Z}^*(j), j = 1, 2, \ldots, K$ which satisfy

$$\mathbf{Z}^*(m+1) = \mathbf{C}^m(\mathbf{Z}^*(1)), \quad m = 1, 2, \ldots, K-1,$$
$$\mathbf{Z}^*(1) = \mathbf{C}^K(\mathbf{Z}^*(1)), \tag{3}$$

is said to form a periodic motion of period K. For convenience we call such a motion a *P-K motion.* We call each of its elements $\mathbf{Z}^*(j)$ a *periodic cell of period K* or simply a *P-K cell.* A $P-1$ cell may also be regarded as an equilibrium cell under the cell mapping $\mathbf{C}$.

Domains of Attraction. A cell $\mathbf{Z}$ is said to be *"r-steps removed from a P-K motion"* if r is the minimum positive integer such that $\mathbf{C}^r(\mathbf{Z}) = \mathbf{Z}^*(j)$ where $\mathbf{Z}^*(j)$ is one of the P-K cells of that P-K motion. In other words, $\mathbf{Z}$ is mapped after r-steps into one of the P-K cells of the P-K motion and any further mapping will lock the evolution of the system in this P-K motion. The set of all cells which are r-steps or less removed from a P-K motion is called the *"r-step domain of attraction"* for that P-K motion. The *total domain of attraction,* or simply the *domain of attraction of a P-K motion* is its r-step domain of attraction with $r \to \infty$.

One of the main purposes of any global analysis is to determine the domains of attraction.

3 An Unravelling Algorithm of Global Analysis

In this section we describe the basic idea of the algorithm which allows us to determine the periodic cells and their domains of attraction in a very effective manner. The algorithm will be explained using one-dimensional cell-to-cell mappings but it will be seen later that its application is not limited to one-dimensional systems.

Consider a one-dimensional cell mapping (2) where the state cell vector $\mathbf{Z}$ becomes now an integer scalar. Before discussing the algorithm let us first introduce a most important and necessary concept of a *"sink cell."*

Sink Cell. For most physical problems once the state variable exceeds a certain positive or negative magnitude, one is no longer interested in the further evolution of the variable. This implies that there is only a range of the value of the state variable which is of practical interest to us. In terms of cells of the state variable, it means that one is only interested in cells lying between a lower limit cell $Z^{(L)}$ and a upper limit cell $Z^{(U)}$. Let $N_c = Z^{(U)} - Z^{(L)} + 1$. Then N_c is the total number of cells one is actually interested in. Let us call these cells the *"regular cells."* We take the view that once the evolution of the system takes it to a cell $Z < Z^{(L)}$ or $Z > Z^{(U)}$, then we are no longer interested in its further evolution. This notion allows us to introduce a *sink cell* which is defined to be a very large all encompassing cell containing all the cells Z with $Z < Z^{(L)}$ or $Z > Z^{(U)}$. Viewed in this framework, the total number of cells one is concerned with will always be *finite* although it could be and usually will be *huge.* Moreover, for convenience one can relabel all the regular cells by positive integers, $1, 2, \ldots, N_c$. The sink cell can be labelled as the 0th cell. In the remainder of this section we shall assume that any cell mapping $\mathbf{C}$ under discussion has been so standardized that the cell designation uses only non-negative integers with $Z = 0$ to be the sink cell. Then for mapping between regular cells we have

$$Z(n+1) = C(Z(n)) \quad \text{for } Z(n+1), \quad Z(n) = 1, 2, \ldots, N_c \tag{4}$$

This mapping is supplemented by two additional rules of mapping:

(*i*) If $Z(n) = 1, 2, \ldots, N_c$ and $C(Z(n)) < 1$ or $C(Z(n)) > N_c$, then $Z(n+1) = C(Z(n))$ is set equal to zero.

(*ii*) $C(0) = 0$. This second rule says that the sink cell is a $P-1$ cell and, therefore, once the system is mapped into the sink cell it stays there.

Once it is recognized that the total number of cells involved, the regular cells and the sink cell, will be finite, several far-reaching consequences follow. These make the proposed algorithm workable and efficient.

(*i*) By definition the sink cell is a $P-1$ cell.

(*ii*) Among the regular cells there could be periodic cells belonging to various periodic motions. The number of the periodic cells could be very large but it cannot exceed N_c. The periodic motions can have various periods but the period of any periodic motion cannot exceed N_c.[1]

(*iii*) The evolution of the system starting with any regular cell Z can lead to only three possible outcomes:

(*iii*-1) Cell Z is itself a periodic cell of a periodic motion. The evolution of the system is simply a periodic motion.

(*iii*-2) Cell Z is mapped into the sink cell in r-steps. Then the cell belongs to the r-step domain of attraction of the sink cell.

(*iii*-3) Cell Z is mapped into a periodic cell of a certain periodic motion in r-steps. Thereafter, the evolution is locked into that periodic motion. In this case the cell belongs to the r-step domain of attraction of that periodic motion.

To delineate the global properties of a cell Z we introduce three numbers. They are the *group number* Gr (Z), *periodicity number* $P(Z)$ and the *steps number* $S(Z)$. To each existing periodic motion we assign a group number. This group number is then given to every periodic cell of that periodic motion and also to every cell in the domain of attraction of that periodic motion. The group numbers, positive integers, can be assigned sequentially as the periodic motions are discovered one by one during the global analysis. Obviously, there will be as many group numbers as there are periodic motions. Each group has an invariant set in the form of a periodic motion which has a definite period. A periodicity number equal to this period is assigned to all the cells in the group. If a cell Z has a periodicity number $P(Z) = K$, then the cell Z is either a P-K cell itself or it belongs to the domain of attraction of a P-K motion. The steps number $S(Z)$ of a cell Z is used to indicate how many steps it takes to map this cell Z into a periodic cell. If Z happens to be a periodic cell then $S(Z) = 0$. In this manner the global properties of a cell Z is entirely characterized by the numbers Gr (Z), $P(Z)$, and $S(Z)$. The purpose of our global analysis is to determine these three numbers for every regular cell when the cell mapping $C(Z)$ is given.

Now we are ready to describe the algorithm of the proposed global analysis which will be seen to be an extremely simple one. Four one-dimensional arrays will be used. They are $C(Z)$, Gr (Z), $P(Z)$, and $S(Z)$ with $Z = 0, 1, 2, \ldots, N_c$. $C(Z)$ is the array which defines the cell mapping, i.e., a cell at Z is mapped into cell $C(Z)$. $Z = 0$ is used as the sink cell; therefore, $C(0) = 0$. The other three arrays give the group number, the periodicity number and the steps number for each cell

[1] We note here that when the new framework of finite number of cells is adopted, chaotic motions no longer have meaning. One must accept the situation of not having chaotic motions but only having $P-K$ motions with possibly very large K.

as explained in the last paragraph, except that the group number Gr (Z) will be used for an additional purpose to be explained shortly.

The algorithm involves essentially calling up a cell and then having it processed in a certain way in order to determine its global character. To be viable and efficient it should have the capability of distinguishing three kinds of cells. Belonging to the first kind are the cells which have not yet been called up by the program. They will be called *virgin cells*. All Gr (Z), $Z = 0, 1, 2, \ldots, N_c$, are set to zero at the beginning of a run of the program. Therefore, a virgin cell Z is characterized by having Gr $(Z) = 0$. Cells of the second kind are those which have been called up and are currently being processed. They will be called *cells under processing*. If we adopt the rule that a virgin cell once called up will have its group number reassigned as -1, then all the cells under processing will be characterized by having -1 as their group number. Here the number -1 is simply used as a symbol of identification. Finally there are cells of the third kind. They are the cells whose global properties have been determined and whose group numbers, periodicity numbers and steps numbers have all been assigned. They will be called *processed cells* and they are identified as having positive integers as their group numbers. In this manner Gr (Z) serves also as an identification flag of processing before it is permanently assigned.

In processing the cells we make repeated use of sequences formed by the mapping. Starting with a cell Z, we form

$$Z \rightarrow C(Z) \rightarrow C^2(Z) \rightarrow \ldots \rightarrow C^m(Z) \tag{5}$$

Such a sequence will be called a *mth-order processing sequence on* Z. In the algorithm we process the cells in a sequential manner, $Z = 0, 1, 2, \ldots, N_c$. The idea is to begin with a virgin cell Z and examine the processing sequence $C^i(Z)$ as defined by (5). At each step in generating this sequence there are three possibilities:

(*i*) The newly generated element $C^i(Z)$ is such that Gr $(C^i(Z)) = 0$ indicating that the cell $C^i(Z)$ is a virgin cell. In this case we continue forward to locate the next cell $C^{i+1}(Z)$ in the processing sequence. Before doing that we first set Gr $(C^i(Z)) = -1$ in order to indicate that $C^i(Z)$ is no longer a virgin cell but a cell under processing.

(*ii*) The newly generated cell $C^i(Z) = Z'$ is found to have a positive integer as its Gr (Z') number. This indicates that $C^i(Z)$ has appeared in one of the previous processing sequences and its global character has already been determined. In this case the current processing sequence is terminated at this point. Since this current processing sequence is mapped into a cell with known global properties, the global character of all the cells in the sequence is easily determined. Obviously all the cells of the present processing sequence will have the same group number and the same periodicity number as that of Z'. The steps number of each cell in the sequence is simply

$$S(C^j(Z)) = S(Z') + i - j, \quad j = 0, 1, 2, \ldots, i. \tag{6}$$

Once these global character numbers have been assigned, the work on this processing sequence is completed and we go back to the cell sequence to pick the next virgin cell to begin a new processing sequence.

(*iii*) The newly generated cell $C^i(Z) = Z''$ is found to have -1 as its group number. This indicates that $C^i(Z)$ has appeared before in the present sequence. Therefore, there is a periodic motion contained in the sequence and, moreover, the periodic motion is a new one. In this case again the processing sequence is terminated. To all cells in the sequence is assigned now a new cell group number which is one larger than the number of groups already determined. Next, it is a simple matter to determine the position in the sequence where the cell $C^i(Z)$ reappears. Let $C^i(Z)$ reappear in the $(j + 1)$th position of the sequence, i.e., $C^i(Z) = C^j(Z)$, $j < i$. The periodicity of the periodic motion is $i - j$, and all cells in the processing sequence are assigned $(i - j)$ as their periodicity numbers. With regard to the steps number, we have

$$S(C^k(Z)) = j - k, \quad k = 0, 1, 2, \ldots, j - 1,$$
$$S(C^k(Z)) = 0, \quad k = j, j + 1, \ldots, i - 1. \tag{7}$$

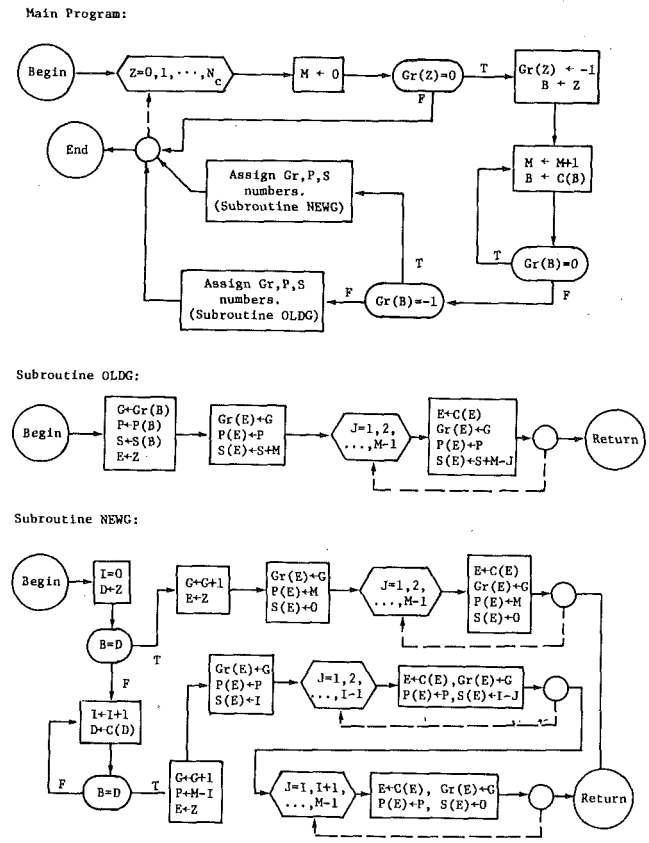


Fig. 1 Flow chart for the algorithm

Once these global character numbers have been assigned, the work on this processing sequence is finished and we go back to the cell sequence to pick the next virgin cell to begin a new processing sequence.

Using these processing sequences starting with virgin cells, the whole cell space $Z = 0, 1, 2, \ldots, N_c$ is covered and the global character of all cells are determined in terms of the numbers Gr (Z), $P(Z)$, and $S(Z)$. What follows are some discussions to tie up the loose ends.

The first processing sequence begins with $Z = 0$. Since $Z = 0$ is the sink cell and is a $P - 1$ cell, the processing sequence has only two elements (one cell):

$$Z = 0, \quad C^1(0) = 0. \tag{8}$$

Group number 1 is assigned to this sequence; therefore, Gr $(0) = 1$. Periodicity number is 1, $P(0) = 1$, and the steps number for $Z = 0$ is 0, $S(0) = 0$. This completes the action on this first processing sequence.

Next, we go back to the cell sequence to take $Z = 1$ as the starting cell for the second processing sequence. This sequence will terminate either because a member $C^i(1) = 0$, i.e., mapped into the sink cell, or because the sequence leads to a periodic motion within itself. In the former case all cells in the sequence will have the cell group number of the sink cell, No. 1. In the latter case a new cell group number, No. 2, is assigned to all cells in the sequence and the periodicity number and the steps numbers are assigned in an appropriate manner as discussed in (*iii*) in the foregoing.

After completing the second processing sequence, we again go to the cell sequence and take the next cell $Z = 2$. However, before we start a new processing sequence with it we have to make sure that it is a virgin cell by checking whether Gr $(2) = 0$ or not. If it is, we start the new sequence. If not, we know that the cell $Z = 2$ has already been processed and we go to the next cell $Z = 3$ and repeat the test procedure.

In Fig. 1 a flow chart is given for the program. One notes here that

the procedure is really not a computing algorithm because no computing is performed. It is basically an algorithm of sorting and unravelling. Therefore, it might be appropriate to call it an *"unravelling algorithm of global analysis."*

In the aforementioned discussion four arrays are used. In actual computation it is desirable to combine $P(Z)$ and $S(Z)$ by using the integral part and the decimal part of one number to represent each, respectively. This reduces the number of arrays required to three. Other programmatical improvements are also possible.

Multidimensional Problems. The implementing program just described is given in the context of a one-dimensional cell mapping. The application need not be so restricted. Consider, as an example, a two-dimensional cell mapping

$$Z_1(n+1) = C_1(Z_1(n), Z_2(n))$$
$$Z_2(n+1) = C_2(Z_1(n), Z_2(n)) \tag{9}$$

Again in practice our interest will be confined to a finite range of cells for each cell variable Z_1 or Z_2. Let the range of interest be covered by

$$Z_1^{(L)} \leqq Z_1 \leqq Z_1^{(U)}$$
$$Z_2^{(L)} \leqq Z_2 \leqq Z_2^{(U)} \tag{10}$$

Cells within these ranges will be called *regular cells*. The total number of the two-dimensional regular cells is

$$N_c = (Z_1^{(U)} - Z_1^{(L)} + 1)(Z_2^{(U)} - Z_2^{(L)} + 1) = N_{c1} \cdot N_{c2}. \tag{11}$$

All the two-dimensional cells outside the ranges of (10) will be lumped together as a sink cell. These $N_c + 1$ cells can then be relabelled as a one-dimensional sequence, $Z = 0, 1, 2, \ldots, N_c$, according to an appropriate relabeling procedure. Once this is done, the algorithm just described can be applied immediately.

Alternately, we can directly work with two-dimensional arrays. Let us assume that Z_1 and Z_2 cell coordinates have been standardized in such a way that the cells in the ranges of interest correspond to

$$Z_1 = 1, 2, \ldots, N_{c1}$$
$$Z_2 = 1, 2, \ldots, N_{2c} \tag{12}$$

These again will be called two-dimensional regular cells. Accompanying these will be the sink cell which will be assigned to be at (0, 0). The sink cell again will be a $P - 1$ cell. The cell mapping will map the sink cell to itself, and map a regular cell either to a regular cell or to the sink cell. To apply the algorithm we need two-dimensional arrays: $C_1(Z_1, Z_2)$, $C_2(Z_1, Z_2)$, Gr (Z_1, Z_2), $P(Z_1, Z_2)$, and $S(Z_1, Z_2)$. Apart from this difference, the basic unravelling procedure remains unchanged. We note here that according to the present scheme, for consistency, $C_1(Z_1, Z_2)$ and $C_2(Z_1, Z_2)$ must be either both nonzero or both zero. In the former case the image of (Z_1, Z_2) under the mapping is a regular cell while in the latter case the image is the sink cell.

4 First Example of Application

As a first example of application consider the following point mapping system.

$$x_1(n+1) = G_1(x_1(n), x_2(n)) = (1-\sigma)x_2(n) + (2 - 2\sigma + \sigma^2)[x_1(n)]^2$$
$$x_2(n+1) = G_2(x_1(n), x_2(n)) = -(1-\sigma)x_1(n) \tag{13}$$

This example is chosen because of the very simple nature of its nonlinearity and also because its global behavior is reasonably well understood. The mapping has a stable spiral point at (0, 0) and a saddle point at $(1, -(1-\sigma))$ when $0 < \sigma < 1$. We shall use the present method to find the domain of attraction of the spiral point at (0, 0).

First we find the cell mapping associated to (13) by using the procedure of discretization given in Section 6 of [6].

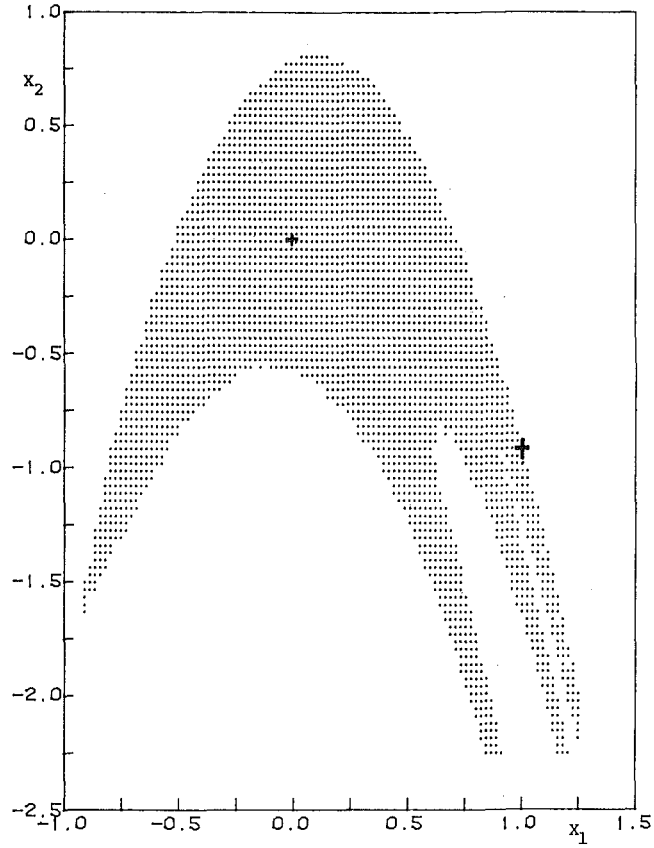


Fig. 2 Domain of attraction for the spiral point of (13) at (0,0) obtained by the unravelling algorithm; $\sigma = 0.1$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

$$Z_1(n+1) = C_1(Z_1(n), Z_2(n)) = \mathrm{Int}\left\{\frac{1}{h_1}[(1-\sigma)h_2 Z_2(n) + (2 - 2\sigma + \sigma^2)(h_1 Z_1(n))^2] + \tfrac{1}{2}\right\}$$
$$Z_2(n+1) = C_2(Z_1(n), Z_2(n)) = \mathrm{Int}\left\{\frac{1}{h_2}[-(1-\sigma)h_1 Z_1(n)] + \tfrac{1}{2}\right\} \tag{14}$$

Applying the algorithm to this mapping, we obtain the domain of attraction as shown in Fig. 2.[2] In the Z_1-direction 101 cells are used covering x_1 from −1.01125 to 1.26125 with an interval size $h_1 = 0.0225$. In the z_2-direction also 101 cells are used covering x_2 from −2.26625 to 1.01625. Thus the total number of regular cells is 10,201. In the figure a blank space (instead of a dot) means that the cell at that position is mapped eventually outside of the region shown and, therefore, is mapped into the sink cell and is lost. Here we note that while the system (13) is a very simple one, yet the domain of attraction has a rather complex shape.

The domain of attraction can of course also be obtained without using cell mappings. As described in [7] one can determine the domain of attraction by either (1) computing one branch of the separatrices from the saddle point located at $(1, -(1-\sigma))$, or (2) by mapping backward (hence expanding) from an assured but small sufficient region of asymptotical stability around the spiral point. As a matter

[2] The program used to produce Figs. 2–6 employs a different discretization procedure leading to a cell mapping which differs slightly from (14) but differs only in that the origin of the **x**-plane is not necessarily at the center of a **Z** cell.

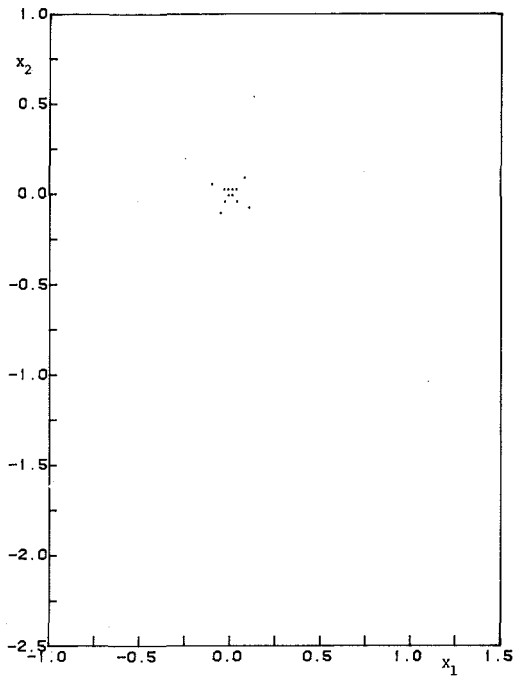


Fig. 3 A cluster of 12 periodic cells near (0,0); the $P-1$ cell at $(1, -.9)$ corresponding to the saddle point of (13); $\sigma = 0.1$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

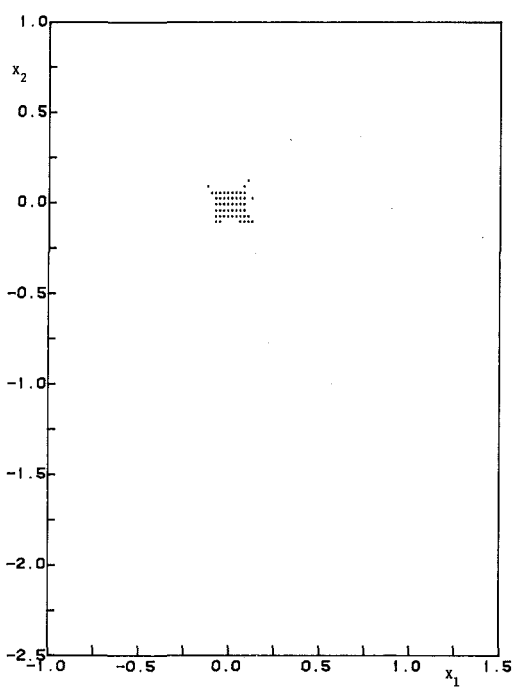


Fig. 4 5-step domain of attraction for the periodic cells near (0,0); $\sigma = 0.1$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

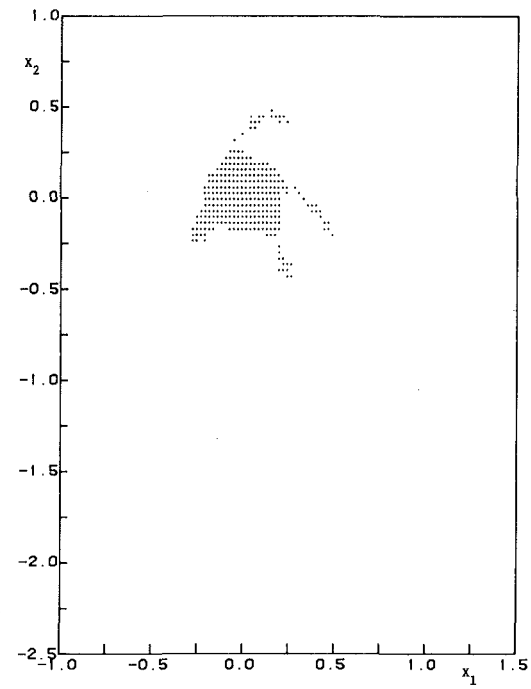


Fig. 5 15-step domain of attraction for the periodic cells near (0,0); $\sigma = 0.1$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

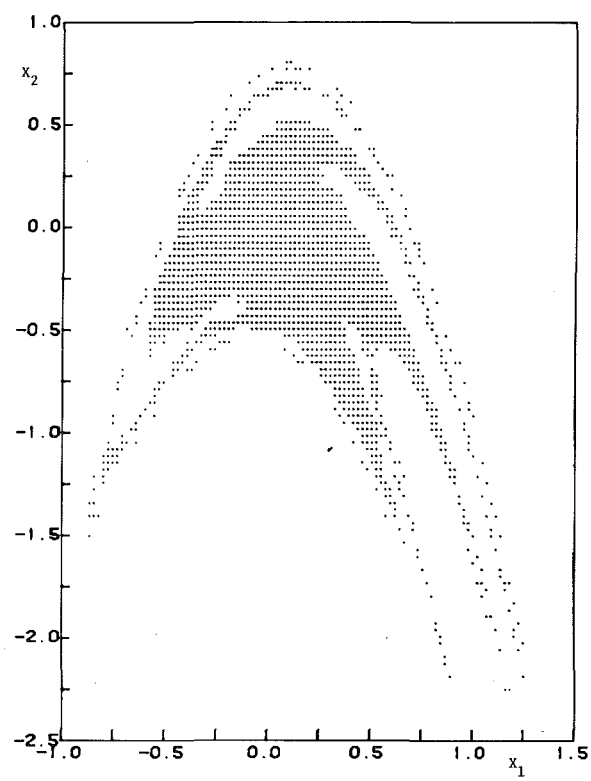


Fig. 6 25-step domain of attraction for the periodic cells near (0,0); $\sigma = 0.1$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

of fact, such calculations were performed and reported in [7]. Comparing Fig. 2 given here with Fig. 3 of [7] where σ is also taken to be 0.1, one sees that the present cell mapping algorithm of global analysis duplicates excellently the domain obtained by using point mapping techniques. Yet, the computer time required for the present method is easily one or two orders of magnitude better than that required by the point mapping techniques. In general, the two point mapping techniques discussed in [7] for determining the domain of attraction demand very delicate computer programming. The new unravelling algorithm is, however, robust and is very simple to implement.

Moreover, by using this algorithm we obtain in one computer run not only the domain of attraction but also $S(Z)$, the number of steps a cell is away from the periodic cells of the domain of attraction. Using

that information we can plot the r-step domains of attraction for various values of r. For Fig. 3, $r = 0$. There are 12 cells in a cluster[3] here, consisting of 3 sets of $P - 4$ motions. These 12 cells form a cluster of periodic cells, replacing the single spiral point of the point mapping. There is also an isolated $P - 1$ cell here corresponding to the saddle point. Figs. 4–6 show the 5-step, 15-step, and 25-step domain of attraction. These diagrams indicate how long it takes "timewise" for a cell to be attracted to the cluster of periodic cells. It might be added here that in Fig. 2 where the total domain of attraction is shown, the largest number of steps any cell is away from the cluster is 43.

5 Second Example of Application

As a second example we consider a nonlinear problem of a hinged bar under the action of a periodic impact load. The problem was treated in [3]. The system is a simple one but because of strong nonlinearity its global behavior is very complicated. The problem does have an *exact* point mapping which governs the dynamic behavior completely. The point mapping for a hinged bar which is damped but otherwise elastically unrestrained is given by

$$x_1(n+1) = x_1(n) - \frac{1-e^{-2\mu}}{2\mu}\alpha \sin x_1(n) + \frac{1-e^{-2\mu}}{2\mu} x_2(n) = G_1(\mathbf{x}(n))$$

$$x_2(n+1) = -e^{-2\mu}\alpha \sin x_1(n) + e^{-2\mu}x_2(n) = G_2(\mathbf{x}(n)) \qquad (15)$$

where x_1 and x_2 are, respectively, the angular displacement and angular velocity of the bar, μ is the damping coefficient, and α is the impact load parameter. For the derivation of the equations and for the general nature of the global behavior, the reader is referred to [3]. Here we apply the present algorithm to this problem and show the results for a typical case.

We take $\mu = 0.1\pi$ and $\alpha = 5.5$. Then from Fig. 3 of [3] it is known that the origin of the phase plane is an unstable equilibrium point and the system has a stable $P - 2$ motion at

$$x_1^*(1) = 1.27280, \quad x_2^*(1) = 1.82907 \qquad (16a)$$

$$x_1^*(2) = -1.27280, \quad x_2^*(2) = -1.82907. \qquad (16b)$$

Besides these periodic solutions discussed in [3], the system has also two "advancing" type of periodic solutions as follows: A point at

$$x_1^* = 0.80063, \quad x_2^* = -4.51463 \qquad (17)$$

is mapped in one step into

$$x_1^* = 0.80063 - 2\pi, \quad x_2^* = -4.51463. \qquad (18)$$

Physically, since the bar is elastically unrestrained, (17) and (18) represent the same state of displacement and velocity for the hinged bar, although going from (17) to (18) the bar has *revolved* once in the negative direction. Regarding (17) and (18) as the same point in the phase plane or, equivalently, taking x_1 by modulo 2π, the point (17) is a $P - 1$ point. We call this an "*advancing-type*" $P - 1$ *point*. Similarly, a point at

$$x_1^* = -0.80063, \quad x_2^* = 4.51463 \qquad (19)$$

is mapped in one step into

$$x_1^* = -0.80063 + 2\pi, \quad x_2^* = 4.51463 \qquad (20)$$

and, therefore, it is also an advancing-type $P - 1$ point. It can be easily shown that these two advancing-type $P - 1$ points are asymptotically stable. Thus we have four asymptotically stable periodic points in (16a), (16b), (17), and (19) and for each periodic solution there is its domain of attraction.

Using the procedure of discretization as given in Section 6 of [6], we can easily construct the cell mapping **C** corresponding to (15). Once the cell mapping **C** has been constructed we can apply the unravelling algorithm to determine the global behavior of the system. The results are shown in Figs. 7–11. In applying the algorithm to generate these particular figures, we have used $\mathbf{G}^2$, instead of simply **G**, to determine the cell mapping. This allows us to obtain the domains of attraction for (16a) and (16b) separately. For this reason the number of steps referred to for Figs. 7–11 should be interpreted accordingly.

For Figs. 7–11, 10,201 regular cells are used covering $-3.14128 \leqq x_1 < 3.14128$ and $-5.07681 \leqq x_2 < 5.07681$ with the cell sizes $h_1 = 0.062204$ and $h_2 = 0.100531$. The part of the phase plane outside these ranges with x_1 modulo 2π is represented by the sink cell. Fig. 7 shows the periodic cells, or the 0-step domains of attraction. Here we find that the unstable $P - 1$ point of (15) at the origin is replaced by an isolated $P - 1$ cell at the origin. In the figure it is designated by the symbol "s." Each of the two $P - 2$ points of (15), i.e., (16a) and (16b),

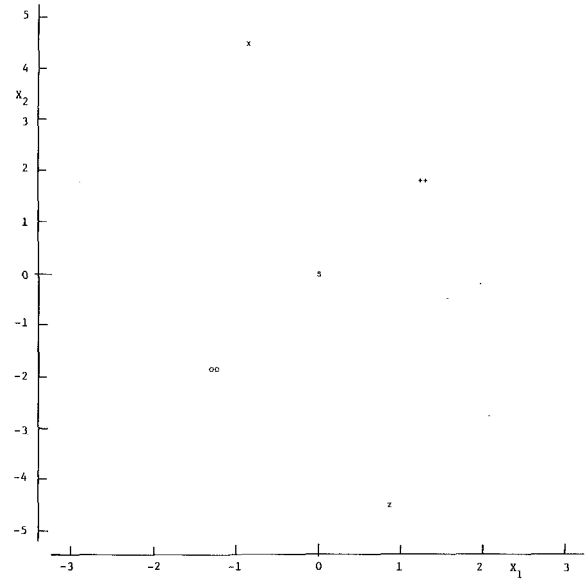


Fig. 7 Periodic cells for (15); $\mu = 0.1\pi$ and $\alpha = 5.5$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

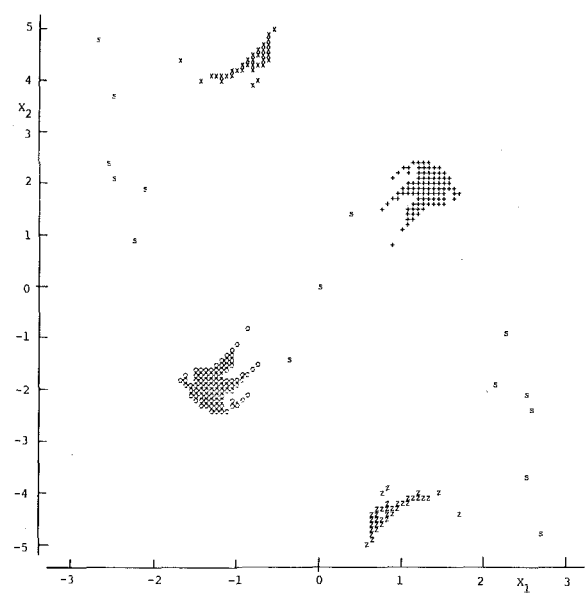


Fig. 8 3-step domains of attraction; $\mu = 0.1\pi$ and $\alpha = 5.5$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

[3] The reader is referred to [6] for the meaning of a cluster of periodic cells.

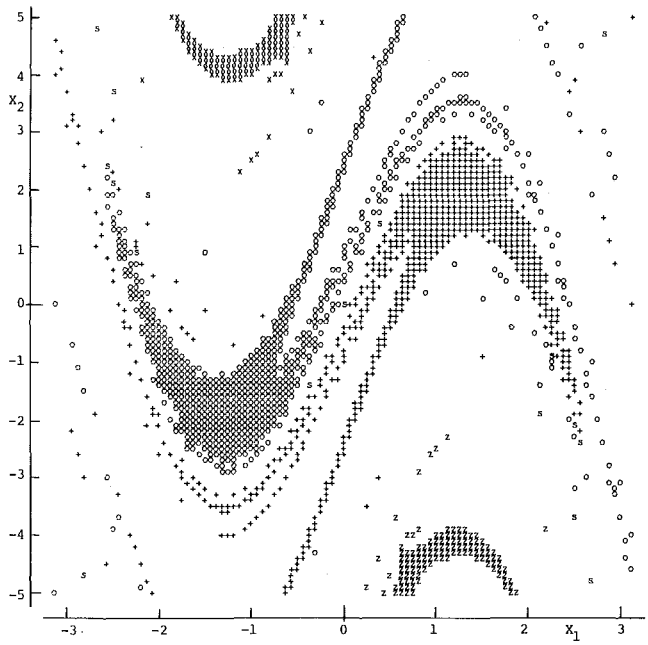


Fig. 9 5-step domains of attraction; $\mu = 0.1\pi$ and $\alpha = 5.5$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

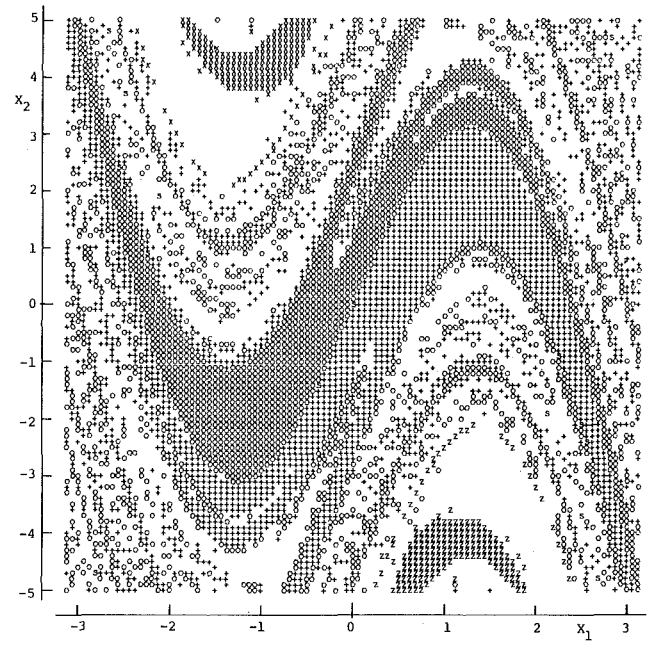


Fig. 10 8-step domains of attraction; $\mu = 0.1\pi$ and $\alpha = 5.5$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

is replaced by a core of 2 periodic cells, with the cells in one core designated by the symbol "+" and the other by "0." The advancing-type $P - 1$ points of (15) at (17) and (19) are replaced by two isolated advancing-type $P - 1$ cells. They are designated by the symbols "x" and "z," respectively.

Fig. 8 shows the 3-step domains of attraction for these four groups of attracting cells. In each domain the attracted cells share the same symbol with the core cell or cells. In Fig. 8 one also finds a few cells which are three steps or less removed from the cell at the origin. They correspond to points on a pair of separatrices approaching the unstable $P - 1$ point at the origin for the point mapping system (15). Figs. 9 and 10 give the 5-step and 8-step domains of attraction. They begin to exhibit the complicated interweaving pattern of the global behavior.

Fig. 11 shows the total domains of attraction obtained with this application of the algorithm. Of the total 10,201 regular cells, 4324 cells belong to (16a), 4324 belong to (16b), 223 cells belong to (17), and 223 belong to (19). 17 cells belogn to the $P - 1$ cell at the origin on account of being on the separatrices. 1090 cells are mapped into the sink cell and their eventual fate is not determined by this run of unravelling. Fig. 11 delineate very nicely the distribution of these four domains of attraction. Of course, the domains of attraction for (16a) and (16b) can be combined into one domain of attraction for the stable $P - 2$ motion which has two cores with two periodic cells in each core. Again, we wish to remark that the algorithm is extremely simple to apply and that the computer run to generate the data for Figs. 7–11 was made on a minicomputer PDP-11/60 consuming very little computer time. The precise time was difficult to determine because of the time sharing aspect of the usage of the computer. It was estimated to be of the order of a few second.

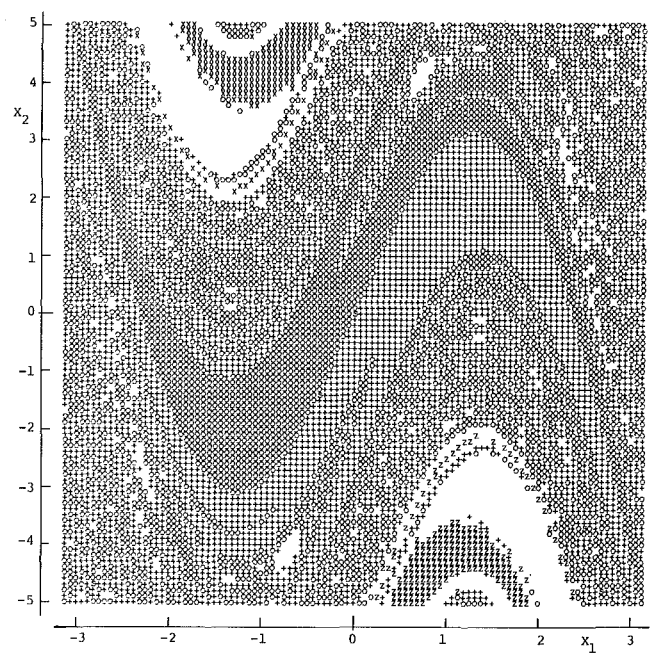


Fig. 11 The total domains of attraction; $\mu = 0.1\pi$ and $\alpha = 5.5$; $N_{c1} = N_{c2} = 101$; the total number of regular cells 10,201

6 Application to Nonlinear Differential Dynamical Systems

In the last two sections we have demonstrated the utility of the method by carrying out the global analysis on cell mapping systems which are obtained by discretizing point mappings. The application of this algorithm is, however, not restricted just to cases of this kind. It can also be applied to nonlinear dynamical systems governed by ordinary differential equations. Consider first the autonomous systems governed by

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t)) \tag{21}$$

We construct a cell mapping for (21) by doing two things. First, we divide the state space into cells according to the procedure described in Section 2 of [6]. A cell $\mathbf{z} = (Z_1, Z_2, \ldots, Z_N)$ is defined to contain all points $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ which satisfy

$$(Z_i - \tfrac{1}{2})h_i \leqq x_i < (Z_i + \tfrac{1}{2})h_i \tag{22}$$

where h_i is the cell width in the x_i-direction. To get the mapping we integrate (21) over a small time interval τ by any convenient and ac-

curate numerical scheme. The cell mapping is therefore obtained in following manner.

(*i*) For a given cell $\mathbf{Z}(n) = (Z_1(n), Z_2(n), \ldots, Z_N(n))$ we compute the x_i coordinates of the center point of the cell and denote them as $x_i^{(d)}(0)$

$$x_i^{(d)}(0) = h_i Z_i(n). \tag{23}$$

This point $\mathbf{x}^{(d)}(0)$ is then taken to be the starting value of the state vector $\mathbf{x}$.

(*ii*) The evolution of the system by (21) taking $\mathbf{x}^{(d)}(0)$ at $t = 0$ to $\mathbf{x}^{(d)}(\tau)$ at $t = \tau$ is obtained by numerical integration. The size of τ will be usually taken to be small but it needs not be the basic time step interval Δt of integration. The vector $\mathbf{x}^{(d)}(\tau)$ could be the result after carrying out several steps of integration.

(*iii*) The construction of the cell mapping is completed when one identifies $\mathbf{x}^{(d)}(\tau)$ by the cell in which it lies according to (22). This cell is denoted by $\mathbf{Z}(n+1)$ and is taken to be the image of $\mathbf{Z}(n)$.

$$Z_i(n+1) = C_i(\mathbf{Z}(n)) = \mathrm{Int}\left\{\frac{1}{h_i} x_i^{(d)}(\tau) + \tfrac{1}{2}\right\}. \tag{24}$$

We note here that once $\mathbf{C}$ has been determined in this manner, we have a cell map which governs the global behavior of the system.

If the nonlinear system is not autonomous but nevertheless is explicitly periodic in t

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t), \tag{25}$$

the aforementioned procedure of constructing the associated cell mapping still applies except that in step (*ii*) it will be necessary to take τ, the interval of numerical integration, to be the period of the system.

Once the associated cell mapping to a nonlinear differential dynamical system has been constructed, the unravelling algorithm can then be used to determine the global behavior of the system. Of course, it is necessary to define the regular cells which are the cells in the ranges of our interest and to introduce a sink cell to take care of all the other cells.

As an example, let us consider the classic problem of van der Pol.

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = \mu(1 - x_1^2)x_2 - x_1 \tag{26}$$

Since the example does demonstrate, we believe, a novel approach to nonlinear problems, we shall describe the steps of analysis, not just the results, so that a reader can easily duplicate the analysis if he so wishes. Let R_1 and R_2 be the ranges of interest in the x_1 and x_2-directions. Let the total interval numbers used in the two directions be N_{c1} and N_{c2}, i.e.,

$$R_1 = h_1 N_{c1}, \quad R_2 = h_2 N_{c2}. \tag{27}$$

Consider a regular cell $\mathbf{Z}(n) = (Z_1(n), Z_2(n))$. Its center position $\mathbf{x}^{(d)}(0)$ can be determined by (23). For integration let us use the 4th-order Runge-Kutta method. Let Δt be the single time step of integration. We integrate (26) a number of steps, say J-steps, to obtain $\mathbf{x}^{(d)}(\tau)$ where $\tau = (\Delta t)J$. The image cell $\mathbf{Z}(n+1)$ of $\mathbf{Z}(n)$ is then determined by (24). If $\mathbf{C}(\mathbf{Z}(n))$ is not a regular cell then $\mathbf{Z}(n+1)$ is assigned to the sink cell.

The two-dimensional cell array can be converted to a one-dimensional array. Let the two-dimensional ranges be

$$Z_1^{(L)} \leqq Z_1 \leqq Z_1^{(L)} + N_{c1} - 1,$$
$$Z_2^{(L)} \leqq Z_2 \leqq Z_2^{(L)} + N_{c2} - 1. \tag{28}$$

The converted one-dimensional array may be defined by

$$Z = (Z_2 - Z_2^{(L)})N_{c1} + Z_1 - Z_1^{(L)} + 1 \tag{29}$$

with $Z = 0$ assigned to be the sink cell.

The global behavior of the van der Pol equation is well understood. There is a stable limit cycle toward which all the trajectories approach.

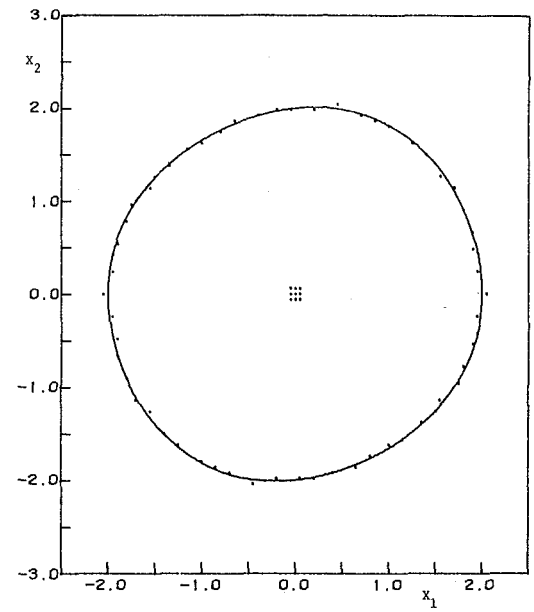


Fig. 12 The case $\mu = 0.1$; the curve represents the limit cycle of the van der Pol equation; dots represent periodic cells obtained by the present algorithm of global analysis; $N_{c1} = N_{c2} = 101$; the total number of regular cells is 10,201

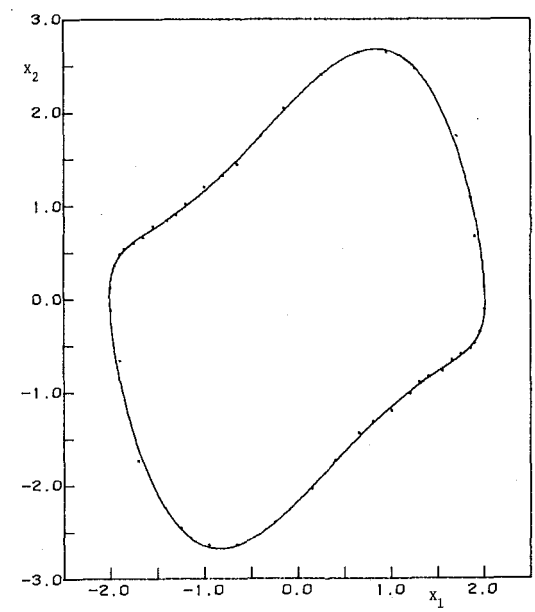


Fig. 13 The case $\mu = 1.0$; the curve represents the limit cycle of the van der Pol equation; dots represent periodic cells obtained by the present algorithm of global analysis; $N_{c1} = N_{c2} = 101$; the total number of regular cells is 10,201

Therefore, the domain of attraction of the limit cycle is the whole phase plane except the origin. The limit cycle of (26) should correspond to a collection of periodic cells when the global analysis algorithm is applied to its associated cell mapping systems. Figs. 12–13 show the results of such an application. Fig. 12 is for $\mu = 0.1$ and Fig. 13 is for $\mu = 1.0$. In both cases the number of regular cells used is 10,201 with $N_{c1} = N_{c2} = 101$. The cell size is determined by $h_1 = 0.05$ and $h_2 = 0.06$; thus the regular cells cover a region $-2.525 \leqq x_1 < 2.525$ and $-3.03 \leqq x_2 < 3.03$. The 4th-order Runge-Kutta numerical procedure used for these cases employs a single integration time step $\Delta t = 0.05$ and 26 steps of integration ($J = 26$) are used to create the cell mapping, i.e., $\tau = (\Delta t)J = 1.30$. The dots in the figures signify periodic cells. In Fig. 12, for the case $\mu = 0.1$, we find one group of 58 periodic

cells forming a "circle" replacing the limit cycle and another group of 9 periodic cells forming a core replacing the unstable equilibrium point at the origin. In Fig. 13, for the case $\mu = 1.0$, we find a group of 46 periodic cells forming a circle replacing the limit cycle and a core of just a single periodic cell replacing the unstable equilibrium point at the origin. In the figures we have also shown the limit cycles themselves. It is remarkable and satisfying to see the present method of global analysis to "duplicate" accurately the limit cycles for both cases of widely different values of μ.

This example is used here merely to show the potential of using this unravelling algorithm to deal with dynamical systems governed by ordinary differential equations. A great deal of theoretical work remains to be done in order to find the best methods of implementation for various kinds of problems in this category.

7 Remarks

From the examples it can be seen that the method of using cell mappings and the global analysis algorithm is a very efficient one. We believe that the efficiency results from three factors.

1 A very efficient algorithm for the desired sorting of an array.
2 The advantage of discretization.
3 The method only requires integration over a time interval once for all for each cell to obtain the global mapping. Other methods require either long term integration repeatedly even for points in the same cell or long term repeated mapping for points in the same cell, resulting in a tremendous amount of duplication of effort.

Of these three factors, it is our belief that factor (3) is the most important one.

In [6] the discretization of linear point maps is discussed in order to provide a qualitative appreciation of the consequences of such a process. In this paper the applications offered are to nonlinear point mappings. One may ask in what way the discretization of a linear point mapping is representative for the discretization of a nonlinear point mapping. In general, when the cells are small, then in the immediate neighborhood of a P-K point the discretization of a nonlinear point map is expected to be essentially that of its locally linearized map. A complete quantitative study of this question is, however, yet to be made.

Finally, it might also be appropriate to remark here that the cell mapping and the new method of analysis can be viewed in another context. For dynamical problems (lumped or continuous systems) there are three kinds of variables: the independent time variable, the independent spatial variables, and the dependent state variables. The classical stepwise time integration is a procedure to discretize the independent time variable. The finite-element method is a procedure to discretize the independent spatial variables. In this spirit the cell mapping and the present method of analysis is a procedure to discretize the dependent state variables.

Acknowledgments

The results presented here were obtained in the course of research supported by a grant from the National Science Foundation.

References

1 Bernussou, J., *Point Mapping Stability,* Pergamon Press, New York, 1977.
2 Marsden, J. E., and McCracken, M, *The Hopf Bifurcation and Its Applications,* Springer-Verlag, New York, 1976.
3 Hsu, C. S., Yee, H. C., and Cheng, W. H., "Steady-State Response of A Nonlinear System Under Impulsive Periodic Parametric Excitation," *Journal of Sound and Vibration,* Vol. 50, 1977, pp. 95–116.
4 Hsu, C. S., "On Nonlinear Parametric Excitation Problems," *Advances in Applied Mechanics,* Vol. 17, 1977, pp. 245–301.
5 Hsu, C. S., "Nonlinear Behavior of Multibody Systems Under Impulsive Parametric Excitation," *Dynamics of Multibody Systems,* ed., Magnus, K., Springer-Verlag, *Proceedings of the IUTAM Symposium on Dynamics of Multibody Systems,* held in Munich, Germany, August 29–September 3, 1977, pp. 63–74.
6 Hsu, C. S., "A Theory of Cell-To-Cell Mapping Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 931–939.
7 Hsu, C. S., Yee, H. C., and Cheng, H. H., "Determination of Global Regions of Asymptotic Stability for Difference Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 147–153.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper which 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.

The Rigid-Wavy-Wall Assumption in Interface Stability Problems[1]

P. B. Joshi[2] and J. A. Schetz[3]

Introduction

Hydrodynamic stability of a gas-liquid interface has been studied extensively [1–12] due to its applications in a variety of phenomena, e.g., the wind generation of ocean waves and liquid film cooling of reentry bodies. One of the key assumptions in the analysis [5–9] is that the interface behaves like a rigid wavy wall relative to the gas flow. Although the rigid-wavy-wall assumption may be appropriate for inviscid high-speed gas [6, 7], its validity is difficult to ascertain for a viscous, high-speed or low-speed gas [5, 8, 9]. The purpose of the present investigation was to directly evaluate the effects of relaxing the rigid-wavy-wall assumption on interface stability. Therefore, the simple case of an incompressible, viscous, laminar flow of a gas over a viscous, laminar, liquid layer was considered for detailed study as a representative example. Both fluids were assumed to be parallel with linear, steady-state velocity profiles. This assumption is felt to be justified in this exploratory study by the profound mathematical simplifications that result. Also

1 It is the simplest, physically possible viscous profile.
2 The Orr-Sommerfeld equation has an exact solution.
3 The linear stability of a plane Couette flow has been studied in great detail, and it has been found to be unconditionally stable.

Consequently, attention is focused exclusively on the stability of the interface during the present work.

[1] This work was supported by the Air Force Office of Scientific Research. Dr. B. T. Wolfson was the Technical Monitor.

[2] Graduate Assistant, now at Tetra Tech, Inc., Jacksonville, Fla. Assoc. Mem. ASME.

[3] Professor and Department Head, Aerospace and Ocean Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061. Fellow ASME.

Manuscript received by ASME Applied Mechanics Division, September, 1978; final revisions, April, 1980.

Problem Formulation

The steady-state mean flow velocity profiles in the liquid (fluid "1") and the gas (fluid "2") have the following dimensionless form:

$$\hat{u}_1(\xi) = 1 + \xi; -1 \le \xi \le 0 \tag{1}$$

$$\hat{u}_2(\eta) = \frac{\eta + \epsilon\bar{\mu}}{1 + \epsilon\bar{\mu}}; 0 \le \eta \le 1 \tag{2}$$

where $\xi = y/h$ and $\eta = y/\delta$. Also, h and δ represent the liquid depth and the gas layer thickness, respectively, and y is measured normal to the interface. The interface velocity u_{if} at $\xi = \eta = 0$ made nondimensional relative to u_e, is given by

$$\bar{u} = \epsilon\bar{\mu}/(1 + \epsilon\bar{\mu}) \tag{3}$$

In equations (1)–(3), $\epsilon = h/\delta$, μ is the viscosity with $\bar{\mu} = \mu_2/\mu_1$.

When the steady-state configuration is distributed, the resulting unsteady, two-dimensional, incompressible motion is governed by the Navier-Stokes equations. Assuming an infinitesimal travelling wave disturbance on the interface to be of the form

$$\eta(x,t) = h \exp\,[ik(x - ct)] \tag{4}$$

where k is the disturbance wave number and c denotes phase velocity, the x-axis coincides with the interface, and t represents time, the corresponding disturbances in the other flow variables (such as velocity components) are given by

$$\bar{q}_j(x,y,t) = q_j(y) \exp\,[ik(x - ct)], \quad j = 1,2 \tag{5}$$

and the stability problem reduces to the solution of Orr-Sommerfeld equations for gas and liquid,

$$\psi_1^{\text{iv}} - 2\alpha_1^2\psi_1'' + \alpha_1^4\psi_1 = i\alpha_1 R_1(\psi_1'' - \alpha_1^2\psi_1)\,(\hat{u}_1 - c_1) \tag{6}$$

$$\ddddot{\psi}_2 - 2\alpha_2^2\ddot{\psi}_2 + \alpha_2^4\psi_2 = i\alpha_2 R_2(\ddot{\psi}_2 - \alpha_2^2\psi_2)\,(\hat{u}_2 - c_2) \tag{7}$$

The governing equations (6) and (7) are subject to the boundary conditions

$$\psi_1(\xi) = 0, \quad \psi_1'(\xi) = 0, \quad \xi = -1 \tag{8, 9}$$

$$\psi_2(\eta) = 0, \quad \dot{\psi}_2(\eta) = 0, \quad \eta = 1 \tag{10, 11}$$

$$\bar{u}[\dddot{\psi}_1'(\xi) - i\alpha_1\hat{u}_1'] = \epsilon[\dot{\psi}_2(\eta) - i\alpha_1\dot{\hat{u}}_2] \quad \text{at} \quad \xi = 0, \quad \eta = 0 \tag{12}$$

$$\bar{u}[\psi_1''(\xi) + \alpha_1^2\psi_1(\xi)] = \bar{\mu}\epsilon^2[\ddot{\psi}_2(\eta) + \alpha_2^2\psi_2(\eta)] \quad \text{at} \quad \xi = \eta = 0 \tag{13}$$

$$\frac{1}{\alpha_2{}^2R_2}\{\dddot{\psi}_2(\eta) - \alpha^2\dot{\psi}_2(\eta)\} - \frac{\bar{u}^2}{\bar{\rho}}\frac{1}{\alpha_1{}^2R_1}\{\psi_1'''(\xi) - \alpha_1{}^2\psi_1'(\xi)\}$$
$$+ \frac{i}{\alpha_2}\{\dot{\hat{u}}_2\psi_2(\eta) - (\hat{u}_2(\eta) - c_2)\dot{\psi}_2(\eta)\}$$
$$- \frac{u^2}{\bar{\rho}}\frac{i}{\alpha_1}\{\hat{u}_1'\psi_1(\xi) - (\hat{u}_1(\xi) - c_1)\psi_1'(\xi)\}$$
$$- \frac{2}{\epsilon\bar{\mu}R_2}\{\epsilon\bar{\mu}\dot{\psi}_2(\eta) - \bar{\mu}\psi_1'(\xi)\}$$
$$= -\frac{\bar{\mu}^2}{\bar{\rho}}\left(\alpha_1{}^2W^2 + \frac{1}{F^2}\right) \quad \text{at} \quad \xi = \eta = 0 \quad (14)$$

$$\psi_1(\xi) - i\alpha_1(\hat{u}_1(\xi) - c_1) = 0 \quad \text{at} \quad \xi = 0 \quad (15)$$

$$\psi_2(\eta) - i\alpha_1(\hat{u}_2(\eta) - c_2) = 0 \quad \text{at} \quad \eta = 0 \quad (16)$$

In equations (6)–(16), ψ_1 and ψ_2 denote amplitudes of the disturbance stream function, $\alpha_1 = kh$ and $\alpha_2 = k\delta$ are dimensionless wave numbers with $\alpha_1 = \epsilon\alpha_2$, $R_1 = u_{if}h/\nu_1$ and $R_2 = u_e\delta/\nu_2$ denote liquid and gas Reynolds numbers respectively, and $c_2 = \bar{u}c_1$ is the dimensionless phase speed in the gas. Further, ν_1 and ν_2 stand for kinematic viscosities of the liquid and gas, respectively. Boundary conditions in equations (8)–(11) express the fact that the disturbance velocity components vanish at $\xi = -1$ and $\eta = 1$. Equations (12) and (13) represent the balance of tangential velocity and shear stress, respectively, at the interface. The difference in normal stresses at the interface equals the surface tension as shown by equation (14). Equations (15) and (16) express the kinematics boundary conditions at the interface for the liquid and gas, respectively. It is noted that the boundary conditions in equations (12), (14)–(16) are nonhomogeneous and the rest are homogeneous. This enables one to convert the problem of interface stability to an eigenvalue problem.

Solution of Stability Problem

Equations (6)–(16) represent the mathematical statement of the stability problem. It should be pointed out that the phase speed c_2 has been retained in the gas disturbance terms, and thus the rigid-wavy-wall assumption has not been made. Had such as assumption been invoked, c_2 would be set identically equal to zero. Yih [10] has considered a somewhat similar problem, but he limited the solution to small wave number disturbances for which $\alpha_1R_1 \ll 1$. The general solutions of equations (6) and (7) for arbitrary wave numbers and Reynolds number are [14],

$$\psi_1(\xi) = C_1 \exp(\alpha_1\xi)$$
$$+ C_2 \exp(-\alpha_1\xi) + \frac{C_3}{\alpha_1}\int_{\xi^*}^{\xi} \sinh\{\alpha_1(\xi - \tilde{t})\}\, Ai\,\{\zeta_1(\tilde{t})E^+\}\, d\tilde{t}$$
$$+ \frac{C_4}{\alpha_1}\int_{\xi^*}^{\xi} \sinh\{\alpha_1(\xi - \tilde{t})\}\, Ai\,\{\zeta_1(\tilde{t})E^-\} d\tilde{t} \quad (17)$$

$$\psi_2(\eta) = C_5 \exp(\alpha_2\eta) + C_6 \exp(-\alpha_2\eta) + \frac{C_7}{\alpha_2}\int_{\eta^*}^{\eta} \sinh\{\alpha_2(\eta - \tilde{t})\}\, Ai$$
$$\{\zeta_2(\tilde{t})E^+\} d\tilde{t} + \frac{C_8}{\alpha_2}\int_{\eta^*}^{\eta} \sinh\{\alpha_2(\eta - \tilde{t})\}\, Ai\{\zeta_2(\tilde{t})E^-\} d\tilde{t} \quad (18)$$

where $\zeta_1(\xi) = -(\alpha_1R_1\hat{u}_1')^{-2/3}\{\alpha_1{}^2 + i\alpha_1R_1(\hat{u}_1(\xi) - c_1)\}$ with ξ^* such that $\zeta_1(\xi^*) = 0$. Similarly, $\zeta_2(\eta) = -(\alpha_2R_2\dot{\hat{u}}_2)^{-2/3}\{\alpha_2{}^2 + i\alpha_2R_2(\hat{u}_2(\eta) - c_2)\}$ with $\zeta_2(\eta^*) = 0$. In equations (17) and (18), C_1–C_8 are the constants of integration, Ai denotes Airy function of the first kind, and $E^{\pm} = \exp(\pm 2\pi i/3)$. Introducing the solutions (17) and (18) into the boundary conditions (8)–(15), the result is a set of eight linear algebraic equations with the constants C_1–C_8 as unknowns. When these constants are evaluated and substituted in equation (16), the result is a frequency equation for c_1 (or c_2) which can be solved for a given set of flow parameters (ϵ, $\bar{\mu}$, R_1, etc.) and a given disturbance wave number α_1 (or α_2). In this sense, the stability problem reduces to an eigenvalue problem. The eigenvalue, c_1, was determined via a combination of graphical and Newton-Raphson iteration procedures which are described in detail in [14]. In general, c_1 is complex, and if $c_1 = c_{1r} + ic_{1i}$, equation (4) shows that $c_{1i} > 0$ corresponds to instability, $c_{1i} < 0$ corresponds to stability, and $C_{1i} = 0$ implies neutral stability.

Table 1 Parameters for numerical example gas/liquid system: air/water

Liquid/gas layer thickness, ϵ	4.302×10^{-2}
Gas density/liquid density, $\bar{\rho}$	1.208×10^{-3}
Gas viscosity/liquid viscosity, $\bar{\mu}$	0.263
Gay Reynolds number, R_2	338.20
Liquid Reynolds number, R_1	35.54
Weber number, W	5.46
Froude number, F	0.93

Results

Description of Conditions for Numerical Computations. The experimental conditions of Craik [5] were employed as a typical case for the computation of eigenvalues. Since the analytical model does not match the conditions of the experiment perfectly, some of the parameters in Craik's investigations were suitably adjusted (see reference [14] for complete details). Table 1 contains the list of various nondimensional parameters employed in the numerical computations.

Examination of the Rigid-Wavy-Wall Assumption. Eigenvalues were computed for the conditions listed in Table 1 and various stability modes were identified. The details are contained in [14], and only the most significant mode, termed the modified Kelvin-Helmholtz mode because its speed of propagation is nearly the same as the classical Kelvin-Helmholtz mode, is discussed here. The amplification curve in Fig. 1(a) shows that this mode is stable for small values of α_1 and then becomes unstable at large α_1. Also included in the figure is the amplification curve obtained upon invoking the rigid-wavy-wall assumption (i.e., $c_2 = \bar{u}c_1 \approx 0$). A comparison of the two sets of curves reveals the following significant results:

1 The assumption of a rigid-wavy interface relative to the gas has no effect at low disturbance wave numbers and affects the neutrally stable wave number only slightly.

2 For wave numbers greater than the neutrally stable value, the rigid-wavy-wall assumption results in an underestimation of the disturbance amplification rate.

3 Most importantly, for sufficiently large wave numbers ($\alpha_1 > 0.4$ in the present case), the rigid-wavy-wall assumption results in the prediction of stability when the interface is actually unstable.

Fig. 1(b) shows that there is no appreciable difference between the phase speed curves for the case $c_2 = 0$ and $c_2 \neq 0$ and that the phase speed is reasonably constant for all wave numbers (at least up to $\alpha_1 = 0.4$). The average dimensional value of the phase speed is 16.1 cm/sec which may be compared with Craink's [5] experimental value of 11.9 cm/sec at the onset of instability. This comparison must be viewed with caution, however, because computations have not been carried out to larger wave numbers at which maximum growth rate occurs. It is believed that the modified Kelvin-Helmholtz mode in Figs. 1(a) and 1(b) is the same as the one obtained by Bordner, et al. [8], using the experimental data of reference [12]. It also appears that the mode can be associated with the class C stability of Landahl [15].

Conclusions

The validity of the rigid-wavy-wall assumption in problems of interface stability has been investigated. As a starting step, an incompressible, viscous, laminar gas adjacent to a liquid layer, with both fluids having linear velocity profiles, has been considered. The frequent assumption of a rigid-wavy interface relative to the gas motion was found valid only for very small values of the disturbance wave number ($\alpha \ll 1$). When the wave number is moderate ($\alpha = 0(1)$), such an assumption not only results in a gross underestimation of the amplification rate but can even erroneously predict stability for an unstable interface. This conclusion is important, and further research

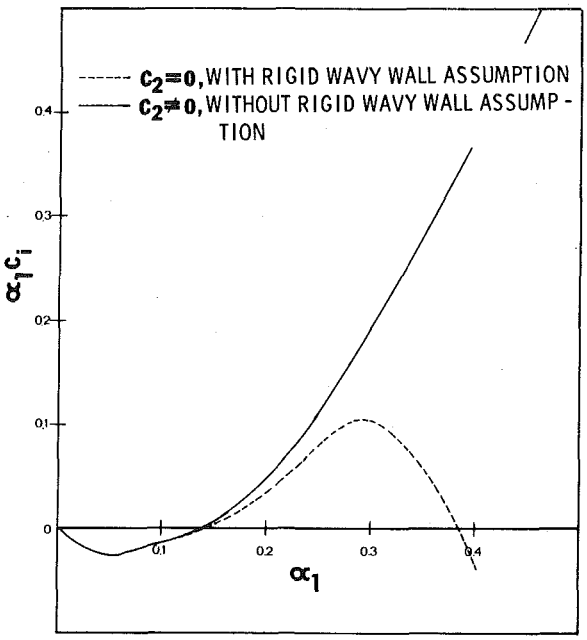


Fig. 1(*a*) Comparison of amplification curves with and without rigid wavy wall assumption

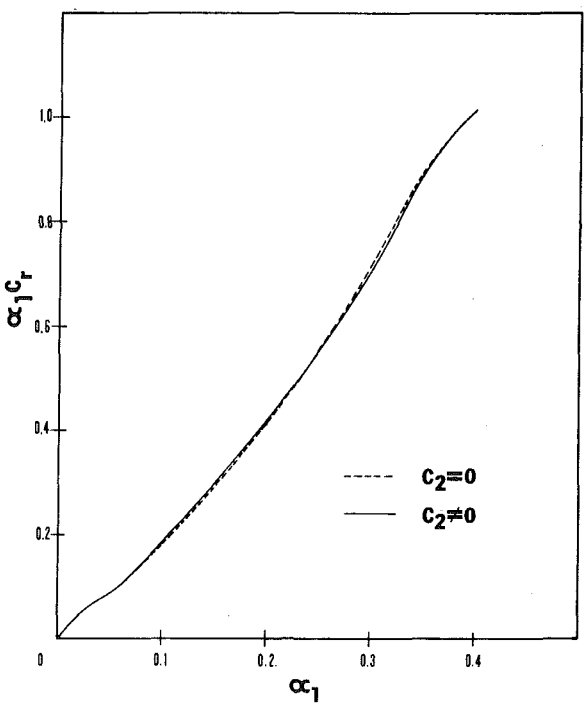


Fig. 1(*b*) Comparison of phase velocity curves with and without rigid wavy wall assumption

is necessary to study the effects of curvature of gas velocity profile and compressibility.

References

1 Lock, R. C., "Hydrodynamic Stability of the Flow in the Laminar Boundary Layer Between Parallel Streams," *Proceedings of the Cambridge Philosophical Society,* Vol. 50, 1954, pp. 105–124.

2 Feldman, S., "On the Hydrodynamic Stability of Two Viscous Incompressible Fluids in Parallel Uniform Shearing Motion," *Journal of Fluid Mechanics,* Vol. 2, 1957, pp. 343–370.

3 Miles, J. W., "The Hydrodynamic Stability of a Thin Film of Liquid in Uniform Shearing Motion," *Journal of Fluid Mechanics,* Vol. 8, 1960, pp. 593–610.

4 Cohen, L. S., and Hanratty, T. J., "Generation of Waves in the Concurrent Flow of Air and a Liquid," *AIChE Journal,* Vol. 31, 1965, pp. 138–144.

5 Craik, A. D. D., "Wind Generated Waves in Thin Liquid Films," *Journal of Fluid Mechanics,* Vol. 26, 1966, pp. 369–392.

6 Nachtscheim, P. R., "Analysis of the Stability of a Thin Liquid Film Adjacent to a High-Speed Gas Stream," NASA TN D-4976, Jan. 1969.

7 Starkenberg, J., "An Analysis of Stability of Thin Liquid Films in Hypersonic Environments," PIBAL Report No. 72-11, Polytechnic Institute of Brooklyn, 1972.

8 Bordner, G. L., Nayfeh, A. H., and Saric, W. S., "Stability of Liquid Films Adjacent to Compressible Streams," VPI Research Report VPI-E-73-3, Virginia Polytechnic Institute and State University, Jan. 1973.

9 Inger, G. R., "Compressible Boundary Layer Flow Past a Swept Wavy Wall With Heat Transfer and Ablation," *Astronautica Acta,* Vol. 16, 1971, pp. 325–339.

10 Yih, C-S., "Instability Due to Viscosity Stratification," *Journal of Fluid Mechanics,* Vol. 27, 1967, pp. 337–352.

11. Hame, W., and Müller, U., "Über die Stabilitat einer Zweischichten-Poiseuille Stromung," *Acta Mechanica,* Vol. 23, 1975, pp. 75–89.

12 Cohen, L. S., and Hanratty, T. J., "Generation of Waves in the Concurrent Flow of Air and a Liquid," *AIChE Journal,* Vol. 31, 1965, pp. 138–144.

13 Benjamin, T. B., "Shearing Flow Over a Wavy Boundary," *Journal of Fluid Mechanics,* Vol. 6, 1959, pp. 161–205.

14 Joshi, P. B., and Schetz, J. A., "Stability of Liquid Films Adjacent to Incompressible Gas Streams Including Effects of Interface Mass Transfer," AFOSR Scientific Report TR-77-0041, 1977.

15 Landahl, M. T., "On the Stability of a Laminar Incompressible Boundary Layer Over a Flexible Surface," *Journal of Fluid Mechanics,* Vol. 13, 1962, pp. 609–632.

The Helical Coordinate System and the Temperature Distribution Inside a Helical Coil

C.-Y. Wang[1]

Introduction

Due to its dependability and ease of operation, electrical heating is now the most important means of small scale heating. In many cases the resistive heating element is in the form of a helical coil and the heat generated is carried off by forced convection. Of interest is the maximum temperature inside the heating element, since this is a necessary design criterion.

Traditional analysis treat the helical coil as if it were a long straight cylinder [1]. The effects of curvature and torsion are neglected. In this paper we shall study whether these factors are important or not through the introduction of the new helical coordinate system which includes these effects.

The Helical Coordinate System

Let the center line be described by the space curve

$$\mathbf{R} = X(s)\mathbf{i} + Y(s)\mathbf{j} + Z(s)\mathbf{k} \tag{1}$$

where s represents arc length along this axis and $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are unit vectors in Cartesian directions. The tangent $\mathbf{T}$, normal $\mathbf{N}$, and binormal $\mathbf{B}$ can be defined as follows:

$$\mathbf{T} = \frac{d\mathbf{R}}{ds}, \quad \mathbf{N} = \frac{1}{\varkappa}\frac{d\mathbf{T}}{ds}, \quad \mathbf{B} = \mathbf{T} \times \mathbf{N} \tag{2}$$

[1] Professor, Department of Mathematics, Michigan State University, East Lansing, Mich. 48824.

Manuscript received by ASME Applied Mechanics Division, February, 1980.

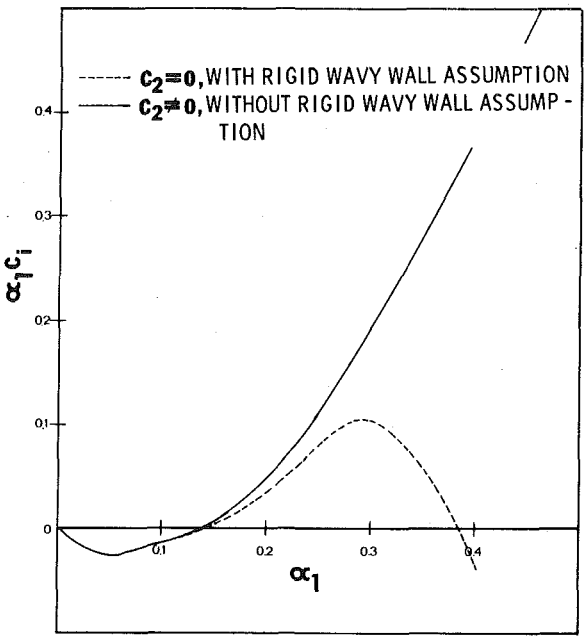


Fig. 1(*a*) Comparison of amplification curves with and without rigid wavy wall assumption

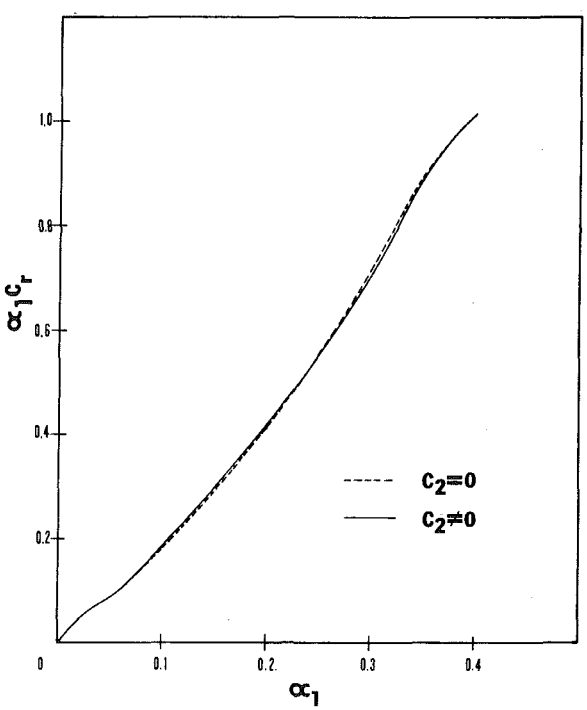


Fig. 1(*b*) Comparison of phase velocity curves with and without rigid wavy wall assumption

is necessary to study the effects of curvature of gas velocity profile and compressibility.

References

1 Lock, R. C., "Hydrodynamic Stability of the Flow in the Laminar Boundary Layer Between Parallel Streams," *Proceedings of the Cambridge Philosophical Society,* Vol. 50, 1954, pp. 105–124.

2 Feldman, S., "On the Hydrodynamic Stability of Two Viscous Incompressible Fluids in Parallel Uniform Shearing Motion," *Journal of Fluid Mechanics,* Vol. 2, 1957, pp. 343–370.

3 Miles, J. W., "The Hydrodynamic Stability of a Thin Film of Liquid in Uniform Shearing Motion," *Journal of Fluid Mechanics,* Vol. 8, 1960, pp. 593–610.

4 Cohen, L. S., and Hanratty, T. J., "Generation of Waves in the Concurrent Flow of Air and a Liquid," *AIChE Journal,* Vol. 31, 1965, pp. 138–144.

5 Craik, A. D. D., "Wind Generated Waves in Thin Liquid Films," *Journal of Fluid Mechanics,* Vol. 26, 1966, pp. 369–392.

6 Nachtscheim, P. R., "Analysis of the Stability of a Thin Liquid Film Adjacent to a High-Speed Gas Stream," NASA TN D-4976, Jan. 1969.

7 Starkenberg, J., "An Analysis of Stability of Thin Liquid Films in Hypersonic Environments," PIBAL Report No. 72-11, Polytechnic Institute of Brooklyn, 1972.

8 Bordner, G. L., Nayfeh, A. H., and Saric, W. S., "Stability of Liquid Films Adjacent to Compressible Streams," VPI Research Report VPI-E-73-3, Virginia Polytechnic Institute and State University, Jan. 1973.

9 Inger, G. R., "Compressible Boundary Layer Flow Past a Swept Wavy Wall With Heat Transfer and Ablation," *Astronautica Acta,* Vol. 16, 1971, pp. 325–339.

10 Yih, C-S., "Instability Due to Viscosity Stratification," *Journal of Fluid Mechanics,* Vol. 27, 1967, pp. 337–352.

11. Hame, W., and Müller, U., "Über die Stabilitat einer Zweischichten-Poiseuille Stromung," *Acta Mechanica,* Vol. 23, 1975, pp. 75–89.

12 Cohen, L. S., and Hanratty, T. J., "Generation of Waves in the Concurrent Flow of Air and a Liquid," *AIChE Journal,* Vol. 31, 1965, pp. 138–144.

13 Benjamin, T. B., "Shearing Flow Over a Wavy Boundary," *Journal of Fluid Mechanics,* Vol. 6, 1959, pp. 161–205.

14 Joshi, P. B., and Schetz, J. A., "Stability of Liquid Films Adjacent to Incompressible Gas Streams Including Effects of Interface Mass Transfer," AFOSR Scientific Report TR-77-0041, 1977.

15 Landahl, M. T., "On the Stability of a Laminar Incompressible Boundary Layer Over a Flexible Surface," *Journal of Fluid Mechanics,* Vol. 13, 1962, pp. 609–632.

The Helical Coordinate System and the Temperature Distribution Inside a Helical Coil

C.-Y. Wang[1]

Introduction

Due to its dependability and ease of operation, electrical heating is now the most important means of small scale heating. In many cases the resistive heating element is in the form of a helical coil and the heat generated is carried off by forced convection. Of interest is the maximum temperature inside the heating element, since this is a necessary design criterion.

Traditional analysis treat the helical coil as if it were a long straight cylinder [1]. The effects of curvature and torsion are neglected. In this paper we shall study whether these factors are important or not through the introduction of the new helical coordinate system which includes these effects.

The Helical Coordinate System

Let the center line be described by the space curve

$$\mathbf{R} = X(s)\mathbf{i} + Y(s)\mathbf{j} + Z(s)\mathbf{k} \tag{1}$$

where s represents arc length along this axis and $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are unit vectors in Cartesian directions. The tangent $\mathbf{T}$, normal $\mathbf{N}$, and binormal $\mathbf{B}$ can be defined as follows:

$$\mathbf{T} = \frac{d\mathbf{R}}{ds}, \quad \mathbf{N} = \frac{1}{\varkappa}\frac{d\mathbf{T}}{ds}, \quad \mathbf{B} = \mathbf{T} \times \mathbf{N} \tag{2}$$

[1] Professor, Department of Mathematics, Michigan State University, East Lansing, Mich. 48824.

Manuscript received by ASME Applied Mechanics Division, February, 1980.

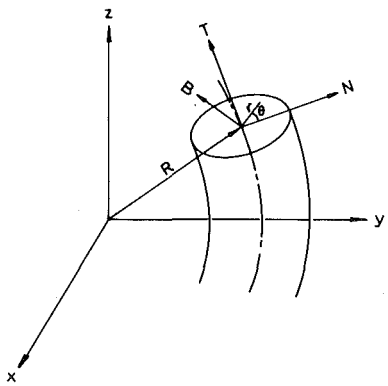


Fig. 1 The helical coordinate system

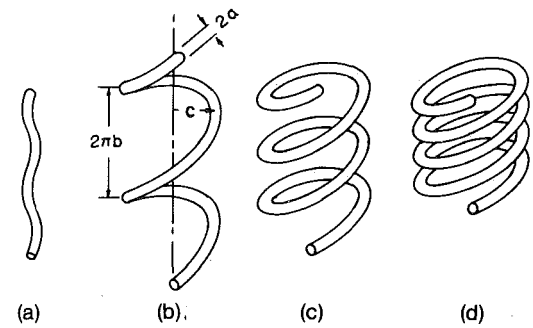


Fig. 2 Helical coils with $\varkappa a = 0.1$, (a) $\tau a = 0.3$, (b) $\tau a = 0.05$, (c) $\tau a = 0.02$, (d) $\tau a = 0.01$

Here $\varkappa$ is the curvature, and **T**, **N**, **B** are orthogonal unit vectors. The Frenet formulas give

$$\frac{d\mathbf{N}}{ds} = \tau\mathbf{B} - \varkappa\mathbf{T}, \quad \frac{d\mathbf{B}}{as} = -\tau\mathbf{N} \tag{3}$$

where τ is the torsion. We then construct a coordinate system (r, θ, s) such that any Cartesian position vector **x** can be expressed as

$$\mathbf{x} = \mathbf{R}(s) + r\cos\theta\mathbf{N}(s) + r\sin\theta\mathbf{B}(s) \tag{4}$$

See Fig. 1. Using equations (2)–(4) we obtain

$$d\mathbf{x}\cdot d\mathbf{x} = (dr)^2 + r^2(d\theta)^2 + [(1-\varkappa r\cos\theta)^2 + \tau^2r^2](ds)^2 + 2\tau r^2 ds d\theta \tag{5}$$

The description of any point in this system is unique for $r \le \varkappa^{-1}$. Notice that the last term in equation (5) indicates the nonorthogonality of (r, θ, s). Using $x^1 \equiv r$, $x^2 \equiv \theta$, $x^3 \equiv s$, the metric tensors g_{ij} and g^{ij} can be derived from equation (5).

$$g_{11} = g^{11} = 1, \quad g_{22} = r^2, \quad g^{22} = G/r^2M,$$

$$g_{33} = G, \quad g^{33} = \frac{1}{M}, \quad g_{23} = \tau r^2, \quad g^{23} = -\tau/M$$

$$g_{12} = g^{12} = g_{13} = g^{13} = 0 \tag{6}$$

$$G \equiv (1-\varkappa r\cos\theta)^2 + \tau^2r^2, \quad M = (1-\varkappa r\cos\theta)^2 \tag{7}$$

The nonzero Christoffel symbols are

$$\Gamma^1_{22} = -r, \Gamma^1_{23} = -\tau r, \Gamma^1_{33} = -\frac{1}{2}\frac{\partial G}{\partial r}, \Gamma^2_{21} = \frac{1}{r},$$

$$\Gamma^2_{13} = \frac{\tau}{M}\left(\frac{G}{r} - \frac{1}{2}\frac{\partial G}{\partial r}\right), \Gamma^2_{23} = \frac{-\tau}{2M}\frac{\partial G}{\partial\theta}, \Gamma^2_{33} = \frac{-G}{2r^2M}\frac{\partial G}{\partial\theta}$$

$$\Gamma^3_{13} = \frac{-r\tau^2}{M} + \frac{1}{2M}\frac{\partial G}{\partial r}, \Gamma^3_{23} = \frac{1}{2M}\frac{\partial G}{\partial\theta}, \Gamma^3_{33} = \frac{\tau}{2M}\frac{\partial G}{\partial\theta} \tag{8}$$

Equations (6) and (8) enable us to write any equation of continuum mechanics in helical coordinates.

The Temperature in a Helical Heating Element

The tensorial equation for heat conduction is

$$g^{ij}T_{ij} = g^{ij}\left(\frac{\partial^2T}{\partial x^i\partial x^j} - \Gamma^k_{ij}\frac{\partial T}{\partial x^k}\right) = -\frac{q}{K} \tag{9}$$

where K is the thermal conductivity, which does not vary much with temperature T, and q is the internal production heat per volume

$$q = q_0[1 + \delta(T - T_0)] \tag{10}$$

Here $q_0 = I^2R_0/V$ is the Joule heating per volume V due to current I and the resistance R_0 at temperature T_0. δ is the temperature coefficient of the electric resistivitiy. For pure metals δ is essentially a constant [1].

We shall assume the helical coil is made of a homogeneous metal wire of circular cross section with radius a. Let the center line of the wire be described by the helix

$$\mathbf{R} = c\cos\frac{s}{\sqrt{b^2+c^2}}\mathbf{i} + c\sin\frac{s}{\sqrt{b^2+c^2}}\mathbf{j} + \frac{bs}{\sqrt{b^2+c^2}}\mathbf{k} \tag{11}$$

where b and c are constants. The center line lies on the surface of a circular cylinder of radius c and the rise angle is $\tan^{-1}(b/c)$. The curvature and torsion are found to be

$$\varkappa = \frac{c}{b^2+c^2}, \quad \tau = \frac{b}{b^2+c^2} \tag{12}$$

Fig. 2 shows some of the geometries with $\varkappa a = 0.1$. It is seen that for practical purposes, both $\varkappa a$ and τa may be regarded as small. We define ϵ and α

$$\varkappa a \equiv \epsilon \ll 1, \quad \tau a \equiv \alpha\epsilon \ll 1 \tag{13}$$

where α is a constant of order unity. We assume the coil is long enough such that end effects could be neglected. This implies the temperature can be considered independent of arc length s. We also assume the surface temperature on the coil is kept constant at the ambient temperature T_0 by forced convection.

Using equations (6)–(8), equation (9) becomes

$$\frac{\partial^2T}{\partial r^2} + \left(1 - \frac{\varkappa r\cos\theta}{1-\varkappa r\cos\theta}\right)\frac{1}{r}\frac{\partial T}{\partial r} + \left(\frac{\varkappa\sin\theta}{1-\varkappa r\cos\theta}\right)\left[1 - \frac{\tau^2r^2}{(1-\varkappa r\cos\theta)^2}\right]\frac{1}{r}\frac{\partial T}{\partial\theta} + \left[1 + \frac{\tau^2r^2}{(1-\varkappa r\cos\theta)^2}\right]\frac{1}{r^2}\frac{^2T}{\partial\theta^2} = \frac{-q_0}{K}[1 + \delta(T - T_0)] \tag{14}$$

Now we normalize the variables to obtain

$$\frac{\partial^2\chi}{\partial\eta^2} + \left(1 - \frac{\epsilon\eta\cos\theta}{1-\epsilon\eta\cos\theta}\right)\frac{1}{\eta}\frac{\partial\chi}{\partial\eta} + \left(\frac{\epsilon\sin\theta}{1-\epsilon\eta\cos\theta}\right)\left[1 - \frac{\epsilon^2\alpha^2\eta^2}{(1-\epsilon\eta\cos\theta)^2}\right]\frac{1}{\eta}\frac{\partial\chi}{\partial\theta} + \left[1 + \frac{\epsilon^2\alpha^2\eta^2}{(1-\epsilon\eta\cos\theta)^2}\right]\frac{1}{\eta^2}\frac{\partial^2\chi}{\partial\theta^2} = -(1 + \lambda^2\chi) \tag{15}$$

where

$$\eta \equiv \frac{r}{a}, \quad \lambda^2 \equiv \frac{a^2q_0\delta}{K}, \quad \chi \equiv \frac{K}{a^2q_0}(T - T_0) \tag{16}$$

The boundary conditions are that χ be bounded inside the heating coil $0 \le \eta \le 1$ and that χ vanishes on the surface.

Since ϵ is small, we can expand χ in a power series

$$\chi = \chi_0 + \epsilon\chi_1 + \epsilon^2\chi_2 + \ldots \tag{17}$$

Equating equal powers of ϵ, we obtain

$$L\chi_0 \equiv \left(\frac{\partial^2}{\partial\eta^2} + \frac{1}{\eta}\frac{\partial}{\partial\eta} + \frac{1}{\eta^2}\frac{\partial^2}{\partial\theta^2} + \lambda^2\right)\chi_0 = -1 \tag{18}$$

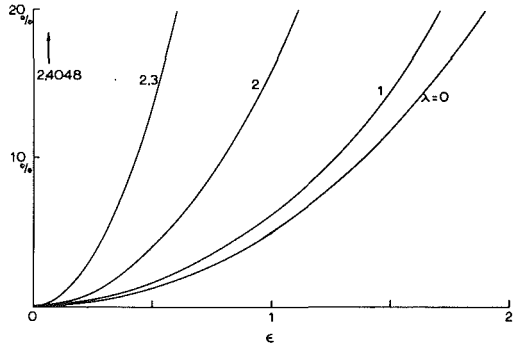


Fig. 3 Percentage increase of maximum temperature $(T_{max} - T_0)/(T_{max}|_{\epsilon=0} - T_0) - 1$

$$L\chi_1 = \cos\theta \frac{\partial \chi_0}{\partial \eta} - \frac{\sin\theta}{\eta}\frac{\partial \chi_0}{\partial \theta} \tag{19}$$

$$L\chi_2 = \cos\theta \frac{\partial \chi_1}{\partial \eta} + \eta\cos^2\theta \frac{\partial \chi_0}{\partial \eta} - \frac{\sin\theta}{\eta}\frac{\partial \chi_1}{\partial \theta} - \sin\theta\cos\theta\frac{\partial \chi_0}{\partial \theta} - \alpha^2\frac{\partial^2 \chi_0}{\partial \theta^2} \tag{20}$$

The boundary conditions are that χ_0, χ_1, χ_2 are zero on $\eta = 1$ and bounded within $\eta \leq 1$. The solutions depend on whether λ is zero or not.

The Uniformly Heated Coil

For the uniformly heated coil the local resistance (and thus heat production) can be considered independent of local temperature. This is true when $\delta(T - T_0) \ll 1$ in equation (10). (For pure copper $\delta = 0.004C^{-1}$). We set $\delta = 0$ in equation (10) or $\lambda = 0$ in equation (15). Equations (18) and (19) yield

$$\chi_0 = \frac{1-\eta^2}{4}, \quad \chi_1 = \frac{1}{16}(\eta - \eta^3)\cos\theta \tag{21}$$

Using these results equation (20) gives

$$\chi_2 = \frac{1}{128}(1 + 2\eta^2 - 3\eta^4) + \frac{5}{192}(\eta^2 - \eta^4)\cos 2\theta \tag{22}$$

χ_0 represents the parabolic temperature distribution for a straight wire. χ_1 and χ_2 are corrections due to curvature.

The Nonuniformly Heated Coil

In this case $\delta\ (T - T_0)$ is large compared to unity and $\lambda \neq 0$. Equations (18) yields

$$\chi_0 = \frac{1}{\lambda^2}\left[\frac{J_0(\lambda\eta)}{J_0(\lambda)} - 1\right] \tag{23}$$

χ_0 was first obtained by Jakob [2] for the temperature distribution inside a straight circular cylinder. He noticed that equation (23) is valid for $\lambda < 2.4048$ which is the first zero of the Bessel function J_0. If $\lambda \geq 2.4048$ or $a \geq 2.4048\ (K/q_0\,\delta)^{1/2}$, the heat generated in the interior cannot be dissipated through the walls fast enough. A subsequent mutual increase in temperature and resistance would lead to very high interior temperature. The higher corrections are found to be

$$\chi_1 = \frac{\cos\theta}{2\lambda^2}\left[\frac{\eta J_0(\lambda\eta)}{J_0(\lambda)} - \frac{J_1(\lambda\eta)}{J_1(\lambda)}\right] \tag{24}$$

$$\chi_2 = \frac{1}{16\lambda^2}\left\{\frac{3\eta^2 J_0(\lambda\eta)}{J_0(\lambda)} - 2\left[\frac{1}{\lambda J_0(\lambda)} + \frac{1}{J_1(\lambda)}\right]\eta J_1(\lambda\eta) - \left[1 - \frac{2J_1(\lambda)}{\lambda J_0(\lambda)}\right]\frac{J_0(\lambda\eta)}{J_0(\lambda)}\right\} \tag{25}$$

$$+ \frac{\cos 2\theta}{16\lambda^2}\left\{\frac{3\eta^2 J_0(\lambda\eta)}{J_0(\lambda)} - \frac{2\eta J_1(\lambda)}{J_1(\lambda)} - \frac{J_2(\lambda\eta)}{J_2(\lambda)}\right\} \quad (Cont.)$$

Results and Discussion

We note that the constant coefficients in χ_1 and χ_2 are geometrically smaller. This means that the perturbation may be valid for larger values of ϵ. In fact, convergence is possible even if $\epsilon = O(1)$. The maximum temperature is located slightly off-centered on $\theta = 0$. The percentage increase of a curved coil over that of a striaght coil is quite significant (Fig. 3). We see that the maximum temperature increases rapidly with curvature. To the order considered, torsion has no effect.

The helical coordinate system was first attempted by Nicholson [3] who erroneously thought the system was orthogonal. It was subsequently published (erroneously) in mathematical tables [4]. The present work presents the correct version for the first time. This system can be applied to other important engineering devices, such as helical springs and helical fluid cooling tubes.

References

1 Jakob, M., *Heat Transfer*, Wiley, New York, Vol. 1, 1949, pp. 173–194.
2 Jakob, M., "Influence of Nonuniform Development of Heat Upon the Temperature Distribution in Electrical Coils and Similar Heat Sources of Simple Form," TRANS. ASME, Vol. 65, 1943, pp. 593–605.
3 Nicholson, J. W., "The Effective Resistance and Inductance of a Helical Coil," *Philosophical Magazine*, Vol. 19, 1910, pp. 77–91.
4 *Smithsonian Mathematical Formulae and Tables of Elliptic Functions*, compiled by E. P. Adams and R. L. Hippisley, The Smithsonian Institution, Washington, D.C., 1922. p. 106.

On the Numerical Solution of Volterra Integral Equations Arising in Linear Viscoelasticity

A. Wineman[1]

Introduction

A numerical method for the solution of linear Volterra integral equations which arise in linear viscoelasticity has been presented by Lee and Rogers [1]. Two results are presented which improve it.

Method of Solution; Time Selection

The equations which arise have the form

$$f(t) + \int_0^t \dot{R}(t-s)\,f(s)ds = g(t), \tag{1}$$

where $f(t)$ is to be found, $g(t)$ and $R(t)$ are specified functions, and $\dot{R}(t) = dR(t)/ds$. $R(t)$ is expressed in terms of the usual material property functions such as creep or relaxation functions in shear or extension. As such, it can be expected to be monotonic. It is assumed that $f(t)$, $g(t)$, and $R(t)$ are dimensionless.

Let $\{t_k\} = \{t_1 = 0, t_2, \ldots, t_k, \ldots, t_n = t\}$ denote the set of times through t_n at which values of the solution to (1), $\{f(t_k)\}$, are to be determined. The integral in (1), denoted by I, is defined on a subset of

[1] Professor, Department of Mechanical Engineering and Applied Mechanics, The University of Michigan, Ann Arbor, Mich. 48109. Mem. ASME.
Manuscript received by ASME Applied Mechanics Division, July, 1979; final revision, May, 1980.

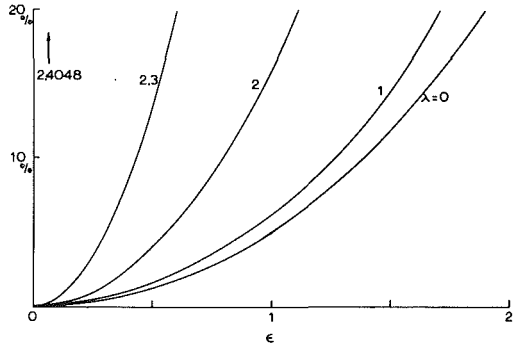


Fig. 3 Percentage increase of maximum temperature $(T_{max} - T_0)/(T_{max|\epsilon=0} - T_0) - 1$

$$L\chi_1 = \cos\theta \frac{\partial \chi_0}{\partial \eta} - \frac{\sin\theta}{\eta}\frac{\partial \chi_0}{\partial \theta} \tag{19}$$

$$L\chi_2 = \cos\theta \frac{\partial \chi_1}{\partial \eta} + \eta\cos^2\theta \frac{\partial \chi_0}{\partial \eta} - \frac{\sin\theta}{\eta}\frac{\partial \chi_1}{\partial \theta} - \sin\theta\cos\theta\frac{\partial \chi_0}{\partial \theta} - \alpha^2\frac{\partial^2 \chi_0}{\partial \theta^2} \tag{20}$$

The boundary conditions are that χ_0, χ_1, χ_2 are zero on $\eta = 1$ and bounded within $\eta \leq 1$. The solutions depend on whether λ is zero or not.

The Uniformly Heated Coil

For the uniformly heated coil the local resistance (and thus heat production) can be considered independent of local temperature. This is true when $\delta(T - T_0) \ll 1$ in equation (10). (For pure copper $\delta = 0.004C^{-1}$). We set $\delta = 0$ in equation (10) or $\lambda = 0$ in equation (15). Equations (18) and (19) yield

$$\chi_0 = \frac{1-\eta^2}{4}, \quad \chi_1 = \frac{1}{16}(\eta - \eta^3)\cos\theta \tag{21}$$

Using these results equation (20) gives

$$\chi_2 = \frac{1}{128}(1 + 2\eta^2 - 3\eta^4) + \frac{5}{192}(\eta^2 - \eta^4)\cos 2\theta \tag{22}$$

χ_0 represents the parabolic temperature distribution for a straight wire. χ_1 and χ_2 are corrections due to curvature.

The Nonuniformly Heated Coil

In this case $\delta\ (T - T_0)$ is large compared to unity and $\lambda \neq 0$. Equations (18) yields

$$\chi_0 = \frac{1}{\lambda^2}\left[\frac{J_0(\lambda\eta)}{J_0(\lambda)} - 1\right] \tag{23}$$

χ_0 was first obtained by Jakob [2] for the temperature distribution inside a straight circular cylinder. He noticed that equation (23) is valid for $\lambda < 2.4048$ which is the first zero of the Bessel function J_0. If $\lambda \geq 2.4048$ or $a \geq 2.4048\ (K/q_0\,\delta)^{1/2}$, the heat generated in the interior cannot be dissipated through the walls fast enough. A subsequent mutual increase in temperature and resistance would lead to very high interior temperature. The higher corrections are found to be

$$\chi_1 = \frac{\cos\theta}{2\lambda^2}\left[\frac{\eta J_0(\lambda\eta)}{J_0(\lambda)} - \frac{J_1(\lambda\eta)}{J_1(\lambda)}\right] \tag{24}$$

$$\chi_2 = \frac{1}{16\lambda^2}\left\{\frac{3\eta^2 J_0(\lambda\eta)}{J_0(\lambda)} - 2\left[\frac{1}{\lambda J_0(\lambda)} + \frac{1}{J_1(\lambda)}\right]\eta J_1(\lambda\eta) - \left[1 - \frac{2J_1(\lambda)}{\lambda J_0(\lambda)}\right]\frac{J_0(\lambda\eta)}{J_0(\lambda)}\right\} \tag{25}$$

$$+ \frac{\cos 2\theta}{16\lambda^2}\left\{\frac{3\eta^2 J_0(\lambda\eta)}{J_0(\lambda)} - \frac{2\eta J_1(\lambda)}{J_1(\lambda)} - \frac{J_2(\lambda\eta)}{J_2(\lambda)}\right\} \quad (Cont.)$$

Results and Discussion

We note that the constant coefficients in χ_1 and χ_2 are geometrically smaller. This means that the perturbation may be valid for larger values of ϵ. In fact, convergence is possible even if $\epsilon = O(1)$. The maximum temperature is located slightly off-centered on $\theta = 0$. The percentage increase of a curved coil over that of a striaght coil is quite significant (Fig. 3). We see that the maximum temperature increases rapidly with curvature. To the order considered, torsion has no effect.

The helical coordinate system was first attempted by Nicholson [3] who erroneously thought the system was orthogonal. It was subsequently published (erroneously) in mathematical tables [4]. The present work presents the correct version for the first time. This system can be applied to other important engineering devices, such as helical springs and helical fluid cooling tubes.

References

1 Jakob, M., *Heat Transfer*, Wiley, New York, Vol. 1, 1949, pp. 173–194.
2 Jakob, M., "Influence of Nonuniform Development of Heat Upon the Temperature Distribution in Electrical Coils and Similar Heat Sources of Simple Form," TRANS. ASME, Vol. 65, 1943, pp. 593–605.
3 Nicholson, J. W., "The Effective Resistance and Inductance of a Helical Coil," *Philosophical Magazine*, Vol. 19, 1910, pp. 77–91.
4 *Smithsonian Mathematical Formulae and Tables of Elliptic Functions*, compiled by E. P. Adams and R. L. Hippisley, The Smithsonian Institution, Washington, D.C., 1922. p. 106.

On the Numerical Solution of Volterra Integral Equations Arising in Linear Viscoelasticity

A. Wineman[1]

Introduction

A numerical method for the solution of linear Volterra integral equations which arise in linear viscoelasticity has been presented by Lee and Rogers [1]. Two results are presented which improve it.

Method of Solution; Time Selection

The equations which arise have the form

$$f(t) + \int_0^t \dot{R}(t-s)\,f(s)ds = g(t), \tag{1}$$

where $f(t)$ is to be found, $g(t)$ and $R(t)$ are specified functions, and $\dot{R}(t) = dR(t)/ds$. $R(t)$ is expressed in terms of the usual material property functions such as creep or relaxation functions in shear or extension. As such, it can be expected to be monotonic. It is assumed that $f(t)$, $g(t)$, and $R(t)$ are dimensionless.

Let $\{t_k\} = \{t_1 = 0, t_2, \ldots, t_k, \ldots, t_n = t\}$ denote the set of times through t_n at which values of the solution to (1), $\{f(t_k)\}$, are to be determined. The integral in (1), denoted by I, is defined on a subset of

[1] Professor, Department of Mechanical Engineering and Applied Mechanics, The University of Michigan, Ann Arbor, Mich. 48109. Mem. ASME.
Manuscript received by ASME Applied Mechanics Division, July, 1979; final revision, May, 1980.

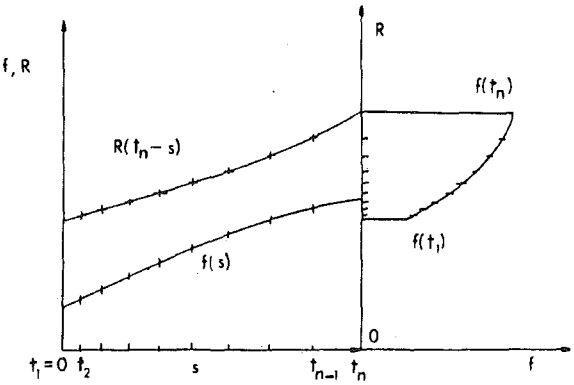


Fig. 1 Construction of $f(s)$ versus $R(t_n - s)$ in the Stieltjes integral from plots of $f(s)$ and $R(t_n - s)$ versus s

the unbounded time axis $[0, \infty)$. The selection criterion for times $\{t_k\}$ is normally based on the anticipated behavior of the solution and on trial and error.

Let (1) be written as a linear Volterra-Stieltjes integral equation in the form

$$f(t) - \int_0^t f(s)\, dR(t - s) = g(t), \tag{2}$$

where I is now defined on a subset of the bounded set $[\min R(s), \max R(s)]$. A graphical interpretation for I is shown in Fig. 1. $f(s)$ and $R(t_n - s)$ are shown plotted against the variable s. $f(s)$ is also shown plotted against $R(t_n - s)$. The value of the integral I is the area under the $f - R$ graph.

A numerical approximation for I can be developed using the definition of a Stieltjes integral [2] and the graphical interpretation of Fig. 1. Let $R(t_n - t_k)$, $(k = 1, 2, \ldots, n)$, be the set of values along the R-axis corresponding to times $\{t_k\}$. If the trapezoidal rule is used to approximate the integral, i.e., the area under the $f(s) - R(t_n - s)$ graph, one obtains the same expression as was presented in [1]. In order to obtain improved accuracy, a Simpson's rule approximation will be presented.

Consider first the part of the integral on the interval $[R(t_n - t_{k-1}), R(t_n - t_{k+1})]$. Let $H = R(t_n - t_k) - R(t_n - t_{k-1})$ and $K = R(t_n - t_{k+1}) - R(t_n - t_k)$. As suggested in Fig. 1, neither the time increments, $t_{k+1} - t_k$, along the s-axis, nor H and K on the R-axis are expected to be equal. Thus Simpson's rule for unequal intervals must be used. For the present problem, it is

$$\int_{t_{k-1}}^{t_{k+1}} f(s) dR(t_n - s) \approx \frac{(H + K)}{6HK} [K(2H - K)f(t_{k-1}) + (H + K)^2 f(t_k) + H(2K - H)f(t_{k+1})] \tag{3}$$

At each time t_n, an approximation for the entire integral can be obtained using (3). When this is substituted into (2), and the term in $f(t_n)$ is isolated, the latter can be found.

At each time t_n, there needs to be some restriction on the size of the increments H and K. For example, if K/H becomes too large, the parabolic interpolation formula on which (3) is based becomes inaccurate. Fig. 1 suggests that on the R-axis, the largest increment is $R(0) - R(t_n - t_{n-1})$. One might thus impose the restriction,

$$R(0) - R(t_n - t_{n-1}) = \alpha[R(t_n - t_{n-1}) - R(t_n - t_{n-2})], \tag{4}$$

where α is chosen according to the accuracy needs of a particular problem. This defines a nonlinear recurrance relation for the selection of times $\{t_k\}$. In particular, it establishes a rational basis for the selection of these times in accordance with the time variation of $R(s)$ and the integral approximation scheme.

Note that if the nondimensionalization is such that $\max R(s) < 1$, then $\Delta R = H < 1$. This insures the avoidance of a large error which might arise if the integral is approximated on the t-axis, and $t = t_{n+1} - t_k$ becomes $O(1)$.

Estimate of Largest t_n

For the class of problems in which $f(t)$, $g(t)$, and $R(t)$ in (1) are monotonic and bounded, an *a priori* estimate can be made for the largest solution time t_n.

Let f^∞ denote $\lim f(t)$ as $t \to \infty$ and introduce the decomposition, $f(t) = f^\infty - \Delta f(t)$. It is assumed that $t\Delta f(t) \to 0$ as $t \to \infty$ and that the integrals of $\Delta f(t)$ and $t\Delta f(t)$ on $[0,\infty)$ exist. Denote these by $A[\Delta f]$ and $A[t\Delta f]$, respectively. The centroid of $f(t)$, defined by $\tau_f = A[t\Delta f]/A[\Delta f]$ is a useful measure of the time needed for $f(t)$ to reach f^∞. Similar statements hold for $R(t)$ and $g(t)$.

If (1) is integrated on $[0,t]$, and then the decompositions of f, g, and R are substituted, it can be shown, after some manipulation, that

$$[f^\infty R^0 - g^\infty]t + \int_0^t \Delta R(t - \alpha) f(\alpha) d\alpha - f^\infty \int_0^t \Delta R(\alpha) d\alpha - R^0 \int_0^t \Delta f(\alpha) d\alpha = - \int_0^t \Delta g(\alpha) d\alpha, \tag{5}$$

where $R^0 = 1 + R(\infty) - R(0)$. Consider the limit as $t \to \infty$. It can be shown that the convolution integral vanishes. Furthermore,

$$f^\infty = g^\infty / R^0, \quad R^0 A[\Delta f] = A[\Delta g] - f^\infty A[\Delta R]. \tag{6}$$

Now let (5) be integrated on $[0,t]$. The integrals in $f(t)$, $\Delta g(t)$ and $\Delta R(t)$ can be integrated by parts and simplified using (5). In the limit as $t \to \infty$, one obtains that

$$R^0 A[t\Delta f] = A[t\Delta g] - f^\infty A[t\Delta R] - A[\Delta R] A[\Delta f], \tag{7}$$

where the last term arises from the convolution integral. Thus τ_f can be expressed in terms of $g(t)$ and $R(t)$ by

$$\tau_f = \frac{\tau_g A_g - f^\infty \tau_R A_R + (f^\infty A_R{}^2 - A_R A_g)/R^0}{A_g - f^\infty A_R} \tag{8}$$

where $A_g = A[\Delta g]$, $A_R = A[\Delta R]$. This generalizes a result of Pipkin [3].

Example

In (1), let $R(t) = 0.25 + 0.75 \exp(-t/2)$ and $g(t) = 1.0 - 0.75 \exp(-t/4)$. By use of the Laplace transform, it is found that $f(t) = 4.0 + 1.5 \exp(-t/4) - 5.25 \exp(-t/8)$.

Since $R(t)$ and $g(t)$ are monotonic, the same could be expected of $f(t)$ on physical grounds. Using (8) an estimate of the rise time of $f(t)$ is found to be $\tau_f = 8.67$. This suggests that $f(t_n)$ will be close to $f^\infty = 4.0$ when $t_n \approx 3\tau_f$.

Equation (1) was solved using the method based on (3). Solution times t_k were determined by (4) with $\alpha = 1.2$. The initial increment $R(0) - R(t_2 - t_1) = 0.01$, which implies $t_2 = 0.0268$. Computation was carried through $t_{90} = 27.064 \approx 3\tau_f$ when $f(t_{90}) = 3.824$. The maximum computational error was $0(10^{-5})$.

The method outlined here, or a variation of it, appears to be useful in solving equations such as (1). It has been used in some problems in nonlinear viscoelasticity [4].

Acknowledgment

The author wishes to express his gratitude to Mr. David Zella for computational assistance and to the National Science Foundation for its support under Grant No. ENG75-17489.

References

1 Lee, E. H., and Rogers, T. G., "Solution of Viscoelastic Stress Analysis Problems Using Measured Creep or Relaxation Functions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 30, 1963, pp. 127–133.

2 Apostol, T. M., *Mathematical Analysis*, Addison-Wesley, Reading, Mass., 1964.

3 Pipkin, A. C. *Lectures on Viscoelasticity Theory*, Springer-Verlag, New York, 1972.

4 Wineman, A. S., "On Axisymmetric Deformations of Nonlinear Viscoelastic Membranes," *Journal of Non-Newtonian Fluid Mechanics*, Vol. 4, 1978, pp. 249–260.

On the Extremal Properties of Hamilton's Action Integral

J. G. Papastavridis[1]

This Note examines the sufficient conditions for the extremization, in particular the minimization, of the action integral in Hamilton's principle for a one-degree-of-freedom nonlinear conservative system. It is usually stated in analytical mechanics that the action is actually minimized only over a short-time interval. Here the quantification of these statements is achieved by obtaining an upper bound for this minimizing interval. This is attained by combining results from the sufficiency variational calculus theory, with oscillation/comparison theorems from differential equations.

Introduction

The question of the characterization of the extremum in Hamilton's principle of "least action," assuming its existence, arises quite naturally to the student of analytical mechanics. Its treatment however is conspicuously absent from most popular texts,[2] and is briefly dealt with even in the better expositions.[3] To quote Landau and Lifshitz [5, p. 2, footnote] "... It should be noted that this "least possible value" formulation of the principle of least action is not always valid for the entire path of the system, but only for any sufficiently short segment of the path. The action integral for the entire path must have an extremum, but not necessarily a minimum. This fact, however, is of no importance as regards the derivation of the equations of motion, since only the extremum condition is used."

This (partly misleading and incorrect remark) implies that the length of these segments depends on the particular system's characteristics, i.e., its inertia, geometry, and force fields. The reason for this is that the equations of motion require only *stationarity* from the Action, i.e., only *first variation.* The extremality (minimality) though is nothing but a *condition,* and its satisfaction or not has to be determined individually, i.e., for every given system or class of systems. And this necessitates the study of the Action's *second variation.*[4]

A relatively recent treatment of this problem, essentially an elaboration of ideas from [6], has appeared in [7]. Both discussions, however, are limited to linear attractive forces; also, both either require the explicit solution of Jacobi's second variation (linear perturbation) equation, or they invoke special analytical tools.[5]

Here an attempt to remove these restrictions is made by first, assuming a nonlinear force-displacement relation, and second by using standard analytical results. In this way *the road to further generalizations and extensions remains open;* also, in the process some new light is shed into the physics of the problem.

Theory

Consider a system of mass m, generalized coordinate $q = q(t)$, and force $Q = Q(q)$. Then, according to Hamilton's variational principle, its equation of motion

$$m\ddot{q} = Q(q), \quad ((\ldots)\cdot = d(\ldots)/dt), \tag{1}$$

is obtained from

$$\delta A(q, \delta q) = 0, \tag{2}$$

where

$$A(q) = \int_{t_0}^{t_1} [T(\dot{q}) - V(q)]dt = \text{action integral}, \tag{3}$$

$$T(\dot{q}) = \tfrac{1}{2}m\dot{q}^2 = \text{kinetic energy}, \tag{4}$$

$$V(q) = V(q_0) + \int_q^{q_0} Q(x)dx = \text{potential energy}, \tag{5}$$

and the (isochronous) variations $\delta q(t)$ satisfy

$$\delta q(t_0) = \delta q(t_1) = 0, \tag{6}$$

for arbitrary times t_0, t_1.

Now, (2) and the resulting (1) are only *necessary* conditions for the minimum of (3). The *sufficient* conditions come from the study of the *second variation* of $A(q)$, $\delta^2 A(q, \delta q)$, or equivalently from Jacobi's functional

$$J(z) = \int_{t_0}^{t_1} [m\dot{z}^2 - (d^2V/dq^2)z^2]dt = \int_{t_0}^{t_1} [m\dot{z}^2 + (dQ/dq)z^2]dt. \tag{7}$$

Here

$$\begin{aligned} \Delta A &= A(q + \delta q) - A(q) \quad (= \text{total variation}) \\ &= \delta A(q, \delta q) + \tfrac{1}{2}\delta^2 A(q, \delta q) + 0_3(\delta q) \\ &\simeq \tfrac{1}{2}\delta^2 A(q, \delta q) \quad (\text{due to } (2)) \\ &= \tfrac{1}{2}J(z), \quad \text{with } \delta q(t) \equiv z(t). \end{aligned} \tag{8}$$

The relevant results are contained in the following fundamental theorem:[6]

In order that the action functional (3), with (6), attain a (*relative weak*) *minimum* in the class of *smooth* functions $q(t)$, it is sufficient that

1 $q(t)$ satisfy the *Euler-Lagrange* equation (1), i.e., that $q(t)$ be actual (dynamical) trajectory.

2 The (strengthened) *Legendre's* condition hold true at every point of the actual trajectory $q(t)$, i.e.,

$$\partial^2/\partial\dot{q}^2[T(\dot{q}) - V(q)] > 0, \quad t_0 \le t \le t_1. \tag{9}$$

3 The (strengthened) *Jacobi's* condition be satisfied: the interval $(t_0, t_1]$ should not contain any points *conjugate* to t_0.

Since 1 and 2 are always satisfied,[7] it is clear that the answer to the minimum question must come from 3. Now, the *Jacobi* or *variational* equation associated with (1) is

$$\ddot{z} + f(t)z = 0, \tag{10}$$

$$z(t_0) = 0, \dot{z}(t_0) = \text{an arbitrary } \textit{nonzero} \text{ constant } \alpha, \tag{11}$$

where

$$f(t) = (m^{-1})d^2V/dq^2 = -(m^{-1})dQ/dq, \tag{12}$$

and both q-derivatives are evaluated at $q = q(t)$, i.e., for the actual motion (1); (10) is the Euler-Lagrange equation for $\delta q = z$, of $J(z)$ in (7).

Therefore, with $\Delta t = t_1 - t_0$, Jacobi's condition 3 yields

---

[1] Assistant Professor, School of Engineering Science and Mechanics, Georgia Institute of Technology, Atlanta, Ga. 30332.

[2] See, for example, [1].

[3] For example, [2]; a more extensive discussion is contained in Routh's classic treatise [3]. Another notable exception is the (little known in the U.S.) excellent textbook of Lur'e [4]; see pp. 649–665.

[4] Compare with the situation in (conservative continuum) statics: equilibrium is answered by stationarity, or first variation, of the potential energy, whereas stability conditions require extremality, i.e., second variation (at least!).

[5] An application of "Wirtinger's inequality." From older references [8, 9] should also be mentioned; their methods and results are equivalent to those of [6, 7].

Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, June, 1980

[6] See, for example, [6 or 10].

[7] Moreover, since $\partial^2/\partial\dot{q}\,(T - V) = m > 0$, for *any* $\dot{q}$, the minimum is also *strong* (Weierstrass' condition), and further, as a consequence of the (Weierstrass-Erdmann) *corner* conditions this strong minimum has to be attained among smooth functions. Legendre's condition physically means, that *locally* the action is *always* a minimum for *any force field;* this happens because the $\dot{z}$ terms always dominate over the z terms for *very short* $(t_1 - t_0)$.

$$\Delta t < \tilde{\tau}. \tag{13}$$

Here $\tilde{\tau} = \tilde{t}_1 - t_0(\tilde{t}_1 > t_0)$, where $\tilde{t}_1$ is the first zero of the solution to the initial-value problem (10), (11), $z(t)$, to the right of t_0, i.e., $\tilde{t}_1$ is conjugate to t_0;[8] $q(t_0)$ and $q(\tilde{t}_1)$ are known as *conjugate kinetic foci.* (13) shows that for sufficiently long intervals the action *may* be a maximum, provided that it attains an extremum.

Since $q(t)$ is unknown, unless one solves the nonlinear equation of motion (1) as in [6, 8, 9] further progress toward the application of (13) can be made only by utilizing some qualitative (inequality) result relating the *distribution of zeros of $z(t)$ with the structure of (10)*, i.e., with the general behavior of $f(t)$.

This leads naturally to the well-known oscillation/comparison theorems for ODEs.[9] Their application to the minimum problem enables us to state the following basic *theorem:* Consider the action integral (3), with (4)–(6). Left $[t_a, t_b] \equiv I$, and $t_a \leq t_0 \leq t_b$. Then

1 If $d^2V/dq^2 = -dQ/dq \leq 0$ on I, the action's minimum occurs for any t_1 ($>t_0$) on I.

2 If $d^2V/dq^2 = -dQ/dq > 0$ on I, the action's minimum occurs for any t_1 ($>t_0$) on I satisfying

$$\Delta t < \tau, \tag{14}$$

where

$$\tau = \pi[m/(d^2V/dq^2)_M]^{1/2} = \pi[m/(-dQ/dq)_M]^{1/2}, \tag{15}$$

and $(\ldots)_M = \max(\ldots)$ on I.

In many cases of interest d^2V/dq^2 (or dQ/dq) keep a constant sign *throughout the entire motion.* In this case, 1 shows that the action is minimized for any t_1 ($>t_0$), i.e., any Δt, whereas 2 provides the upper bound

$$\tau = \min[\pi(m/(d^2V/dq^2))^{1/2}] = \min[\pi(m/(-dQ/dq)^{1/2}]; \tag{16}$$

so long as $\Delta t < \tau$ the action (3) can be *guaranteed* to be minimized by the actual motion (1).

In the harmonic oscillator case $Q(q) = -kq(k > 0)$, so (16) gives

$$\tau = \pi(m/k)^{1/2} = \text{half period of free oscillation}, \tag{17}$$

thus rediscovering the result of [4, 6–9].

For the harmonic oscillator with linear friction $-f\dot{q}(f > 0)$, and "damped" action[10]

$$A(q) = \int_{t_0}^{t_1} \exp(m^{-1}ft)[\tfrac{1}{2}m\dot{q}^2 - \tfrac{1}{2}kq^2]dt, \tag{18}$$

one can easily deduce that: 1 if $4km \leq f^2$, then (18) is minimized for *any* t_1, and 2 if $4km > f^2$, then (18) is minimized for any t_1 satisfying $\Delta t < \tau$, where

$$\begin{aligned}\tau &= (2\pi m)(4km - f^2)^{-1/2} \\ &= \text{half period of free (underdamped) oscillation;}\end{aligned} \tag{19}$$

(19) has been derived in [9] by *ad hoc* means.

Discussion

The first part of the theorem states that *constant* ($dQ/dq = 0$) or *repulsive* ($dQ/dq > 0$) forces always minimize the action; for example, $Q(q) = \beta^2 q^n$, β: constant and n: odd.

The second states that *attractive* ($dQ/dq < 0$) forces may minimize (3) or not depending on (14)–(16). It also indicates that *τ is of the order of the system's half period of oscillation;* if the system cannot oscillate $\tau \to \infty$.

In (15) and (16) $Q(q)$ is the force applied by the "spring" to the particle. If instead, one uses the force applied *on* the spring $Q_s = -Q$, then as the "constitutive" $Q_s - q$ diagram shows, one can replace $(-dQ/dq)_M$ with:

1 The *initial* slope $(dQ_s/dq)_{q_0}$ for *soft* springs.

2 The *maximum amplitude* slope $(dQ_s/dq)_{q_{\max}}$ for *hard* springs; the energy equation and initial conditions supply $q_{\max}$.

The oscillation/comparison theorems methodology is worth extending to several degrees-of-freedom and continuous systems, as well as to other mechanics areas, such as elastic stability; [13, 14] seem to furnish appropriate tools.

Finally, the interest in extremum principles has been recently renewed, due to their relations with the "inverse problem" of mechanics; see [15].

[8] The $\tilde{t}_1$ location is independent of α in (11). If $\Delta t = \tilde{\tau}$ further investigation is needed to answer the extremum question.

[9] See, for example, [11, pp. 636–640], or [12, pp. 171–172].

[10] Actually, only the friction need be linear.

References

1 Goldstein, H., *Classical Mechanics,* Addison-Wesley, Reading, Mass., 1950.

2 Whittaker, E. T., *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies,* Dover, New York, 1944 (originally published 1904).

3 Routh, E. J., *A Treatise on the Dynamics of a Particle,* Cambridge University Press, Cambridge, 1898.

4 Lur'e, A. I., *Analytical Mechanics* (in Russian), Gosudarstvennoe Izdat'elstvo Fiziko-Mat'emat' iskoi Literaturi, 1961, Moscow.

5 Landau, L., and Lifshitz, E., *Mechanics,* Addison-Wesley, Reading, Mass., 1960.

6 Gelfand, I. M., and Fomin, S. V., *Calculus of Variations,* Prentice-Hall, Englewood Cliffs, N. J., 1963.

7 Smith, D. R., and Smith, C. V., "When is Hamilton's Principle an Extremum Principle?, *AIAA Journal,* Vol 12, 1974, pp. 1573–1576.

8 Bottema, O., "Beispiele zum Hamiltonschen Prinzip," *Monatshefte für Mathematik,* Vol. 66, 1962, pp. 97–104.

9 Leitmann, G., "Some Remarks on Hamilton's Principle," ASME JOURNAL OF APPLIED MECHANICS, Vol. 30, 1963, pp. 623–625.

10 Petrov, I. P., *Variational Methods in Optimum Control Theory,* Academic Press, N.Y., 1968.

11 Myškis, A. D., *Advanced Mathematics for Engineers,* Mir, Moscow, 1975.

12 Bieberbach, L., *Theorie der Differentialgleichungen,* Springer-Verlag, Berlin, 1930.

13 Protter, M. H., and Weinberger, H. F., *Maximum Principles in Differential Equations,* Prentice-Hall, Englewood-Cliffs, N.J., 1967.

14 Kloetzler, R., *Mehrdimensionale Variationsrechnung,* Birkhäuser, Basel-Stuttgart, 1970.

15 Santilli, R. M., *Foundations of Theoretical Mechanics: Part I,* Springer-Verlag, NYC/Berlin, 1978.

Theory of the Constant Force Spring

C. -Y. Wang[1] and L. T. Watson[2]

Introduction

The constant force spring, or Negator, is a coil of precurved flat spring which can be opened by opposite forces F' (Fig 1). These springs have been used in balancing mechanisms, feeding devices, brush holders for electric motors and as clamps for model makers [1].

Let the spring have a natural radius of R and an equivalent flexural rigidity EI (= (Young's modulus) (width) × (thickness)^3/12 (1 − (Poisson's ratio)^2). We assume the thickness is negligible in comparison to the other dimensions. Using the strain energy of flattening a precurved spring, Votta [2] suggested the force F' is independent of displacement

[1] Professor, Department of Mathematics, Michigan State University, E. Lansing, Mich. 48824.

[2] Department of Computer Science, Virginia Polytechnic Institute, Blacksburg, Va. 24061.

Manuscript received by ASME Applied Mechanics Division, March, 1980.

$$\Delta t < \tilde{\tau}. \tag{13}$$

Here $\tilde{\tau} = \tilde{t}_1 - t_0(\tilde{t}_1 > t_0)$, where $\tilde{t}_1$ is the first zero of the solution to the initial-value problem (10), (11), $z(t)$, to the right of t_0, i.e., $\tilde{t}_1$ is conjugate to t_0;[8] $q(t_0)$ and $q(\tilde{t}_1)$ are known as *conjugate kinetic foci.* (13) shows that for sufficiently long intervals the action *may* be a maximum, provided that it attains an extremum.

Since $q(t)$ is unknown, unless one solves the nonlinear equation of motion (1) as in [6, 8, 9] further progress toward the application of (13) can be made only by utilizing some qualitative (inequality) result relating the *distribution of zeros of $z(t)$ with the structure of (10)*, i.e., with the general behavior of $f(t)$.

This leads naturally to the well-known oscillation/comparison theorems for ODEs.[9] Their application to the minimum problem enables us to state the following basic *theorem:* Consider the action integral (3), with (4)–(6). Left $[t_a, t_b] \equiv I$, and $t_a \leq t_0 \leq t_b$. Then

1 If $d^2V/dq^2 = -dQ/dq \leq 0$ on I, the action's minimum occurs for any t_1 ($>t_0$) on I.

2 If $d^2V/dq^2 = -dQ/dq > 0$ on I, the action's minimum occurs for any t_1 ($>t_0$) on I satisfying

$$\Delta t < \tau, \tag{14}$$

where

$$\tau = \pi[m/(d^2V/dq^2)_M]^{1/2} = \pi[m/(-dQ/dq)_M]^{1/2}, \tag{15}$$

and $(\ldots)_M = \max(\ldots)$ on I.

In many cases of interest d^2V/dq^2 (or dQ/dq) keep a constant sign *throughout the entire motion.* In this case, 1 shows that the action is minimized for any t_1 ($>t_0$), i.e., any Δt, whereas 2 provides the upper bound

$$\tau = \min[\pi(m/(d^2V/dq^2))^{1/2}] = \min[\pi(m/(-dQ/dq)^{1/2}]; \tag{16}$$

so long as $\Delta t < \tau$ the action (3) can be *guaranteed* to be minimized by the actual motion (1).

In the harmonic oscillator case $Q(q) = -kq(k > 0)$, so (16) gives

$$\tau = \pi(m/k)^{1/2} = \text{half period of free oscillation}, \tag{17}$$

thus rediscovering the result of [4, 6–9].

For the harmonic oscillator with linear friction $-f\dot{q}(f > 0)$, and "damped" action[10]

$$A(q) = \int_{t_0}^{t_1} \exp(m^{-1}ft)[\tfrac{1}{2}m\dot{q}^2 - \tfrac{1}{2}kq^2]dt, \tag{18}$$

one can easily deduce that: 1 if $4km \leq f^2$, then (18) is minimized for *any* t_1, and 2 if $4km > f^2$, then (18) is minimized for any t_1 satisfying $\Delta t < \tau$, where

$$\begin{aligned}\tau &= (2\pi m)(4km - f^2)^{-1/2} \\ &= \text{half period of free (underdamped) oscillation;}\end{aligned} \tag{19}$$

(19) has been derived in [9] by *ad hoc* means.

Discussion

The first part of the theorem states that *constant* ($dQ/dq = 0$) or *repulsive* ($dQ/dq > 0$) forces always minimize the action; for example, $Q(q) = \beta^2 q^n$, β: constant and n: odd.

The second states that *attractive* ($dQ/dq < 0$) forces may minimize (3) or not depending on (14)–(16). It also indicates that *τ is of the order of the system's half period of oscillation;* if the system cannot oscillate $\tau \to \infty$.

In (15) and (16) $Q(q)$ is the force applied by the "spring" to the particle. If instead, one uses the force applied *on* the spring $Q_s = -Q$, then as the "constitutive" $Q_s - q$ diagram shows, one can replace $(-dQ/dq)_M$ with:

1 The *initial* slope $(dQ_s/dq)_{q_0}$ for *soft* springs.

2 The *maximum amplitude* slope $(dQ_s/dq)_{q_{\max}}$ for *hard* springs; the energy equation and initial conditions supply $q_{\max}$.

The oscillation/comparison theorems methodology is worth extending to several degrees-of-freedom and continuous systems, as well as to other mechanics areas, such as elastic stability; [13, 14] seem to furnish appropriate tools.

Finally, the interest in extremum principles has been recently renewed, due to their relations with the "inverse problem" of mechanics; see [15].

References

1 Goldstein, H., *Classical Mechanics,* Addison-Wesley, Reading, Mass., 1950.

2 Whittaker, E. T., *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies,* Dover, New York, 1944 (originally published 1904).

3 Routh, E. J., *A Treatise on the Dynamics of a Particle,* Cambridge University Press, Cambridge, 1898.

4 Lur'e, A. I., *Analytical Mechanics* (in Russian), Gosudarstvennoe Izdat'elstvo Fiziko-Mat'emat' iskoi Literaturi, 1961, Moscow.

5 Landau, L., and Lifshitz, E., *Mechanics,* Addison-Wesley, Reading, Mass., 1960.

6 Gelfand, I. M., and Fomin, S. V., *Calculus of Variations,* Prentice-Hall, Englewood Cliffs, N. J., 1963.

7 Smith, D. R., and Smith, C. V., "When is Hamilton's Principle an Extremum Principle?, *AIAA Journal,* Vol 12, 1974, pp. 1573–1576.

8 Bottema, O., "Beispiele zum Hamiltonschen Prinzip," *Monatshefte für Mathematik,* Vol. 66, 1962, pp. 97–104.

9 Leitmann, G., "Some Remarks on Hamilton's Principle," ASME JOURNAL OF APPLIED MECHANICS, Vol. 30, 1963, pp. 623–625.

10 Petrov, I. P., *Variational Methods in Optimum Control Theory,* Academic Press, N.Y., 1968.

11 Myškis, A. D., *Advanced Mathematics for Engineers,* Mir, Moscow, 1975.

12 Bieberbach, L., *Theorie der Differentialgleichungen,* Springer-Verlag, Berlin, 1930.

13 Protter, M. H., and Weinberger, H. F., *Maximum Principles in Differential Equations,* Prentice-Hall, Englewood-Cliffs, N.J., 1967.

14 Kloetzler, R., *Mehrdimensionale Variationsrechnung,* Birkhäuser, Basel-Stuttgart, 1970.

15 Santilli, R. M., *Foundations of Theoretical Mechanics: Part I,* Springer-Verlag, NYC/Berlin, 1978.

[8] The $\tilde{t}_1$ location is independent of α in (11). If $\Delta t = \tilde{\tau}$ further investigation is needed to answer the extremum question.

[9] See, for example, [11, pp. 636–640], or [12, pp. 171–172].

[10] Actually, only the friction need be linear.

Theory of the Constant Force Spring

C. -Y. Wang[1] and L. T. Watson[2]

Introduction

The constant force spring, or Negator, is a coil of precurved flat spring which can be opened by opposite forces F' (Fig 1). These springs have been used in balancing mechanisms, feeding devices, brush holders for electric motors and as clamps for model makers [1].

Let the spring have a natural radius of R and an equivalent flexural rigidity EI (= (Young's modulus) (width) × (thickness)^3/12 (1 − (Poisson's ratio)^2). We assume the thickness is negligible in comparison to the other dimensions. Using the strain energy of flattening a precurved spring, Votta [2] suggested the force F' is independent of displacement

[1] Professor, Department of Mathematics, Michigan State University, E. Lansing, Mich. 48824.

[2] Department of Computer Science, Virginia Polytechnic Institute, Blacksburg, Va. 24061.

Manuscript received by ASME Applied Mechanics Division, March, 1980.

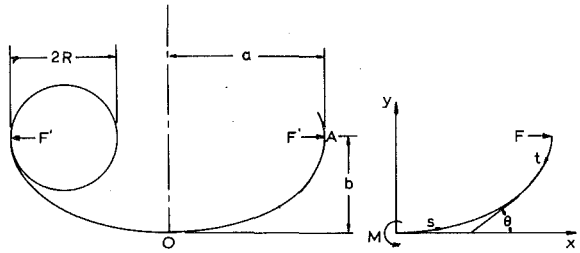


Fig. 1

$$F' = \frac{EI}{2R^2} \tag{1}$$

Experimental data [2, 3] depicts the force rising abruptly, then increasing very slowly as the spring is further extended.

Is the force really constant for a "constant-force" spring? The simplified theory of Votta could not be valid for all displacements, especially small displacements, since no part of the spring is truly flattened. We expect equation (1) serves as an asymptote of the force as the displacement tends to infinity. The present Note presents a concise theory for the constant-force spring.

Formulation

Since the thickness of the spring is small, we can use the theory of elastica to describe the spring. The equations [4] are

$$EI\frac{d\theta}{ds'} = \frac{EI}{R} - M' + F'y' \tag{2}$$

$$\frac{dx'}{ds'} = \cos\theta, \quad \frac{dy'}{ds'} = \sin\theta \tag{3}$$

where s' is the arc length, θ is the local angle of inclination, M' is the maximum moment occurring at the point of symmetry, and x' and y' are Cartesian coordinates. Let L be the arc length of the unwound spring OA. The boundary conditions are

$$x'(0) = y'(0) = \theta(0) = 0 \tag{4}$$

$$\theta(L) = \frac{\pi}{2}, \quad y'(L) = M'/F' \tag{5}$$

The unknowns are M', F', x', y', θ. We normalize all variables as follows:

$$s = s'/R,\ x = x'/R,\ y = y'/R,\ M = M'R/EI,\ F = F'R^2/EI \tag{6}$$

Equations (2)–(5) become

$$\frac{d\theta}{ds} = 1 - M + Fy \tag{7}$$

$$\frac{dx}{ds} = \cos\theta, \quad \frac{dy}{ds} = \sin\theta \tag{8}$$

$$x(0) = y(0) = \theta(0) = 0,\ \theta(L/R) = \pi/2,\ y(L/R) = M/F \tag{9}$$

We differentiate equation (7) once, multiply by $d\theta/ds$ and integrate to get

$$\frac{1}{2}\left(\frac{d\theta}{ds}\right)^2 = -F\cos\theta + \frac{1}{2} \tag{10}$$

where the constant of integration is obtained by the boundary conditions at $s = L/R$. Integrating again, we have

$$s = \int_0^\theta \frac{d\theta}{\sqrt{1 - 2F\cos\theta}} = \frac{2}{\sqrt{1+2F}} \times \left[\mathcal{F}\left(\sqrt{\frac{4F}{1+2F}}\ \frac{\pi}{2}\right) - \mathcal{F}\left(\sqrt{\frac{4F}{1+2F}}, \frac{\pi-\theta}{2}\right)\right] \tag{11}$$

where $\mathcal{F}$ is the elliptic function of the first kind.

The solution in terms of elliptic functions is extremely inconvenient. Furthermore, the accuracy is limited by the accuracy of elliptic functions themselves, either obtained from tables or from numerical integration.

Asymptotic Character as $L/R \to \infty$

This corresponds to Votta's case when the spring is pulled far apart such that $L \gg R$. We set $t = L/R - s$ in equation (10). As $t \to \infty, \theta \to 0, d\theta/dt \to 0$ we find

$$F \to 1/2 \tag{12}$$

This yields exactly Votta's result. Also, since $y \to 0$ as $t \to \infty$, equation (7) yields

$$M \to 1 \tag{13}$$

The maximum moment is thus EI/R. Equations (9), (12), and (13) give

$$b/R \equiv y\left(\frac{L}{R}\right) \to 2 \tag{14}$$

As the spring is pulled apart, the height b increase from R to $2R$. The exact form of $\theta(t)$ is

$$\theta = 4\tan^{-1}\left[\tan\frac{\pi}{8}\exp(-t/\sqrt{2})\right] \tag{15}$$

Numerical Integration

With the advent of the computer, it is simpler and more accurate to integrate equations (7)–(9) directly. One usually guesses M, F, and shoots for the boundary conditions at L/R. Newton's method is used to refine the guesses. However, this method may not be convergent unless the guesses M and F are extremely close to the true solution. We shall propose a method for which no initial guess is required.

We transform the governing equations further by

$$\bar{s} = s(1-M), \quad \bar{y} = y(1-M), \quad \bar{x} = x(1-M) \tag{16}$$

Equations (7)–(9) become

$$\frac{d\theta}{ds} = 1 + K\bar{y}, \frac{d\bar{x}}{d\bar{s}} = \cos\theta, \frac{d\bar{y}}{d\bar{s}} = \sin\theta \tag{17}$$

$$\bar{x}(0) = \bar{y}(0) = \theta(0) = 0 \tag{18}$$

$$\theta_{|\bar{s}=(1-M)L/R} = \frac{\pi}{2}, \quad \bar{y}_{|\bar{s}=(1-M)L/R} = \frac{(1-M)M}{F} \tag{19}$$

where

$$K \equiv F/(1-M)^2 \tag{20}$$

The method is to choose any K, integrate equations (17) and (18) until θ reaches $\pi/2$ say at $\bar{s} = s^*$. Then we can solve for F, M from equations (19) and (20):

$$F = K/(1 + \bar{y}(s^*)K)^2, \quad M = \bar{y}(s^*)K/(1 + \bar{y}(s^*)K) \tag{21}$$

The displacement is

$$\frac{a}{R} \equiv \bar{x}(L/R) = \frac{\bar{x}(s^*)}{1-M} = \bar{x}(s^*)(1 + \bar{y}(s^*)K) \tag{22}$$

The arc length is

$$\frac{L}{R} = s^*(1 + \bar{y}(s^*)K) \tag{23}$$

If the exact shape of the spring is needed, one can integrate equation (7)–(9) with the correct values from equations (21), (23).

Results and Discussion

Fig. 2 is a graph of nondimensional force F and maximum nondimensional moment M versus displacement a/R. It is obvious the force is not constant. However, if $a/R > 2.62$, F will be within 5 percent of its asymptotic value of ½. The maximum moment (a design paramter) approaches its asymptotic value much slower. Unfortunately, the

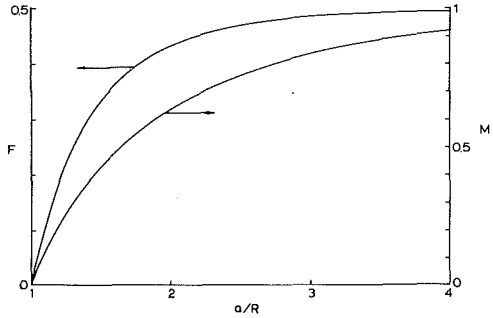


Fig. 2

natural radii are not given in references [2, 3] and we are unable to compare their experimental results to our universal curve.

References

1 Wahl, A. M., *Mechanical Springs*, McGraw-Hill, New York, 1963.
2 Votta, F. A., Jr., "The Theory and Design of Long-Deflection Constant—Force Spring Elements," TRANS. ASME, Vol. 74, 1952, pp. 439–450.
3 *Design Information for Constant Force Springs*, John Evans' Sons, Inc., Lansdale, Pa., 1979.
4 Frisch-Fay, R., *Flexible Bars*, Butterworths, 1962.

Transient Analysis of Stress Waves Around a Rectangular Crack Under Impact Load

S. Itou[1]

Introduction

In an earlier paper [1], the steady-state problem with the time factor exp $(i\omega t)$, has been resolved for an infinite elastic body weakened by a rectangular crack. In this paper, expressions are presented for the three-dimensional transient response of a solid containing a plane rectangular crack subjected to impact load. The problem is first formulated with the aid of Fourier and Laplace transforms and then reduced to the solution of dual integral equations in terms of the Laplace transform variable. These equations are solved with the same manner which is employed in reference [1], namely, the surface displacement is expanded in a double series of Jacobi polynomials and the Schmidt method is used. A numerical Laplace inversion technique, which is developed in reference [2] and is used in references [3, 4], is also taken to obtain the stresses in the physical space.

Numerical calculations are carried out for the dynamic stress-intensity factors and compared with those of the corresponding static values given by Weaver [5] and Itou [1].

Formulation of Problem

Consider an infinite elastic space with Cartesian coordinate x_i, $i = 1, 2, 3$. A rectangular crack is located along the x_1-axis from $-a$ to a and along the x_3-axis from $-b$ to b as shown in Fig. 1.

For an incident displacement wave which impinges on the rectangular crack, the boundary condition equations are as follows:

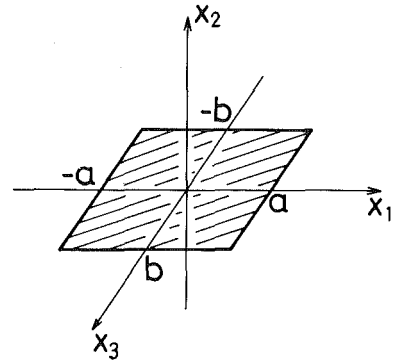


Fig. 1 Geometry and coordinate system

$$\begin{aligned} \tau_{22} &= -P\,\mathrm{H}(t), \quad \text{for } x_2 = 0, \quad |x_1| < a, \quad |x_3| < b, \\ u_2 &= 0, \quad \text{for } x_2 = 0, \quad |x_1| > a, \quad |x_3| > b, \\ \tau_{12} &= \tau_{23} = 0, \quad \text{for } x_2 = 0, \quad |x_1| < \infty, \quad |x_3| < \infty, \end{aligned} \tag{1}$$

with

$$P = (\lambda + 2\mu)\epsilon_0, \tag{2}$$

where $\mathrm{H}(t)$ is the Heaviside unite step function, τ_{ij} and u_i are the stress and displacement components, respectively, λ and μ are the Lamé elastic constants, and ϵ_0 is a constant.

Stress-Intensity Factors

By comparing the Laplace transformed equation of motion with equation (4) in reference [6], we know that all expressions in a time-harmonic steady-state space can be used in the Laplace transformed domain if we replace ω^2 by $-s^2$. Therefore, the stress-intensity factors $K_a{}^*$ along $x_1 = a$ and $K_b{}^*$ along $x_3 = b$ in the Laplace transform domain are easily defined in a manner similar to that employed by the author to solve the corresponding steady-state time-harmonic problem [1],

$$\begin{aligned} K_a{}^* &= \sqrt{2\pi(x_1 - a)}\,\tau_{22}{}^*|_{x_1 \to a+} \\ &= \frac{4K(\delta,\zeta)}{\sqrt{\pi a}\,\zeta} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_{mn}(s) \left[\frac{(-1)^{2m+n}\Gamma(2m - \frac{1}{2})\Gamma(2n - \frac{1}{2})}{(2m-2)!(2n-1)!} \right. \\ &\qquad \left. \times \cos\{(2n-1)\sin^{-1}(x_3/b)\} \right], \\ K_b{}^* &= \sqrt{2\pi(x_s - b)}\,\tau_{22}{}^*|_{x_3 \to b+} \\ &= \frac{4K(\xi,\gamma)}{\sqrt{\pi b}\,\gamma} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_{mn}(s) \left[\frac{(-1)^{m+2n}\Gamma(2m - \frac{1}{2})\Gamma(2n - \frac{1}{2})}{(2m-1)!(2n-2)!} \right. \\ &\qquad \left. \times \cos\{2(m-1)\sin^{-1}(x_1/a)\} \right], \end{aligned} \tag{3}$$

with

$$\begin{aligned} K(\delta, \zeta)/\delta = K(\xi,\gamma)/\gamma &= (1 - \alpha^2)/\alpha^2, \\ \alpha^2 &= (\lambda + 2\mu)/\mu, \end{aligned} \tag{4}$$

where $c_{mn}(s)$ are the coefficients and can be solved by the Schmidt method [1], and $\Gamma(m)$ is Gamma function.

The Laplace inverse transformations in equation (3) are carried out by the numerical method given by Miller and Guy [2].

Numerical Examples and Results

For a numerical calculation, Poisson's ratio ν is taken as 0.2 and the shape of the crack is assumed to be a regular square, namely, $a/b = 1.0$. In Table 1, the values of K_a at $x_3 = 0$ and K_b at $x_1 = 0$ are given where c_2 is the shear wave velocity. Since the ratio b/a is 1.0, it should hold that $K_a = K_b$. There is a difference between these. Since it is very

[1] Associate Professor, Department of Mechanical Engineering, Hachinohe Institute of Technology, Hachinohe 031, Japan.
Manuscript received by ASME Applied Mechanics Division, April, 1980.

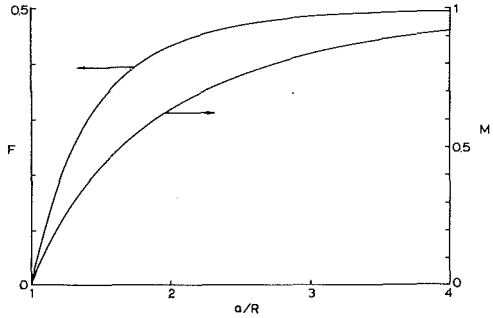


Fig. 2

natural radii are not given in references [2, 3] and we are unable to compare their experimental results to our universal curve.

References

1 Wahl, A. M., *Mechanical Springs*, McGraw-Hill, New York, 1963.
2 Votta, F. A., Jr., "The Theory and Design of Long-Deflection Constant—Force Spring Elements," TRANS. ASME, Vol. 74, 1952, pp. 439–450.
3 *Design Information for Constant Force Springs*, John Evans' Sons, Inc., Lansdale, Pa., 1979.
4 Frisch-Fay, R., *Flexible Bars*, Butterworths, 1962.

Transient Analysis of Stress Waves Around a Rectangular Crack Under Impact Load

S. Itou[1]

Introduction

In an earlier paper [1], the steady-state problem with the time factor exp $(i\omega t)$, has been resolved for an infinite elastic body weakened by a rectangular crack. In this paper, expressions are presented for the three-dimensional transient response of a solid containing a plane rectangular crack subjected to impact load. The problem is first formulated with the aid of Fourier and Laplace transforms and then reduced to the solution of dual integral equations in terms of the Laplace transform variable. These equations are solved with the same manner which is employed in reference [1], namely, the surface displacement is expanded in a double series of Jacobi polynomials and the Schmidt method is used. A numerical Laplace inversion technique, which is developed in reference [2] and is used in references [3, 4], is also taken to obtain the stresses in the physical space.

Numerical calculations are carried out for the dynamic stress-intensity factors and compared with those of the corresponding static values given by Weaver [5] and Itou [1].

Formulation of Problem

Consider an infinite elastic space with Cartesian coordinate x_i, $i = 1, 2, 3$. A rectangular crack is located along the x_1-axis from $-a$ to a and along the x_3-axis from $-b$ to b as shown in Fig. 1.

For an incident displacement wave which impinges on the rectangular crack, the boundary condition equations are as follows:

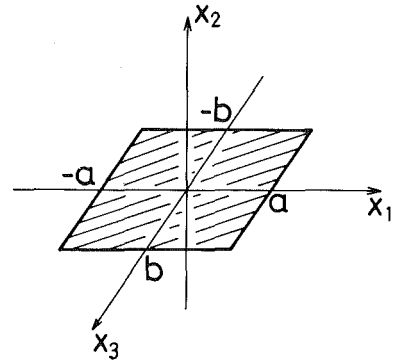


Fig. 1 Geometry and coordinate system

$$\tau_{22} = -P\,\mathrm{H}(t), \quad \text{for } x_2 = 0, \quad |x_1| < a, \quad |x_3| < b,$$
$$u_2 = 0, \quad \text{for } x_2 = 0, \quad |x_1| > a, \quad |x_3| > b,$$
$$\tau_{12} = \tau_{23} = 0, \quad \text{for } x_2 = 0, \quad |x_1| < \infty, \quad |x_3| < \infty, \tag{1}$$

with

$$P = (\lambda + 2\mu)\epsilon_0, \tag{2}$$

where $\mathrm{H}(t)$ is the Heaviside unite step function, τ_{ij} and u_i are the stress and displacement components, respectively, λ and μ are the Lamé elastic constants, and ϵ_0 is a constant.

Stress-Intensity Factors

By comparing the Laplace transformed equation of motion with equation (4) in reference [6], we know that all expressions in a time-harmonic steady-state space can be used in the Laplace transformed domain if we replace ω^2 by $-s^2$. Therefore, the stress-intensity factors $K_a{}^*$ along $x_1 = a$ and $K_b{}^*$ along $x_3 = b$ in the Laplace transform domain are easily defined in a manner similar to that employed by the author to solve the corresponding steady-state time-harmonic problem [1],

$$K_a{}^* = \sqrt{2\pi(x_1 - a)}\,\tau_{22}{}^*|_{x_1 \to a+}$$
$$= \frac{4K(\delta,\zeta)}{\sqrt{\pi a}\,\zeta} \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} c_{mn}(s) \left[\frac{(-1)^{2m+n}\Gamma(2m-\frac{1}{2})\Gamma(2n-\frac{1}{2})}{(2m-2)!(2n-1)!} \times \cos\{(2n-1)\sin^{-1}(x_3/b)\}\right],$$

$$K_b{}^* = \sqrt{2\pi(x_s - b)}\,\tau_{22}{}^*|_{x_3 \to b+}$$
$$= \frac{4K(\xi,\gamma)}{\sqrt{\pi b}\,\gamma} \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} c_{mn}(s) \left[\frac{(-1)^{m+2n}\Gamma(2m-\frac{1}{2})\Gamma(2n-\frac{1}{2})}{(2m-1)!(2n-2)!} \times \cos\{2(m-1)\sin^{-1}(x_1/a)\}\right], \tag{3}$$

with

$$K(\delta,\zeta)/\delta = K(\xi,\gamma)/\gamma = (1-\alpha^2)/\alpha^2,$$
$$\alpha^2 = (\lambda + 2\mu)/\mu, \tag{4}$$

where $c_{mn}(s)$ are the coefficients and can be solved by the Schmidt method [1], and $\Gamma(m)$ is Gamma function.

The Laplace inverse transformations in equation (3) are carried out by the numerical method given by Miller and Guy [2].

Numerical Examples and Results

For a numerical calculation, Poisson's ratio ν is taken as 0.2 and the shape of the crack is assumed to be a regular square, namely, $a/b = 1.0$. In Table 1, the values of K_a at $x_3 = 0$ and K_b at $x_1 = 0$ are given where c_2 is the shear wave velocity. Since the ratio b/a is 1.0, it should hold that $K_a = K_b$. There is a difference between these. Since it is very

[1] Associate Professor, Department of Mechanical Engineering, Hachinohe Institute of Technology, Hachinohe 031, Japan.
Manuscript received by ASME Applied Mechanics Division, April, 1980.

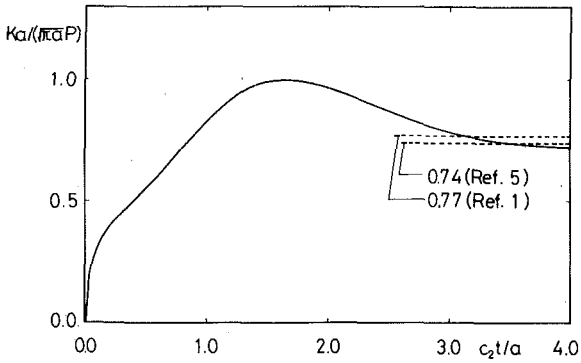


Fig. 2 Dynamic stress-intensity factor K_a at $x_3/b = 0.0$

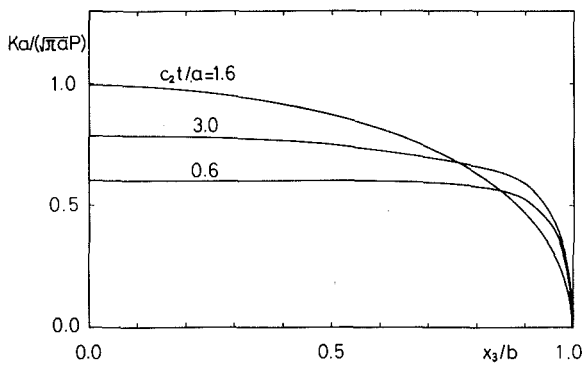


Fig. 3 Dynamic stress-intensity factor K_a for $c_2t/a = 0.6, 1.6, 3.0$ versus x_3

Table 1 The values K_a at $x_3/b = 0.0$ and K_b at $x_1/a = 0.0$

c_2t/a	0.01	1.00	1.60	2.00	3.00
$K_a/(\sqrt{\pi a}P)$	0.037	0.830	0.995	0.963	0.786
$K_b/(\sqrt{\pi b}P)$	0.035	0.825	0.986	0.965	0.794

small, it is considered that the accuracy of the present numerical calculations are satisfactory from an engineering viewpoint.

In Fig. 2, K_a at $x_3 = 0$ is plotted versus t, in which the broken lines are the corresponding static values given by Weaver [5] and Itou [1]. There is a small difference between these static solutions. Therefore, it is considered that the dynamic stress-intensity factor exceeds the corresponding static value by the range from 29 to 35 percent. Fig. 3 shows K_a versus x_3/b for $c_2t/a = 0.6, 1.6, 3.0$. At the corner of the crack the stress-intensity factor drops to zero. This means that the singularity of the stress at the corner is different from the usual one. When we write τ_{22} at the corners as the following:

$$\tau_{22} = k_a/(x_1^2 - a^2)^{\epsilon_1} + k_b/(x_3^2 - b^2)^{\epsilon_2}, \tag{5}$$

with constants k_a, k_b, ϵ_1, ϵ_2, the values of ϵ_1 and ϵ_2 are smaller than 0.5. To obtain these exact values is beyond the present numerical analysis. However, an existing crack has a dull corner rather than a sharp one, and then the corner singularity is not so much interesting in fracture mechanics. For the contrary case of an external crack, ϵ_1 and ϵ_2 may have the larger values than 0.5.

Acknowledgments

The author wishes to express thanks to Prof. A. Atsumi, Tohoku University, for his suggestions. A part of this research was carried with a Scientific Research Grant from the Ministry of Education, to which the author is deeply grateful.

References

1 Itou, S., "Dynamic Stress Concentration Around a Rectangular Crack in an Infinite Elastic Medium," *ZAMM*, in press.

2 Miller, M. K., and Guy, W. T., "Numerical Inversion of the Laplace Transform by Use of Jacobi polynomials," *SIAM Journal of Numerical Analysis*, Vol. 3, 1966, pp. 624–635.

3 Chen, E. P., "Impact Response of a Finite Crack in a Finite Strip Under Antiplane Shear," *Engineering Fracture Mechanics*, Vol. 9, 1977, pp. 719–724.

4 Itou, S., "Transient Analysis of Stress Waves Around Two Coplanar Griffith Cracks Under Impact Load," *Engineering Fracture Mechanics*, Vol. 13, 1980, pp. 349–356.

5 Weaver, J., "Three-Dimensional Crack Analysis," *International Journal of Solids and Structures*, Vol. 13, 1977, pp. 321–330.

6 Itou, S., "Three-Dimensional Wave Propagation in Cracked Elastic Solids," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 807–811.

On the Analysis of First and Second-Order Shear Deformation Effects for Isotropic Elastic Plates[1]

E. Reissner[2]

Introduction

In what follows we consider once more briefly the problem of transverse shear deformations for isotropic plates, within the framework of the two-dimensional sixth-order theory as derived from three-dimensional theory by a variational method [2] or, alternately, by means of self-contained two-dimensional considerations [3].

Specifically, we are here concerned with the fact that it is possible to distinguish between first and second-order shear deformation effects, with the determination of the first-order effects depending on a rational analysis of edge-zone behavior and with the second-order effects requiring no such analysis [5]. As regards the nature of the two kinds of effects we note, in particular, that the second-order effect is a natural generalization of Timoshenko's analysis of shear deformation in beams while the first-order effect disappears in a specialization of the plate problem to the corresponding problem of the beam.

As regards the objects of this Note, these are as follows.

Recent considerations by Simmonds [4], including a description of results by Goldenveiser [1] concerning the asymptotic derivation of a fourth-order plate theory in which first-order shear correction terms are accounted for by a modification of the classical Kirchhoff boundary conditions, make it seem worthwhile to indicate that results of the same nature are in fact implied by the writer's sixth-order two-dimensional plate theory.[3]

Our results on modified Kirchhoff boundary conditions in [3] were stated for the case of straight edges only. It seems desirable to present a derivation of the corresponding results for the case of curved edges inasmuch as edge curvature brings with it a significant supplementary term in the relevant formulas.

[1] Supported by the Office of Naval Research.

[2] Department of Applied Mechanics and Engineering Sciences, University of California, San Diego, La Jolla, Calif. 92093. Fellow ASME.

[3] We note that we did not think of this possibility in the presentation of our earlier work [2], and that our later analysis by a direct two-dimensional approach [3] led us to this possibility without consciousness of its relation to Goldenveiser's results.

Manuscript received by ASME Applied Mechanics Division, March, 1980.

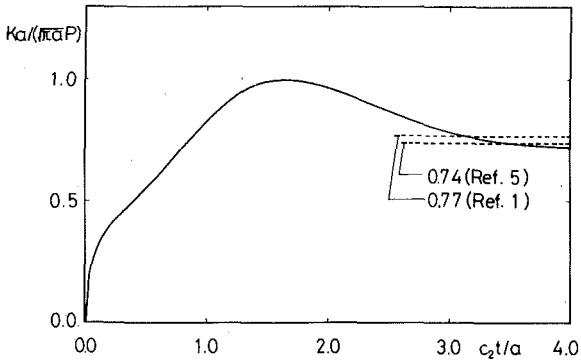


Fig. 2 Dynamic stress-intensity factor K_a at $x_3/b = 0.0$

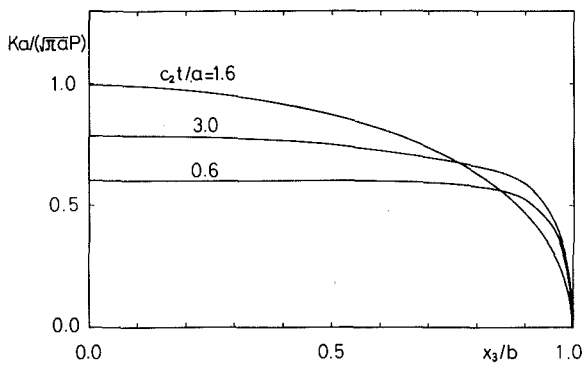


Fig. 3 Dynamic stress-intensity factor K_a for $c_2t/a = 0.6, 1.6, 3.0$ versus x_3

Table 1 The values K_a at $x_3/b = 0.0$ and K_b at $x_1/a = 0.0$

c_2t/a	0.01	1.00	1.60	2.00	3.00
$K_a/(\sqrt{\pi a}P)$	0.037	0.830	0.995	0.963	0.786
$K_b/(\sqrt{\pi b}P)$	0.035	0.825	0.986	0.965	0.794

small, it is considered that the accuracy of the present numerical calculations are satisfactory from an engineering viewpoint.

In Fig. 2, K_a at $x_3 = 0$ is plotted versus t, in which the broken lines are the corresponding static values given by Weaver [5] and Itou [1]. There is a small difference between these static solutions. Therefore, it is considered that the dynamic stress-intensity factor exceeds the corresponding static value by the range from 29 to 35 percent. Fig. 3 shows K_a versus x_3/b for $c_2t/a = 0.6, 1.6, 3.0$. At the corner of the crack the stress-intensity factor drops to zero. This means that the singularity of the stress at the corner is different from the usual one. When we write τ_{22} at the corners as the following:

$$\tau_{22} = k_a/(x_1^2 - a^2)^{\epsilon_1} + k_b/(x_3^2 - b^2)^{\epsilon_2}, \tag{5}$$

with constants k_a, k_b, ϵ_1, ϵ_2, the values of ϵ_1 and ϵ_2 are smaller than 0.5. To obtain these exact values is beyond the present numerical analysis. However, an existing crack has a dull corner rather than a sharp one, and then the corner singularity is not so much interesting in fracture mechanics. For the contrary case of an external crack, ϵ_1 and ϵ_2 may have the larger values than 0.5.

Acknowledgments

The author wishes to express thanks to Prof. A. Atsumi, Tohoku University, for his suggestions. A part of this research was carried with a Scientific Research Grant from the Ministry of Education, to which the author is deeply grateful.

References

1 Itou, S., "Dynamic Stress Concentration Around a Rectangular Crack in an Infinite Elastic Medium," *ZAMM*, in press.

2 Miller, M. K., and Guy, W. T., "Numerical Inversion of the Laplace Transform by Use of Jacobi polynomials," *SIAM Journal of Numerical Analysis*, Vol. 3, 1966, pp. 624–635.

3 Chen, E. P., "Impact Response of a Finite Crack in a Finite Strip Under Antiplane Shear," *Engineering Fracture Mechanics*, Vol. 9, 1977, pp. 719–724.

4 Itou, S., "Transient Analysis of Stress Waves Around Two Coplanar Griffith Cracks Under Impact Load," *Engineering Fracture Mechanics*, Vol. 13, 1980, pp. 349–356.

5 Weaver, J., "Three-Dimensional Crack Analysis," *International Journal of Solids and Structures*, Vol. 13, 1977, pp. 321–330.

6 Itou, S., "Three-Dimensional Wave Propagation in Cracked Elastic Solids," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 807–811.

On the Analysis of First and Second-Order Shear Deformation Effects for Isotropic Elastic Plates[1]

E. Reissner[2]

Introduction

In what follows we consider once more briefly the problem of transverse shear deformations for isotropic plates, within the framework of the two-dimensional sixth-order theory as derived from three-dimensional theory by a variational method [2] or, alternately, by means of self-contained two-dimensional considerations [3].

Specifically, we are here concerned with the fact that it is possible to distinguish between first and second-order shear deformation effects, with the determination of the first-order effects depending on a rational analysis of edge-zone behavior and with the second-order effects requiring no such analysis [5]. As regards the nature of the two kinds of effects we note, in particular, that the second-order effect is a natural generalization of Timoshenko's analysis of shear deformation in beams while the first-order effect disappears in a specialization of the plate problem to the corresponding problem of the beam.

As regards the objects of this Note, these are as follows.

Recent considerations by Simmonds [4], including a description of results by Goldenveiser [1] concerning the asymptotic derivation of a fourth-order plate theory in which first-order shear correction terms are accounted for by a modification of the classical Kirchhoff boundary conditions, make it seem worthwhile to indicate that results of the same nature are in fact implied by the writer's sixth-order two-dimensional plate theory.[3]

Our results on modified Kirchhoff boundary conditions in [3] were stated for the case of straight edges only. It seems desirable to present a derivation of the corresponding results for the case of curved edges inasmuch as edge curvature brings with it a significant supplementary term in the relevant formulas.

[1] Supported by the Office of Naval Research.

[2] Department of Applied Mechanics and Engineering Sciences, University of California, San Diego, La Jolla, Calif. 92093. Fellow ASME.

[3] We note that we did not think of this possibility in the presentation of our earlier work [2], and that our later analysis by a direct two-dimensional approach [3] led us to this possibility without consciousness of its relation to Goldenveiser's results.

Manuscript received by ASME Applied Mechanics Division, March, 1980.

Two-Dimensional Plate Equations in Polar Coordinate Form

We depart from our earlier Cartesian-coordinate statement of sixth-order two-dimensional theory for plates which are two-dimensionally isotropic and homogeneous [3] and rewrite these (for the case of absent distributed surface loads) with reference to polar coordinates r, θ in the form

$$M_{rr} = -D\left[v_{,rr} + \nu\left(\frac{v_{,r}}{r} + \frac{v_{,\theta\theta}}{r^2}\right)\right] + \frac{2}{\lambda^2}\left(\frac{\chi_{,\theta r}}{r} - \frac{\chi_{,\theta}}{r^2}\right) \tag{1}$$

$$M_{\theta\theta} = -D\left[\frac{v_{,r}}{r} + \frac{v_{,\theta\theta}}{r^2} + \nu v_{,rr}\right] - \frac{2}{\lambda^2}\left(\frac{\chi_{,\theta r}}{r} - \frac{\chi_{,\theta}}{r^2}\right) \tag{2}$$

$$M_{r\theta} = -(1-\nu)\,D\left[\frac{v_{,\theta r}}{r} - \frac{v_{,\theta}}{r^2}\right] + \frac{2}{\lambda^2}\left(\frac{\chi_{,r}}{r} + \frac{\chi_{,\theta\theta}}{r^2}\right) - \chi \tag{3}$$ [4]

$$Q_r = -D(\nabla^2 v)_{,r} + \frac{\chi_{,\theta}}{r}, \quad Q_\theta = -D\frac{(\nabla^2 v)_{,\theta}}{r} - \chi_{,r}, \tag{4a,b}$$

$$\phi_r = -v_{,r} + B\frac{\chi_{,\theta}}{r}, \quad \phi_\theta = \frac{v_{,\theta}}{r} - B\chi_{,r}. \tag{5a,b}$$

In this we have

$$\frac{1}{\lambda^2} = \frac{1-\nu}{2}BD, \quad v = w + BD\nabla^2 w, \tag{6a,b}$$

with D and B being transverse bending and shear deformation factors, and w and χ being solutions of the differential equations

$$D\,\nabla^2\nabla^2 w = 0, \quad \nabla^2\chi - \lambda^2\chi = 0. \tag{7a,b}$$

The factors D and B are, for the case of a plate which is also homogeneous in thickness direction

$$D = \frac{Eh^3}{12(1-\nu^2)}, \quad B = \frac{6}{5Gh}. \tag{8a,b}$$

Therewith, for this case

$$\frac{1}{\lambda^2} = \frac{Eh^2}{20(1+\nu)G}, \tag{8c}$$

and then $\lambda = \sqrt{10}/h$ for a three-dimensionally isotropic material as considered in [2].

Asymptotic Analysis

Given a circular ring plate with inner edge $r = a$ we consider the system of stress boundary conditions,

$$r = a; \quad M_{rr} = \overline{M}_{rr}, \quad M_{r\theta} = \overline{M}_{r\theta}, \quad Q_r = \overline{Q}_r, \tag{9}$$

and, alternately, the system of displacement boundary conditions,

$$r = a; \quad w = \overline{w}, \quad \phi_r = \overline{\phi}_r, \quad \phi_\theta = \overline{\phi}_\theta. \tag{10}$$

The possibility of an asymptotic analysis is given upon making the fundamental assumption

$$1 \ll \lambda a, \tag{11}$$

and upon making use of the order-of-magnitude relations

$$w_{,r} = O\left(\frac{w}{a}\right), \quad \frac{w_{,\theta}}{r} = O\left(\frac{w}{a}\right), \quad \frac{\chi_{,\theta}}{r} = O\left(\frac{w}{a}\right), \quad w_{,r} = O(\lambda\chi), \tag{12}$$

with these depending on restrictive assumptions of the form $\overline{M}_{rr,\theta} = O(\overline{M}_{rr})$, etc.

It follows from (12) and (6) that now

$$v = w + O\left(\frac{w}{\lambda^2 a^2}\right), \tag{13}$$

and, if we designate v-contributions to M_{rr}, etc., by a superscript i, the boundary conditions (9) may be written in the form

$$M^i_{rr} + \frac{2}{\lambda^2}\left(\frac{\chi_{,\theta r}}{r} - \frac{\chi_{,\theta}}{r^2}\right) = \overline{M}_{rr}, \tag{14a}$$

$$M^i_{r\theta} + \frac{2}{\lambda^2}\left(\frac{\chi_{,r}}{r} + \frac{\chi_{,\theta\theta}}{r^2}\right) - \chi = \overline{M}_{r\theta}, \tag{14b}$$

$$Q^i_r + \frac{\chi_{,\theta}}{r} = \overline{Q}_r, \tag{14c}$$

for $r = a$, while the boundary conditions (10) may be written in the form

$$w = \overline{w}, \tag{15a}$$

$$-Dv_{,r} + \frac{2}{(1-\nu)\lambda^2}\frac{\chi_{,\theta}}{r} = D\overline{\phi}_r, \tag{15b}$$

$$-D\frac{v_{,\theta}}{r} - \frac{2}{(1-\nu)\lambda^2}\chi_{,r} = D\overline{\phi}_\theta. \tag{15c}$$

Having the systems (14) and (15) we now proceed to deduce from them a system of abbreviated relations, in such a way that *terms of relative order $1/\lambda a$ are retained, while terms of relative order $1/(\lambda a)^2$ are being disregarded.* To accomplish this reduction, it is necessary to stipulate at the outset a suitable order-of-magnitude relation between the dependent variables v and χ.

Inspection of the system (14) indicates that a reduction of this system is accomplished upon stipulating that

$$a^2\chi = O(Dv). \tag{16}$$

Therewith equations (14a,b) become, except for terms of relative order $1/(a\lambda)^2$,

$$M^i_{rr} + \frac{2}{\lambda^2}\frac{\chi_{,\theta r}}{r} = \overline{M}_{rr}, \quad M^i_{r\theta} + \frac{2}{\lambda^2}\frac{\chi_{,r}}{r} - \chi = \overline{M}_{r\theta}, \tag{17a,b}$$

with the relevant expressions for M^i_{rr}, $M^i_{r\theta}$, Q^i_r now involving w rather than v directly, as a consequence of equation (13), and with equation (14c) remaining unchanged.

The corresponding order-of-magnitude stipulation regarding v and χ for the system (15) is

$$a^2\chi = O(\lambda a\, Dv). \tag{18}$$

Therewith the χ-term in (15c) is of the same order of magnitude as the v-term, with the χ-contribution in (15b) now being of relative order $1/\lambda a$. At the same time, because of (13), equations (15b,c) may be written in terms of w rather than in terms of v, as follows:

$$-Dw_{,r} + \frac{2}{1-\nu}\frac{\chi_{,\theta}}{\lambda^2 r} = D\overline{\phi}_r, \quad -D\frac{w_{,\theta}}{r} - \frac{2}{1-\nu}\frac{\chi_{,r}}{\lambda^2} = D\overline{\phi}_\theta. \tag{19a,b}$$

Having the systems (17a,b) and (14c), and (19a,b) and (15a), we now use these for the derivation of equivalent systems which are of such nature as to allow a *sequential* determination of w and χ, with the w-problem being the desired generalization of Kirchhoff's problem in which first-order transverse shear deformation terms are taken into account of entirely by modification of Kirchhoff's boundary conditions.

In order to derive from the given systems of three boundary conditions for w and χ separate systems of two conditions for w and one condition for χ we make use of the differential equation $\nabla^2\chi - \lambda^2\chi = 0$ in the asymptotic form $\chi_{,rr} - \lambda^2\chi = 0$, from which it follows that $\chi = f(\theta)e^{-\lambda(r-a)}$ and therewith, except for terms of relative order $1/\lambda a$,

$$\chi_{,r} = -\lambda\chi. \tag{20}$$

The introduction of (20) into (17a,b) changes these relations into

$$M^i_{rr} - \frac{2}{\lambda a}\chi_{,\theta} = \overline{M}_{rr}, \quad M^i_{r\theta} - \left(\frac{2}{\lambda a} + 1\right)\chi = \overline{M}_{r\theta}, \tag{21a,b}$$

for $r = a$. Equation (21b) may be rewritten in the form

$$r = a, \quad \chi = \left(1 - \frac{2}{\lambda a}\right)(M^i_{r\theta} - \overline{M}_{r\theta}). \tag{22}$$

[4] The restatement of this expression involves a useful transformation with the help of the differential equation (7b).

A substitution of this in (21a) and (14c) then gives as modified Kirchhoff's boundary conditions, involving w alone

$$M^i_{rr} - \overline{M}_{rr} - \frac{2}{\lambda a}(M_{r\theta} - \overline{M}_{r\theta})_{,\theta} = 0, \tag{23a}$$

$$Q^i_r - \overline{Q}_r + \left(1 - \frac{2}{\lambda a}\right)\frac{(M^i_{r\theta} - \overline{M}_{r\theta})_{,\theta}}{a} = 0, \tag{23b}$$

for $r = a$.[5] It is evident from (23a,b) and (22), in conjunction with the differential equations (7a,b), that the asymptotic determination of w and χ up to terms of relative order $1/\lambda a$ is now in fact of a sequential nature. The previous result [3] for the case of a straight boundary follow from (23a,b) by first setting $()_{,\theta}/a = ()_{,2}$ and by then going to the limit $a \to \infty$, with the term $2/\lambda a$ in (23b) disappearing in this process.

The analogous reduction of the displacement boundary conditions (19a,b) comes out as follows. We first combine (20) and (19b) in the form

$$r = a, \quad \chi = \frac{1-\nu}{2} D\lambda \left(\frac{w_{,\theta}}{a} + \overline{\phi}_\theta\right), \tag{24}$$

and then use this relation in equation (19a) so as to obtain as second displacement boundary condition for w alone, in complementation of (15a),

$$(w_{,r} + \overline{\phi}_r) - \frac{1}{\lambda a}\left(\frac{w_{,\theta}}{a} + \overline{\phi}_\theta\right)_{,\theta} = 0, \tag{25}$$

for $r = a$. Equations (25), (24), and (15a) reduce directly to the corresponding conditions in [3] for the case of a straight boundary.

References

1 Goldenveiser, A. L., "The Principles of Reducing Three-Dimensional Problems of Elasticity to Two-Dimensional Problems of the Theory of Plates and Shells," *Proceedings of the 11th International Congress on Applied Mechanics*, 1966, pp. 306–311.

2 Reissner, E., "The Effect of Transverse Shear Deformation on the Bending of Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 12, 1945, pp. A69–A77.

3 Reissner, E., "On the Theory of Transverse Bending of Elastic Plates," *International Journal of Solids and Structures*, Vol. 12, 1976, pp. 545–554.

4 Simmonds, J. G., "Recent Advances in Shell Theory," *Proceedings of the 13th Annual Meeting Soc. Eng. Science*, 1976, pp. 617–626.

5 Timoshenko, S., and Woinowsky-Krieger, S., *Theory of Plates and Shells*, 1959, pp. 98–104.

[5] We note that for the case of a *three-dimensionally* isotropic plate, which is the case considered by Goldenveiser [1, 4], we have in the foregoing $2/\lambda a = (2/\sqrt{10})\, h/a$. The numerical factor $2/\sqrt{10} = 0.632\ldots$ corresponds to a factor $0.630\ldots$ in Goldenveiser's three-dimensional asymptotic analysis.

On the Stokes Flow of Viscous Fluids Through Corrugated Pipes

N. Phan-Thien[1]

Introduction

The Stokes flow problem of an incompressible viscous fluid in corrugated channels was recently considered by Wang [1, 2] who obtained solutions for flow parallel to the corrugations [1] and transverse to them [2]. In the latter case it was shown that, for sinusoidal corrugations of small amplitude ϵ, the mean pressure drop increased over the noncorrugated value by a factor of $1 + \epsilon^2 B(\lambda,\phi)$, where B is a function of the frequency (λ) and the phase shift (ϕ) of the corrugations in the channel walls. Numerical solutions to the same problem but allowing for inertia and diffusion effects were obtained by Chow and Soda [3] with a view of modeling blood oxygenators.

An approximate solution to the corresponding pipe flow problem was derived by Langlois [4] who used the lubrication approximation to calculate the mean pressure drop. The prediction of this simple method agrees well with the exact value provided that the pipe radius varies slowly, i.e., the frequency of the corrugations should be small. Since Langlois' [4] paper, this present pipe flow problem has been considered by a number of authors. Tanner [5] extended the perturbation analysis of Blassius to predict the kinetic energy losses of viscometric capillary tubes. However, as Manton [6] has pointed out, Tanner neglected second-order terms in the momentum equations so that the pressure was assumed to vary in the axial direction only. In applying the boundary condition (equation (2.6a) of [6])

$$\frac{\partial \psi}{\partial r} = 0 \quad \text{on the radius of the pipe,}$$

Manton [6] also neglected terms of first and second-order; the exact expression is given in equation (3) of this communication.

In view of the importance of this flow in Biomechanics [7] we attempt here to develop a perturbation solution to this problem. We also report the predictions of a lubrication approximation which is similar to that used by Langlois [4].

Analysis

A good starting point of the analysis is the Stokes flow equation in cylindrical coordinates (r, θ, z) for an incompressible viscous fluid

$$E^4\psi = 0, \quad E^2 \equiv r\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}, \tag{1}$$

where ψ is the Stokes dimensionless stream function.

The boundary conditions on ψ are

$$\text{At } r = 0, \quad \psi = 0, \quad \frac{1}{r}\frac{\partial \psi}{\partial z} = 0 \tag{2}$$

$$\text{At } r = 1 + \epsilon \sin \lambda z, \quad \psi = 1, \quad \frac{\partial \psi}{\partial p} = 0, \tag{3}$$

where p is the direction normal to the boundary of the pipe. Note that we have nondimensionalized all length variables with respect to a, the mean radius of the pipe.

We seek a perturbation solution of the form

$$\psi = \psi_0 + \epsilon\psi_1 + \epsilon^2\psi_2 + O(\epsilon^3) \tag{4}$$

for which the boundary conditions (3) become

$$\psi_0(1) + \epsilon\left(\psi_1(1) + \sin\lambda z\,\frac{\partial\psi_0}{\partial r}(1)\right) + \epsilon^2\left(\psi_2(1) + \sin\lambda z\,\frac{\partial\psi_1}{\partial r}(1) + \frac{1}{2}\sin^2\lambda z\,\frac{\partial^2\psi_0}{\partial r^2}(1)\right) + O(\epsilon^3) = 1 \tag{5}$$

and

$$\begin{aligned}\frac{\partial\psi_0}{\partial r}(1) &+ \epsilon\left(\frac{\partial\psi_1}{\partial r}(1) + \sin\lambda z\,\frac{\partial^2\psi_0}{\partial r^2}(1) - \lambda\cos\lambda z\,\frac{\partial\psi_0}{\partial z}(1)\right)\\ &+ \epsilon^2\left(\frac{\partial\psi_2}{\partial r}(1) + \sin\lambda z\,\frac{\partial^2\psi_1}{\partial r^2}(1) + \frac{1}{2}\sin^2\lambda z\,\frac{\partial^2\psi_0}{\partial r^3}(1)\right.\\ &- \frac{1}{2}\lambda^2\cos^2\lambda z\,\frac{\partial\psi_0}{\partial r}(1) - \lambda\cos\lambda z\,\frac{\partial\psi_1}{\partial z}(1)\\ &\left.- \lambda\sin\lambda z\cos\lambda z\,\frac{\partial^3\psi_0}{\partial z\,\partial r}(1)\right) + O(\epsilon^3) = 0\end{aligned} \tag{6}$$

where the argument 1 denotes an evaluation at $r = 1$.

[1] Lecturer, Department of Mechanical Engineering, University of Sydney, N.S.W., 2006, Australia.

Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, April, 1980.

A substitution of this in (21a) and (14c) then gives as modified Kirchhoff's boundary conditions, involving w alone

$$M^i_{rr} - \overline{M}_{rr} - \frac{2}{\lambda a}(M_{r\theta} - \overline{M}_{r\theta})_{,\theta} = 0, \tag{23a}$$

$$Q^i_r - \overline{Q}_r + \left(1 - \frac{2}{\lambda a}\right)\frac{(M^i_{r\theta} - \overline{M}_{r\theta})_{,\theta}}{a} = 0, \tag{23b}$$

for $r = a$.[5] It is evident from (23a,b) and (22), in conjunction with the differential equations (7a,b), that the asymptotic determination of w and χ up to terms of relative order $1/\lambda a$ is now in fact of a sequential nature. The previous result [3] for the case of a straight boundary follow from (23a,b) by first setting $()_{,\theta}/a = ()_{,2}$ and by then going to the limit $a \to \infty$, with the term $2/\lambda a$ in (23b) disappearing in this process.

The analogous reduction of the displacement boundary conditions (19a,b) comes out as follows. We first combine (20) and (19b) in the form

$$r = a, \quad \chi = \frac{1-\nu}{2} D\lambda \left(\frac{w_{,\theta}}{a} + \overline{\phi}_\theta\right), \tag{24}$$

and then use this relation in equation (19a) so as to obtain as second displacement boundary condition for w alone, in complementation of (15a),

$$(w_{,r} + \overline{\phi}_r) - \frac{1}{\lambda a}\left(\frac{w_{,\theta}}{a} + \overline{\phi}_\theta\right)_{,\theta} = 0, \tag{25}$$

for $r = a$. Equations (25), (24), and (15a) reduce directly to the corresponding conditions in [3] for the case of a straight boundary.

References

1 Goldenveiser, A. L., "The Principles of Reducing Three-Dimensional Problems of Elasticity to Two-Dimensional Problems of the Theory of Plates and Shells," *Proceedings of the 11th International Congress on Applied Mechanics*, 1966, pp. 306–311.

2 Reissner, E., "The Effect of Transverse Shear Deformation on the Bending of Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 12, 1945, pp. A69–A77.

3 Reissner, E., "On the Theory of Transverse Bending of Elastic Plates," *International Journal of Solids and Structures*, Vol. 12, 1976, pp. 545–554.

4 Simmonds, J. G., "Recent Advances in Shell Theory," *Proceedings of the 13th Annual Meeting Soc. Eng. Science*, 1976, pp. 617–626.

5 Timoshenko, S., and Woinowsky-Krieger, S., *Theory of Plates and Shells*, 1959, pp. 98–104.

[5] We note that for the case of a *three-dimensionally* isotropic plate, which is the case considered by Goldenveiser [1, 4], we have in the foregoing $2/\lambda a = (2/\sqrt{10})\, h/a$. The numerical factor $2/\sqrt{10} = 0.632\ldots$ corresponds to a factor $0.630\ldots$ in Goldenveiser's three-dimensional asymptotic analysis.

On the Stokes Flow of Viscous Fluids Through Corrugated Pipes

N. Phan-Thien[1]

Introduction

The Stokes flow problem of an incompressible viscous fluid in corrugated channels was recently considered by Wang [1, 2] who obtained solutions for flow parallel to the corrugations [1] and transverse to them [2]. In the latter case it was shown that, for sinusoidal corrugations of small amplitude ϵ, the mean pressure drop increased over the noncorrugated value by a factor of $1 + \epsilon^2 B(\lambda,\phi)$, where B is a function of the frequency (λ) and the phase shift (ϕ) of the corrugations in the channel walls. Numerical solutions to the same problem but allowing for inertia and diffusion effects were obtained by Chow and Soda [3] with a view of modeling blood oxygenators.

An approximate solution to the corresponding pipe flow problem was derived by Langlois [4] who used the lubrication approximation to calculate the mean pressure drop. The prediction of this simple method agrees well with the exact value provided that the pipe radius varies slowly, i.e., the frequency of the corrugations should be small. Since Langlois' [4] paper, this present pipe flow problem has been considered by a number of authors. Tanner [5] extended the perturbation analysis of Blassius to predict the kinetic energy losses of viscometric capillary tubes. However, as Manton [6] has pointed out, Tanner neglected second-order terms in the momentum equations so that the pressure was assumed to vary in the axial direction only. In applying the boundary condition (equation (2.6a) of [6])

$$\frac{\partial \psi}{\partial r} = 0 \quad \text{on the radius of the pipe,}$$

Manton [6] also neglected terms of first and second-order; the exact expression is given in equation (3) of this communication.

In view of the importance of this flow in Biomechanics [7] we attempt here to develop a perturbation solution to this problem. We also report the predictions of a lubrication approximation which is similar to that used by Langlois [4].

Analysis

A good starting point of the analysis is the Stokes flow equation in cylindrical coordinates (r, θ, z) for an incompressible viscous fluid

$$E^4\psi = 0, \quad E^2 \equiv r\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}, \tag{1}$$

where ψ is the Stokes dimensionless stream function.

The boundary conditions on ψ are

$$\text{At } r = 0, \quad \psi = 0, \quad \frac{1}{r}\frac{\partial \psi}{\partial z} = 0 \tag{2}$$

$$\text{At } r = 1 + \epsilon \sin \lambda z, \quad \psi = 1, \quad \frac{\partial \psi}{\partial p} = 0, \tag{3}$$

where p is the direction normal to the boundary of the pipe. Note that we have nondimensionalized all length variables with respect to a, the mean radius of the pipe.

We seek a perturbation solution of the form

$$\psi = \psi_0 + \epsilon\psi_1 + \epsilon^2\psi_2 + O(\epsilon^3) \tag{4}$$

for which the boundary conditions (3) become

$$\psi_0(1) + \epsilon\left(\psi_1(1) + \sin\lambda z\frac{\partial\psi_0}{\partial r}(1)\right) + \epsilon^2\left(\psi_2(1) + \sin\lambda z\frac{\partial\psi_1}{\partial r}(1) + \frac{1}{2}\sin^2\lambda z\frac{\partial^2\psi_0}{\partial r^2}(1)\right) + O(\epsilon^3) = 1 \tag{5}$$

and

$$\begin{aligned}\frac{\partial\psi_0}{\partial r}(1) &+ \epsilon\left(\frac{\partial\psi_1}{\partial r}(1) + \sin\lambda z\frac{\partial^2\psi_0}{\partial r^2}(1) - \lambda\cos\lambda z\frac{\partial\psi_0}{\partial z}(1)\right)\\ &+ \epsilon^2\left(\frac{\partial\psi_2}{\partial r}(1) + \sin\lambda z\frac{\partial^2\psi_1}{\partial r^2}(1) + \frac{1}{2}\sin^2\lambda z\frac{\partial^2\psi_0}{\partial r^3}(1)\right.\\ &- \frac{1}{2}\lambda^2\cos^2\lambda z\frac{\partial\psi_0}{\partial r}(1) - \lambda\cos\lambda z\frac{\partial\psi_1}{\partial z}(1)\\ &\left.- \lambda\sin\lambda z\cos\lambda z\frac{\partial^3\psi_0}{\partial z\partial r}(1)\right) + O(\epsilon^3) = 0\end{aligned} \tag{6}$$

where the argument 1 denotes an evaluation at $r = 1$.

[1] Lecturer, Department of Mechanical Engineering, University of Sydney, N.S.W., 2006, Australia.

Manuscript received by ASME Applied Mechanics Division, January, 1980; final revision, April, 1980.

The zeroth-order (ϵ^0) solution is the usual Poiseuille flow

$$\psi_0 = r^2(2 - r^2) \tag{7}$$

The ϵ^1-solution is governed by

$$E^4\psi_1 = 0$$

$$\psi_1(0) = 0, \quad \frac{1}{r}\frac{\partial\psi_1}{\partial z}(0) = 0$$

$$\psi_1(1) = 0, \quad \frac{\partial\psi_1}{\partial r}(1) = 8\sin\lambda z \tag{8}$$

In view of the boundary conditions, the solution to (8) is

$$\psi_1 = \phi_1(r)\sin\lambda z, \tag{9}$$

where

$$(D^2 - \lambda^2)^2\phi_1 = 0, \quad D^2 \equiv r\frac{d}{dr}\frac{1}{r}\frac{d}{dr} \tag{10, 11}$$

Rejecting the K_ν-solutions due to the boundedness of the velocity field, it can be verified that ϕ_1 is given by

$$\phi_1 = \alpha r I_1(\lambda r) + \beta r^2 I_0(\lambda r) \tag{12}$$

where $I_\nu(x)$, $K_\nu(x)$ are the Bessel functions of order ν.

After applying the boundary conditions (8*b–c*), we have

$$\psi_1(r,z) = 8\frac{r^2 I_1(\lambda)I_0(\lambda r) - rI_0(\lambda)I_1(\lambda r)}{\lambda[I_1{}^2(\lambda) - I_0(\lambda)I_2(\lambda)]}\sin\lambda z \tag{13}$$

Next, the ϵ^2-solution is governed by

$$E^4\psi_2 = 0,$$

$$\psi_2(0) = 0, \quad \frac{1}{r}\frac{\partial\psi_2}{\partial z}(0) = 0,$$

$$\psi_2(1) = -4\sin^2\lambda z, \quad \frac{\partial\psi_2}{\partial r}(1) = C\sin^2\lambda z,$$

$$C = 4 - 16I_1{}^2(\lambda)/[I_1{}^2(\lambda) - I_0(\lambda)I_2(\lambda)] \tag{14}$$

A solution of (14) is

$$\psi_2 = \phi_2(r)\sin^2\lambda z + \zeta_2(r)\cos^2\lambda z, \tag{15}$$

where

$$D^4(\phi_2 + \zeta_2) = 0$$

$$(D^2 - 4\lambda^2)^2(\phi_2 - \zeta_2) = 0 \tag{16}$$

Thus ϕ_2 and ζ_2 also involve $rI_1(2\lambda r)$ and $r^2I_0(2\lambda r)$ and can be written down easily. However, if one wishes to calculate the mean value then from (15) and (16*a*) we have

$$\langle\psi_2\rangle = \tfrac{1}{2}(\phi_2 + \zeta_2)$$

$$= \tfrac{1}{2}br^2 - Br^4, \tag{17}$$

where $\langle\cdot\rangle$ denotes the z-average of $(\cdot)$.

In effect, we have preaveraged the "biharmonic" equation (14*a*) as Wang did in his paper [2] for the two-dimensional case. However, in proceeding this way, full information about ψ_2 is also available if needed.

In (17) we have $b = -4 + 2B$ and

$$B = -3 + 4I_1{}^2(\lambda)/[I_1{}^2(\lambda) - I_0(\lambda)I_2(\lambda)] \tag{18}$$

The mean dimensional pressure gradient is given by

$$\left\langle\frac{\partial P}{\partial z}\right\rangle = -\frac{\eta Q}{2\pi a^4}\left\langle\frac{1}{r}\frac{\partial}{\partial r}E^2\psi\right\rangle$$

$$= -\frac{8\eta Q}{\pi a^4}(1 + \epsilon^2 B), \tag{19}$$

where Q is the flow rate and η is the fluid viscosity.

As sketched in Fig. 1, $B(\lambda)$ is always positive indicating an increase in the pressure drop. At low frequency λ, B is quadratic in λ and is given by $5 + \tfrac{2}{3}\lambda^2$. At large λ, $B(\lambda)$ tends asymptotically to 4λ which agrees qualitatively with the two-dimensional value (6λ) derived by Wang [2]. This dependence on λ is seen here primarily as an area effect; that is, the wetted area increases with λ and gives rise to an increase in the pressure drop.

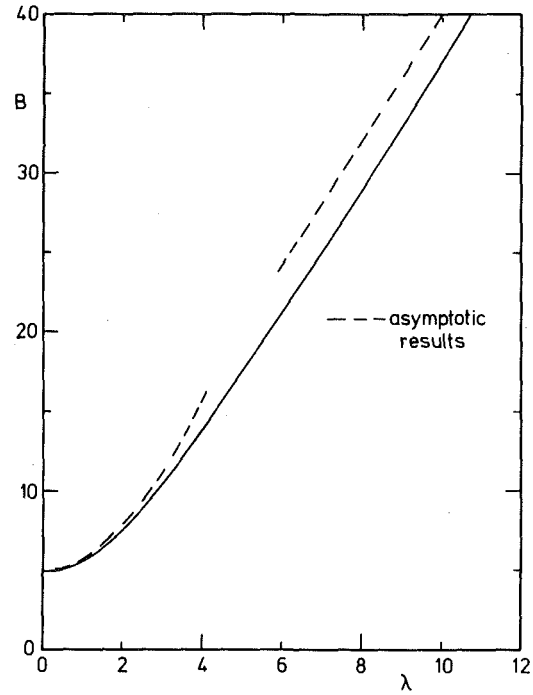


Fig. 1 $B(\lambda)$ **as a function of the frequency of the corrugations**

Note that $B(\lambda \to 0) \neq 0$ is an artifice of the method of average chosen here (i.e. $\langle\sin^2\lambda x\rangle = \tfrac{1}{2}$, all $\lambda \neq 0$) and by no means is it a contradiction to the Poiseuille result.

Lubrication Approximation

The essence of the lubrication argument [4] is that the axial velocity profile at any station z can be approximated by the Poiseuille profile

$$u = \frac{2Q}{\pi R^4}(R^2 - r^2), \tag{20}$$

where Q is the constant flow rate (which can be shown by integrating the continuity equation) and R is the radius of the tube

$$R = a[1 + \epsilon n(z)], \quad \epsilon \ll 1 \tag{21}$$

It should be noted that n can be allowed to depend on θ as well in which case $\langle\cdot\rangle$ denotes an average with respect to z and θ.

To obtain information about the pressure drop, we assume that P is essential uniform over the cross section of the pipe in which case we have

$$-\frac{\partial}{\partial z}(PR^2) = 2R\tau(1 + R'^2)^{1/2}, \tag{22}$$

where τ is the wall shear stress and $R' = \partial R/\partial z$. Here, we have neglected any normal stress that may arise.

It is not unreasonable to define an area-average pressure gradient by

$$\overline{\frac{\partial P}{\partial z}} = \frac{1}{\pi R^2}\frac{\partial}{\partial z}(\pi R^2 P)$$

from which (22) becomes

$$\overline{\frac{\partial P}{\partial z}} = \frac{2\tau}{R}(1 + R'^2)^{1/2} \tag{23}$$

When $R' = 0$ (noncorrugated pipe), (23) is seen to be the customary global balance of momentum.

For a Newtonian fluid, $\tau = \eta K$, where K is the shear rate at the wall,

$$K = -\frac{4Q}{\pi R^3},$$

we have

$$\overline{\frac{\partial P}{\partial z}} = -\frac{8Q\eta}{\pi a^4}[1 - 4\epsilon n + \epsilon^2(10n^2 + \tfrac{1}{2}n'^2) + O(\epsilon^3)] \tag{24}$$

Thus we can approximate the mean pressure gradient by

$$\left\langle\overline{\frac{\partial P}{\partial z}}\right\rangle = -\frac{8Q\eta}{\pi a^4}[1 + \epsilon^2(10\langle n^2\rangle + \tfrac{1}{2}\langle n'^2\rangle)] \tag{25}$$

For the sinusoidal corrugations (25) shows that the mean pressure drop increases over its noncorrugated value by a factor of $1 + (5 + \tfrac{1}{4}\lambda^2)\epsilon^2$. For low values of λ ($\lambda \leq 1$) this approximation agrees well with the exact value (at $\lambda = 1$ this underestimates the true pressure drop by 8 percent). This reasonably good agreement at low values of λ lends more confidence to the lubrication argument and we propose that (25) is a good estimate of the pressure drop in corrugated pipes carrying viscous fluids as long as the fundamental frequency of $n(z)$ is less than 1. This simple approach can be easily extended to the corresponding non-Newtonian problem once a constitutive equation for the fluid is nominated. However, since normal stresses may be significant, this approach may not give a good estimate of the pressure drop and we close the subject here.

Acknowledgment

I wish to thank Dr. B. L. Karihaloo for helpful discussions.

References

1 Wang, C-Y., "Parallel Flow Between Corrugated Plates," *ASCE Journal of the Engineering Mechanics Division*, Vol. 102, 1976, pp. 1088–1090.

2 Wang, C-Y., "On Stokes Flow Between Corrugated Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 462–464.

3 Chow, J. C. F., and Soda, K., "Laminar Flow and Blood Oxygeneration in Channels With Boundary Irregularities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, 1973, pp. 843–850.

4 Langlois, W. E., "Creeping Viscous Flow Through a Two-Dimensional Channel of Varying Gap," *Proceedings, 3rd National Congress of Applied Mechanics*, 1958, pp. 777–783.

5 Tanner, R. I., "Pressure Losses in Viscometric Capillary Tubes of Slowly Varying Diameter," *Journal of Applied Physics*, Vol. 17, 1966, pp. 633–670.

6 Manton, M. J., "Low Reynolds Number Flow in Slowly Varying Axisymmetric Tubes," *Journal of Fluid Mechanics*, Vol. 49, 1971, pp. 451–459.

7 Rodkiewicz, Cg. M., and Otto, W. J., "On the Newtonian Behaviour of Bile," *Journal of Biomechanics*, Vol. 12, 1979, pp. 609–612.

Dynamic Stress-Intensity Factors for Penny-Shaped Crack in Twisted Plate

B. M. Singh,[1], J. B. Haddow,[2] J. Vrbik,[3] and T. B. Moodie[4]

In this Note we determine the time-dependent stress-intensity factor due to the sudden formation of a penny-shaped crack at the middle surface of an elastic plate subjected to torsion about an axis normal to its faces. This is an extension of the work of Sih and Embley [1] who considered the sudden twisting of a penny-shaped crack in infinite elastic medium.

The analysis presented is applicable to the twisting of a thin disk whose faces are bonded to rigid dies and the method of solution provides an indication of how certain problems of dynamic fracture in a finite body can be analyzed.

Formulation of Problem

Stresses and displacements are referred to cylindrical polar coordinates (r, ϕ, z) and the usual notation is used.

The plate occupies the region $-h \leq z \leq h$, $0 \leq r \leq R$ and the penny-shaped crack is assumed to form suddenly and occupy the region $z = 0$, $a \geq r$. It is further assumed that $R \gg a$ and $R \gg h$.

Torsion of the plate may be achieved by the prescribed nonzero tangential displacement

$$u_\phi(r, \pm h, t) = \pm r\theta$$

on $z = \pm h$ which gives nonzero stress and displacement components

$$\tau_{z\phi} = \frac{\mu r\theta}{h}, \quad u_\phi = \frac{rz\theta}{h},$$

before the crack forms, where μ is the shear modulus and θ is constant. The stress-intensity factor is the same as for an equivalent problem in which the plate is initially undeformed, and at rest, and the stress

$$\tau_{z\phi} = -\frac{\mu r\theta}{h}H(t),$$

where $H(t)$ is the unit function, is applied to the faces of the crack with the faces $z = \pm h$ of the plate held fixed. Boundary conditions for the equivalent problem are

$$u_\phi(r, 0, t) = 0, \quad a \leq r \tag{1}$$

$$u_\phi(r, \pm h, t) = 0, \quad 0 \leq r \tag{2}$$

$$\tau_{z\phi}(r, 0, t) = -\frac{\tau r}{a}H(t), \quad 0 \leq r \leq a \tag{3}$$

where $\tau = \mu a\theta/h$.

Torsion of the plate may also be achieved by the nonzero shearing stress

$$\tau_{z\phi}(r, \pm h, t) = \frac{\tau r}{a},$$

on $z = \pm h$, where τ is constant and this gives nonzero stress and displacement components

$$\tau_{z\phi} = \frac{\tau r}{a}, \quad u_\phi = \frac{\tau rz}{a\mu}$$

before the crack forms. Again the stress-intensity factor is the same as for an equivalent problem. In the equivalent problem the stress

$$\tau_{\phi z} = -\frac{\tau r}{a}H(t),$$

is applied to the faces of the crack and the faces $z = \pm h$ are stress-free. Boundary conditions for this equivalent problem are

$$\tau_{z\phi}(r, 0, t) = -\frac{\tau r}{a}H(t), \quad 0 \leq r \leq a \tag{4}$$

$$u_\phi(r, 0, t) = 0, \quad a \leq r \tag{5}$$

$$\tau_{z\phi}(r, \pm h, t) = 0, \quad 0 \leq r \tag{6}$$

When boundary conditions (1)–(3) are applicable the problem is henceforth called Problem I and when (4)–(6) are applicable it is called Problem II.

The nonzero displacement component is u_ϕ and the nonzero stress components are given by

[1] 30/3 Gandhi Colony, Nazaftar Nagar, U.P., India.

[2] Professor, Mechanical Engineering Department, University of Alberta, Edmonton, Alberta, Canada T6G 2G8. Assoc. Mem. ASME.

[3] Computing Services Department, University of Calgary, Alberta, Canada.

[4] Mathematics Department, University of Alberta, Edmonton, Alberta, Canada T6G 2G8.

Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, June, 1980.

When $R' = 0$ (noncorrugated pipe), (23) is seen to be the customary global balance of momentum.

For a Newtonian fluid, $\tau = \eta K$, where K is the shear rate at the wall,

$$K = -\frac{4Q}{\pi R^3},$$

we have

$$\overline{\frac{\partial P}{\partial z}} = -\frac{8Q\eta}{\pi a^4}[1 - 4\epsilon n + \epsilon^2(10n^2 + \tfrac{1}{2}n'^2) + O(\epsilon^3)] \tag{24}$$

Thus we can approximate the mean pressure gradient by

$$\left\langle\overline{\frac{\partial P}{\partial z}}\right\rangle = -\frac{8Q\eta}{\pi a^4}[1 + \epsilon^2(10\langle n^2\rangle + \tfrac{1}{2}\langle n'^2\rangle)] \tag{25}$$

For the sinusoidal corrugations (25) shows that the mean pressure drop increases over its noncorrugated value by a factor of $1 + (5 + \frac{1}{4}\lambda^2)\epsilon^2$. For low values of λ ($\lambda \leq 1$) this approximation agrees well with the exact value (at $\lambda = 1$ this underestimates the true pressure drop by 8 percent). This reasonably good agreement at low values of λ lends more confidence to the lubrication argument and we propose that (25) is a good estimate of the pressure drop in corrugated pipes carrying viscous fluids as long as the fundamental frequency of $n(z)$ is less than 1. This simple approach can be easily extended to the corresponding non-Newtonian problem once a constitutive equation for the fluid is nominated. However, since normal stresses may be significant, this approach may not give a good estimate of the pressure drop and we close the subject here.

Acknowledgment

I wish to thank Dr. B. L. Karihaloo for helpful discussions.

References

1 Wang, C-Y., "Parallel Flow Between Corrugated Plates," *ASCE Journal of the Engineering Mechanics Division*, Vol. 102, 1976, pp. 1088–1090.

2 Wang, C-Y., "On Stokes Flow Between Corrugated Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 462–464.

3 Chow, J. C. F., and Soda, K., "Laminar Flow and Blood Oxygeneration in Channels With Boundary Irregularities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, 1973, pp. 843–850.

4 Langlois, W. E., "Creeping Viscous Flow Through a Two-Dimensional Channel of Varying Gap," *Proceedings, 3rd National Congress of Applied Mechanics*, 1958, pp. 777–783.

5 Tanner, R. I., "Pressure Losses in Viscometric Capillary Tubes of Slowly Varying Diameter," *Journal of Applied Physics*, Vol. 17, 1966, pp. 633–670.

6 Manton, M. J., "Low Reynolds Number Flow in Slowly Varying Axisymmetric Tubes," *Journal of Fluid Mechanics*, Vol. 49, 1971, pp. 451–459.

7 Rodkiewicz, Cg. M., and Otto, W. J., "On the Newtonian Behaviour of Bile," *Journal of Biomechanics*, Vol. 12, 1979, pp. 609–612.

Dynamic Stress-Intensity Factors for Penny-Shaped Crack in Twisted Plate

B. M. Singh,[1], J. B. Haddow,[2] J. Vrbik,[3] and T. B. Moodie[4]

In this Note we determine the time-dependent stress-intensity factor due to the sudden formation of a penny-shaped crack at the middle surface of an elastic plate subjected to torsion about an axis normal to its faces. This is an extension of the work of Sih and Embley [1] who considered the sudden twisting of a penny-shaped crack in infinite elastic medium.

The analysis presented is applicable to the twisting of a thin disk whose faces are bonded to rigid dies and the method of solution provides an indication of how certain problems of dynamic fracture in a finite body can be analyzed.

Formulation of Problem

Stresses and displacements are referred to cylindrical polar coordinates (r, ϕ, z) and the usual notation is used.

The plate occupies the region $-h \leq z \leq h$, $0 \leq r \leq R$ and the penny-shaped crack is assumed to form suddenly and occupy the region $z = 0$, $a \geq r$. It is further assumed that $R \gg a$ and $R \gg h$.

Torsion of the plate may be achieved by the prescribed nonzero tangential displacement

$$u_\phi(r, \pm h, t) = \pm r\theta$$

on $z = \pm h$ which gives nonzero stress and displacement components

$$\tau_{z\phi} = \frac{\mu r\theta}{h}, \quad u_\phi = \frac{rz\theta}{h},$$

before the crack forms, where μ is the shear modulus and θ is constant. The stress-intensity factor is the same as for an equivalent problem in which the plate is initially undeformed, and at rest, and the stress

$$\tau_{z\phi} = -\frac{\mu r\theta}{h}H(t),$$

where $H(t)$ is the unit function, is applied to the faces of the crack with the faces $z = \pm h$ of the plate held fixed. Boundary conditions for the equivalent problem are

$$u_\phi(r, 0, t) = 0, \quad a \leq r \tag{1}$$

$$u_\phi(r, \pm h, t) = 0, \quad 0 \leq r \tag{2}$$

$$\tau_{z\phi}(r, 0, t) = -\frac{\tau r}{a}H(t), \quad 0 \leq r \leq a \tag{3}$$

where $\tau = \mu a\theta/h$.

Torsion of the plate may also be achieved by the nonzero shearing stress

$$\tau_{z\phi}(r, \pm h, t) = \frac{\tau r}{a},$$

on $z = \pm h$, where τ is constant and this gives nonzero stress and displacement components

$$\tau_{z\phi} = \frac{\tau r}{a}, \quad u_\phi = \frac{\tau rz}{a\mu}$$

before the crack forms. Again the stress-intensity factor is the same as for an equivalent problem. In the equivalent problem the stress

$$\tau_{\phi z} = -\frac{\tau r}{a}H(t),$$

is applied to the faces of the crack and the faces $z = \pm h$ are stress-free. Boundary conditions for this equivalent problem are

$$\tau_{z\phi}(r, 0, t) = -\frac{\tau r}{a}H(t), \quad 0 \leq r \leq a \tag{4}$$

$$u_\phi(r, 0, t) = 0, \quad a \leq r \tag{5}$$

$$\tau_{z\phi}(r, \pm h, t) = 0, \quad 0 \leq r \tag{6}$$

When boundary conditions (1)–(3) are applicable the problem is henceforth called Problem I and when (4)–(6) are applicable it is called Problem II.

The nonzero displacement component is u_ϕ and the nonzero stress components are given by

[1] 30/3 Gandhi Colony, Nazaftar Nagar, U.P., India.

[2] Professor, Mechanical Engineering Department, University of Alberta, Edmonton, Alberta, Canada T6G 2G8. Assoc. Mem. ASME.

[3] Computing Services Department, University of Calgary, Alberta, Canada.

[4] Mathematics Department, University of Alberta, Edmonton, Alberta, Canada T6G 2G8.

Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, June, 1980.

$$\tau_{r\phi} = \mu\left(\frac{\partial u_\phi}{\partial r} - \frac{u_\phi}{r}\right),$$

$$\tau_{z\phi} = \frac{\partial u_\phi}{\partial z} \tag{7}$$

for both problems, and u_ϕ satisfies the equation of motion,

$$\frac{\partial^2 u_\phi}{\partial r^2} + \frac{1}{r}\frac{\partial u_\phi}{\partial r} - \frac{u_\phi}{r^2} + \frac{\partial^2 u_\phi}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 u_\phi}{\partial t^2}, \tag{8}$$

where $c^2 = \mu/\rho$ and ρ is the material density.

We adopt the usual definition of the Laplace transform with respect to the variable t and with this definition the solution of equation (8) in the Laplace transform domain is

$$\bar{u}_\phi(r, z, p) = \int_0^\infty [A(s, p)\exp(-\gamma z) + B(s, p)\exp(\gamma z)]J_1(sr)ds \tag{9}$$

where $J_\nu(\ )$ denotes a Bessel function of the first kind and order ν, a superposed bar denotes the Laplace transform,

$$\gamma = (s^2 + p^2/c^2)^{1/2},$$

and A and B are functions to be determined.

Since the problem is symmetric about $z = 0$ we consider the region $0 \le z \le h$.

The stress-intensity factor k is defined as

$$k_{1,2}(t) = \lim_{r\to a^+} \{2(r-a)\}^{1/2}\tau_{z\phi}(r, 0, t) \tag{10}$$

where the subscript 1 is used for Problem I and subscript 2 for Problem II.

Problem I

Equations (1)–(3) become

$$\bar{u}_\phi(r, 0, p) = 0, \quad a \le r \tag{11}$$

$$\bar{u}_\phi(r, h, p) = 0, \quad 0 \le r \tag{12}$$

$$\bar{\tau}_{z\phi}(r, 0, p) = \frac{\tau r}{ap}, \quad 0 \le r \le a \tag{13}$$

in the Laplace transform domain. Condition (12) is satisfied if

$$B(s, p) = -A(s, p)\exp(-2\gamma h) \tag{14}$$

where A and B are the functions which appear in equation (9). Conditions (11) and (13) are then satisfied if

$$\int_0^\infty D(s, p)\,\gamma\coth(\gamma h)J_1(rs)ds = \frac{\tau r}{a\mu p}, \quad 0 \le r \le a, \tag{15}$$

$$\int_0^\infty D(s, p)J_1(rs)ds = 0, \quad a \le r, \tag{16}$$

where

$$D(s, p) = [1\text{-}\exp(-2\gamma h)]A(sp). \tag{17}$$

The solution of the dual integral equations (15) and (16) can be obtained by using the method described by Copson [2] and is

$$D(s, p) = \frac{4\tau a^{5/2}s^{1/2}}{3\mu p(2\pi)^{1/2}}\int_0^1 \xi^{1/2}\bar{\psi}(\xi, \hat{p})J_{3/2}(\xi as)d\xi, \tag{18}$$

where $\bar{\psi}(\xi, \hat{p})$ satisfies the following Fredholm integral equation of the second kind:

$$\bar{\psi}(\xi, \hat{p}) + \int_0^1 \bar{\psi}(u, \hat{p})K_1(u, \hat{p}, \xi)du = \xi^2, \quad 0 \le \xi \le 1, \tag{19}$$

$$K_1(u, \hat{p}, \xi) = (\xi u)^{1/2}\int_0^\infty \{(\hat{s}^2 + \hat{p}^2)^{1/2}\coth[(\hat{s}^2 + \hat{p}^2)^{1/2}/\epsilon] - \hat{s}\} \times J_{3/2}(u\hat{s})J_{3/2}(\xi\hat{s})d\hat{s}, \tag{20}$$

and

$$\hat{p} = \frac{pa}{c}, \quad \epsilon = \frac{a}{h}.$$

It follows from equations (7) and (9) that

$$\bar{\tau}_{z\phi}(r, 0, p) = -\mu\left\{\int_0^\infty sD(s, p)J_1(rs)ds + \int_0^\infty [\gamma\coth(\gamma h) - s]D(s, p)J_1(rs)ds\right\}, \quad 0 \le r < \infty$$

and with the use of equation (18) this becomes

$$\bar{\tau}_{z\phi}(r, 0, p) = \frac{2\mu}{\pi p}\left[\frac{2a^2\bar{\psi}(1, \hat{p})}{3r(r^2 - a^2)^{1/2}} - \frac{2a}{3r}\int_0^a \frac{\bar{\psi}'\left(\frac{\eta}{a}, \hat{p}\right)d\eta}{(r^2 - \eta^2)^{1/2}}\right] - \frac{2a\mu}{3p}\left(\frac{2}{\pi}\right)^{1/2}\left\{\int_0^a u^{1/2}\bar{\psi}\left(\frac{u}{a}, p\right)du\int_0^\infty s^{1/2} \times [\gamma\coth(\gamma h) - s]J_1(rs)J_{3/2}(us)ds\right\}, \quad a < r, \tag{21}$$

where the prime denotes differentiation with respect to η.

The stress-intensity factor in the Laplace transform domain is given by

$$\bar{k}_1 = \lim_{r\to a^+}\{2(r-a)\}^{1/2}\bar{\tau}_{z\phi}(r, 0, p). \tag{22}$$

It follows from equations (21) and (22) that the stress-intensity factor in the nondimensional Laplace transform domain is given by

$$\bar{k}_1(\hat{p}) = \frac{4\tau a^{1/2}}{3\pi\hat{p}}\bar{\psi}(1, p),$$

consequently the stress-intensity factor as a function of the nondimensional time $\hat{t} = ct/a$ is given by

$$\frac{3\pi k_1(\hat{t})}{4\tau a^{1/2}} = L^{-1}\left\{\frac{\bar{\psi}(1, \hat{p})}{\hat{p}}\right\}, \tag{23}$$

where L^{-1} denotes the inverse Laplace transform from the $\hat{p}$ domain to the $\hat{t}$ domain.

Problem II

Problem II is solved by the same method as for Problem I, except that boundary conditions (4)–(6) are used, to obtain

$$\frac{3k_2(\hat{t})}{4\tau a^{1/2}} = L^{-1}\left\{\frac{F(1, \hat{p})}{\hat{p}}\right\}, \tag{24}$$

where the function $F(\xi, \hat{p})$ satisfies the Fredholm integral equation

$$F(\xi, \hat{p}) + \int_0^1 F(u, \hat{p})K(u, \hat{p}, \xi)du = \xi^2, \quad 0 \le \xi \le 1 \tag{25}$$

with

$$K(u, \hat{p}, \xi) = (\xi u)^{1/2}\int_0^\infty \{(\hat{s}^2 + \hat{p}^2)^{1/2}\tanh[(\hat{s}^2 + \hat{p}^2)^{1/2}/\epsilon] - \hat{s}\} \times J_{3/2}(u\hat{s})J_{3/2}(\xi\hat{s})d\hat{s}. \tag{26}$$

Numerical Results

In order to obtain the relation between the nondimensional stress-intensity factor $3\pi k_{1/2}(\hat{t})/(4\tau a^{1/2})$ and nondimensional time $\hat{t}$ an integral equation, equation (19) for Problem I and equation (25), for Problem II, is solved numerically. Discrete values of the function $\psi(1, \hat{p})$ and $F(1, p)$ for Problems I and II, respectively, are thus obtained and the Laplace transform inversions indicated in equations (23) and (24) are then done numerically using the method of Miller and Guy [3]. The integrals appearing in the kernals (20) and (26) pose a numerical difficulty due to the slowly decaying integrands. To resolve this problem the part of each integrand in curly brackets is expanded as a power series in $1/\hat{s}$ and the contributions of $O(\hat{s}^{-1})$ and $O(\hat{s}^{-3})$ are evaluated analytically. The remaining function converges

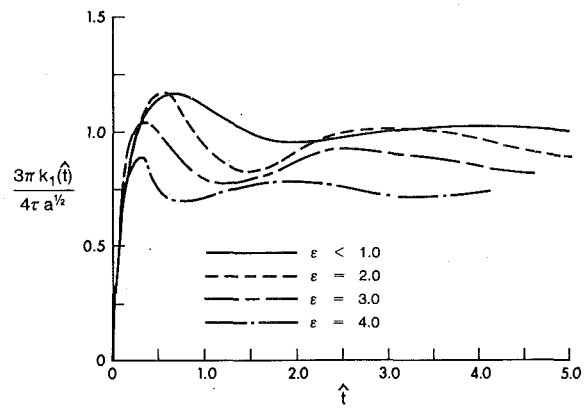


Fig. 1 Stress-intensity factors for Problem I

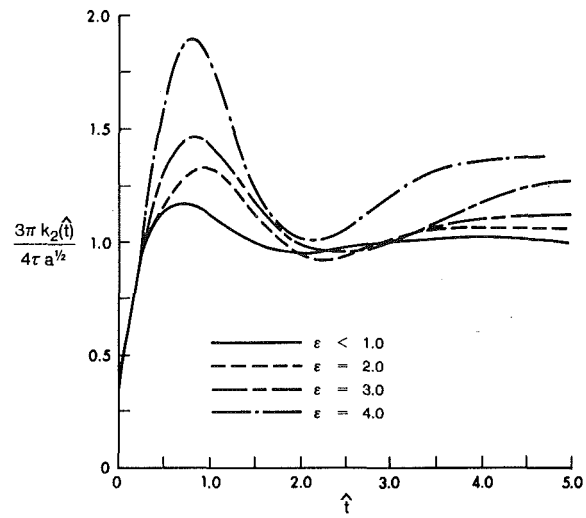


Fig. 2 Stress-intensity factors for Problem II

much more rapidly, and is evaluated by adding contributions between successive zeros of $J_{3/2}(u\hat{s})$ and $J_{3/2}(\xi\hat{s})$.

Results are shown graphically in Figs. 1 and 2 for different values of ϵ. For $\epsilon < 1.0$ the results for the two problems are almost identical and independent of ϵ. The time-dependent stress-intensity factor rises rapidly from zero in each case and then has the form of a damped oscillation about the static value. Results for $\epsilon > 1.0$ are quantitatively different for the two problems since the limiting value of the stress-intensity factor, as $t \to \infty$, decreased with increasing ϵ for Problem I and increases for Problem II.

References

1 Sih, G. C., and Embly, G. T., "Sudden Twisting of a Penny-Shaped Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, 1972, pp. 395–400.

2 Copson, E. T., "On Certain Dual Integral Equations," *Proceedings of the Glasgow Mathematical Association*, Vol. 5, 1961, pp. 21–24.

3 Miller, M. K., and Guy, W. T., "Numerical Inversion of the Laplace Transform by Use of Jacobi Polynomials," *SIAM Journal of Numerical Analysis*, Vol. 3, 1966, pp. 624–635.

Heat Transfer Due to the Flow Between Infinite Plates—One Rotating and the Other at Rest Under Transverse Magnetic Field

A. K. Borkakati[1]

The magnetohydrodynamic flow between two infinite parallel plates, one rotating and the other at rest, has been analyzed by Srivastava and Sharma [1], who have imposed the condition that the Reynolds number $R \ll$ the Hartmann number M^2. Stephenson [2] has pointed out that in finding out the value of radial current, the foregoing workers have omitted one term with the results that (*i*) their expansions are valid only if the stationary disk is a perfect conductor, or has infinite thickness (and is not an insulator) and (*ii*) their expressions are completely erroneous. He has calculated the modified expressions of the velocity components.

We consider the motion of an incompressible electrically conducting viscous fluid between two infinite parallel disks at a distance d apart, one of which ($z = 0$) is rotating with constant angular velocity about an axis ($r = 0$) pependicular to the plates and the other ($z = d$) is at rest. A transverse magnetic field $\mathbf{B}_0$ is imposed perpendicular to the plates. The induced magnetic field being small in comparison to the imposed magnetic field, is neglected which is valid for small magnetic Reynolds number. The rotating plate is maintained at a temperature T_0 while the other plate is kept at $T_1 (T_1 > T_0)$. The boundary conditions are

$$T = T_0, \quad u_r = u_z = 0, \quad u_\theta = r\Omega \quad \text{at} \quad z = 0,$$
$$T = T_1, \quad u_r = u_\theta = u_z = 0 \quad \text{at } z = d, \tag{1}$$

where u_r, u_θ, u_z are the velocity components in the directions of r, θ, z, respectively.

The equations of motion and continuity are same with those of Stephenson [2]. The energy equation is

$$\rho C_p \left[u_r \frac{\partial T}{\partial r} + u_z \frac{\partial T}{\partial z}\right] = k \left[\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial z^2}\right] + \phi + \frac{\mathbf{J}^2}{\sigma}, \tag{2}$$

where C_p is the specific heat at constant pressure, k is the thermometric conductivity, σ is the electrical conductivity, and $\mathbf{J} = (J_r, J_\theta, J_z)$ is the current density.

The dissipation function ϕ is

$$\phi = 2\rho\nu \left[\left(\frac{\partial u_r}{\partial r}\right)^2 + \frac{u_r^2}{r^2} + \left(\frac{\partial u_z}{\partial z}\right)^2 + \frac{1}{2}\left(\frac{\partial u_\theta}{\partial z}\right)^2 + \frac{1}{2}\left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r}\right)^2 + \frac{1}{2}\left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r}\right)^2\right] \tag{3}$$

Let us consider the following similarity solutions:

$$u_r = r\partial W F(\eta),$$
$$u_\theta = r\Omega G(\eta),$$
$$u_z = d\Omega H(\eta),$$
$$T = T_0 + \frac{\nu\Omega}{C_p}\left[\nu(\eta) + \frac{r^2}{d^2}\psi(\eta)\right],$$
$$p = \rho\nu\Omega\left[-P_1(\eta) + \frac{\lambda r^2}{d^2}\right],$$
$$\lambda = \text{constant}, \tag{4}$$

where $\eta = z/d$ = dimensionless axial distance.

[1] Department of Mathematics, Dibrugarh University, Dibrugarh 786004 (Assam) India.

Manuscript received by ASME Applied Mechanics Division, June, 1978; final revision, June, 1980.

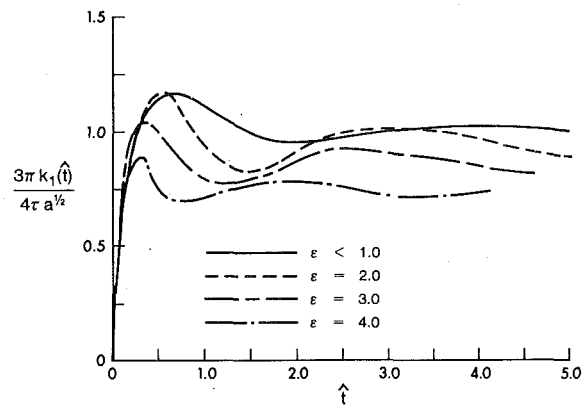


Fig. 1 Stress-intensity factors for Problem I

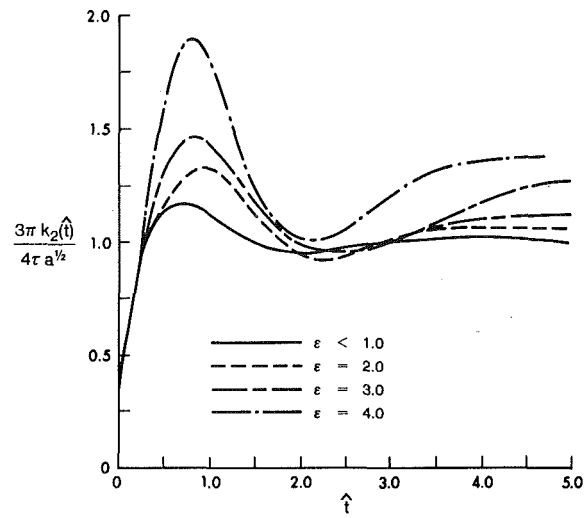


Fig. 2 Stress-intensity factors for Problem II

much more rapidly, and is evaluated by adding contributions between successive zeros of $J_{3/2}(u\hat{s})$ and $J_{3/2}(\xi\hat{s})$.

Results are shown graphically in Figs. 1 and 2 for different values of ϵ. For $\epsilon < 1.0$ the results for the two problems are almost identical and independent of ϵ. The time-dependent stress-intensity factor rises rapidly from zero in each case and then has the form of a damped oscillation about the static value. Results for $\epsilon > 1.0$ are quantitatively different for the two problems since the limiting value of the stress-intensity factor, as $t \to \infty$, decreased with increasing ϵ for Problem I and increases for Problem II.

References

1 Sih, G. C., and Embly, G. T., "Sudden Twisting of a Penny-Shaped Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, 1972, pp. 395–400.

2 Copson, E. T., "On Certain Dual Integral Equations," *Proceedings of the Glasgow Mathematical Association*, Vol. 5, 1961, pp. 21–24.

3 Miller, M. K., and Guy, W. T., "Numerical Inversion of the Laplace Transform by Use of Jacobi Polynomials," *SIAM Journal of Numerical Analysis*, Vol. 3, 1966, pp. 624–635.

Heat Transfer Due to the Flow Between Infinite Plates—One Rotating and the Other at Rest Under Transverse Magnetic Field

A. K. Borkakati[1]

The magnetohydrodynamic flow between two infinite parallel plates, one rotating and the other at rest, has been analyzed by Srivastava and Sharma [1], who have imposed the condition that the Reynolds number $R \ll$ the Hartmann number M^2. Stephenson [2] has pointed out that in finding out the value of radial current, the foregoing workers have omitted one term with the results that (*i*) their expansions are valid only if the stationary disk is a perfect conductor, or has infinite thickness (and is not an insulator) and (*ii*) their expressions are completely erroneous. He has calculated the modified expressions of the velocity components.

We consider the motion of an incompressible electrically conducting viscous fluid between two infinite parallel disks at a distance d apart, one of which ($z = 0$) is rotating with constant angular velocity about an axis ($r = 0$) pependicular to the plates and the other ($z = d$) is at rest. A transverse magnetic field $\mathbf{B}_0$ is imposed perpendicular to the plates. The induced magnetic field being small in comparison to the imposed magnetic field, is neglected which is valid for small magnetic Reynolds number. The rotating plate is maintained at a temperature T_0 while the other plate is kept at $T_1 (T_1 > T_0)$. The boundary conditions are

$$T = T_0, \quad u_r = u_z = 0, \quad u_\theta = r\Omega \quad \text{at} \quad z = 0,$$
$$T = T_1, \quad u_r = u_\theta = u_z = 0 \quad \text{at } z = d, \tag{1}$$

where u_r, u_θ, u_z are the velocity components in the directions of r, θ, z, respectively.

The equations of motion and continuity are same with those of Stephenson [2]. The energy equation is

$$\rho C_p \left[u_r \frac{\partial T}{\partial r} + u_z \frac{\partial T}{\partial z}\right] = k \left[\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial z^2}\right] + \phi + \frac{\mathbf{J}^2}{\sigma}, \tag{2}$$

where C_p is the specific heat at constant pressure, k is the thermometric conductivity, σ is the electrical conductivity, and $\mathbf{J} = (J_r, J_\theta, J_z)$ is the current density.

The dissipation function ϕ is

$$\phi = 2\rho\nu \left[\left(\frac{\partial u_r}{\partial r}\right)^2 + \frac{u_r^2}{r^2} + \left(\frac{\partial u_z}{\partial z}\right)^2 + \frac{1}{2}\left(\frac{\partial u_\theta}{\partial z}\right)^2 + \frac{1}{2}\left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r}\right)^2 + \frac{1}{2}\left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r}\right)^2\right] \tag{3}$$

Let us consider the following similarity solutions:

$$u_r = r\partial W F(\eta),$$
$$u_\theta = r\Omega G(\eta),$$
$$u_z = d\Omega H(\eta),$$
$$T = T_0 + \frac{\nu\Omega}{C_p}\left[\nu(\eta) + \frac{r^2}{d^2}\psi(\eta)\right],$$
$$p = \rho\nu\Omega\left[-P_1(\eta) + \frac{\lambda r^2}{d^2}\right],$$
$$\lambda = \text{constant}, \tag{4}$$

where $\eta = z/d$ = dimensionless axial distance.

[1] Department of Mathematics, Dibrugarh University, Dibrugarh 786004 (Assam) India.

Manuscript received by ASME Applied Mechanics Division, June, 1978; final revision, June, 1980.

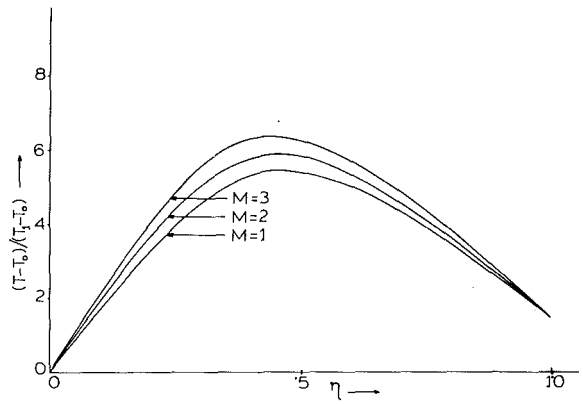


Fig. 1 Graph of $(T - T_0)/(T_1 - T_0)$ for $EPX^2 = 20$

Imposing axial symmetry, and using curl $\mathbf{E} = 0$, we get

$$E_\theta = 0 \text{ everywhere}, \tag{5}$$

where $\mathbf{E} = (E_r, E_\theta, E_z)$ is the electric field.

Neglecting induced magnetic field, we get

$$J_\theta = -\sigma B_0 u_r = -\sigma B_0 \Omega r F(\eta),$$
$$J_z = \sigma E_z, \tag{6}$$

where B_0 is the magnitude of the imposed magnetic field.

The tangential equation of motion may be written as

$$\frac{\nu}{r^2}\frac{\partial^2}{\partial r^2}(u_\theta) - \left\{\frac{2}{r^2}u_r u_\theta + \frac{u_z}{r}\frac{\partial}{\partial z}(u_\theta)\right\} = \frac{B_0 J_r}{r}, \tag{7}$$

which with (4), shows that

$$J_r = \sigma[E_r - B_0 u_\theta] = r \times \text{function of } z \tag{8}$$

Using div $\mathbf{J} = 0$, we get

$$\frac{\partial J_z}{\partial z} = -\frac{1}{r}\frac{\partial}{\partial r}(rJ_r) = \text{function of } z \tag{9}$$

Therefore

$$J_z = \sigma E_z = \text{function of } r + \text{function of } z \tag{10}$$

Now we consider a surface where a conducting medium adjoins a nonconducting medium; such surfaces exist (*a*) between a conducting fluid and an insulating disk and (*b*) between a conducting disk if present and the environment. It has been seen that there will always be at least one surface of this kind. At such a surface, we have $J_z = 0$ and therefore in the conducting medium

$$J_z = \sigma E_z = \text{function of } z \tag{11}$$

From curl $\mathbf{E} = 0$, we obtain

$$\frac{\partial}{\partial r}(E_z) = \frac{\partial}{\partial z}(E_r), \tag{12}$$

which gives, E_r = function of r. Hence, with (8),

$$E_r = r \times \text{constant} \tag{13}$$

Now, we have

$$E_r = r \times \text{constant} = -\chi B_0 \Omega r \quad \text{(say)}, \tag{14}$$
$$E_\theta = 0, \tag{15}$$
$$E_z = \text{function of } z, \tag{16}$$
$$J_r = r\Omega\sigma B_0[G(\eta) - \chi], \tag{17}$$
$$J_\theta = -\sigma B_0 r \Omega F(\eta), \tag{18}$$
$$J_z = \sigma E_z = \text{function of } z \tag{19}$$

where χ is the constant to be determined.

Integrating (17), and substituting $\int_0^z J_r dz = 0$ we get

$$\int_0^1 G d\eta = \chi. \tag{20}$$

Again, integrating equation (9), we have

$$J_z = -2\sigma B_0 \Omega (L - K), \tag{21}$$

where

$$L = \int_0^\eta (G - \chi) d\eta = \text{function of } \eta$$

and

$$K = \int_0^1 L d\eta = \text{constant}. \tag{22}$$

Substituting (4) in the equations of motion, continuity and energy, we get

$$F'' = R[F^2 - G^2 + F'H + \lambda] + M^2 F,$$
$$G'' = R[2FG + G'H] + M^2[G - \chi],$$
$$H' + 2F = 0$$
$$R[H\vartheta' - 4F^2 - 2H'^2 - 4M^2(L - K)^2] = \frac{1}{P}[4\psi + \vartheta''],$$
$$R[2F\psi + H\psi' - (F'^2 + G'^2) - M^2\{(G - \chi)^2 + F^2\}] = \frac{1}{P}\psi'' \tag{23}$$

where $R(= \Omega d^2/\nu)$ is the Reynolds number, $M = B_0 d(\sigma/\rho\nu)^{1/2}$ is Hartmann number and $P = C_p \nu\rho/k$ is the Prandtl number.

If we consider $R \ll M^2$ and substitute $N = M^2/R$, the unknown quantities may be formally expanded in powers of N as follows:

$$\alpha = \sum_{n=0}^{\infty} \alpha_n / N^n, \tag{24}$$

where α can be replaced by $F, G, H, \vartheta, \psi, \lambda, \chi$.

Substituting these expressions in (23) and equating the coefficients of like powers of N from both sides of the equations, we get linear differential equations in

$$f_0, g_0, h_0, \psi, \vartheta_0, f_1, g_1, h_1, \psi_1, \vartheta_1, \text{ etc.}$$

This method of expansion has been adopted by Srivastava in solving similar problems in Reiner-Rivlin fluid [3], in electrically conducting fluid [1] as well as in second-order fluid [4].

Boundary conditions reduces to

$$\left.\begin{aligned} &\left.\begin{aligned} f_0 = h_0 = \vartheta_0 = \psi_0 = \vartheta_1 = \psi_1 = 0, \quad g_0 = 1 \\ f_1 = g_1 = h_1 = 0, \end{aligned}\right\} \text{at } \eta = 0 \\ &\left.\begin{aligned} f_0 = g_0 = h_0 = f_1 = g_1 = h_1 = \psi_0 = \vartheta_1 = \psi_1 = 0 \\ \vartheta_0 = S \end{aligned}\right\} \text{at } \eta = 1 \end{aligned}\right] \tag{25}$$

where

$$S = \frac{(T_1 - T_0)C_p}{\nu\Omega} = \text{constant}.$$

Solving these equations subject to the boundary conditions (25), $f_0, h_0, g_0, f_1, h_1, g_1, \psi_0, \psi_1, \vartheta_0, \vartheta_1$ are determined. Substituting them in

$$\frac{T - T_0}{T_1 - T_0} = \frac{1}{S}\left[\vartheta_0 + \frac{1}{N}\vartheta_1 + \frac{r^2}{\alpha^2}\left(\psi_0 + \frac{1}{N}\psi_1\right)\right] \tag{26}$$

We can get the expression for the dimensionless temperature distribution which, for moderate distances from the axis of rotation, reduces to

$$\frac{T - T_0}{T_1 - T_0} = \eta - \frac{\rho E X^2}{16 \sinh^2 M}[\cosh 2M(1 - \eta) + \cosh 2M\eta + 2\cosh M(1 - 2\eta) - \cos 2M - 2\cosh M - 1] \tag{27}$$

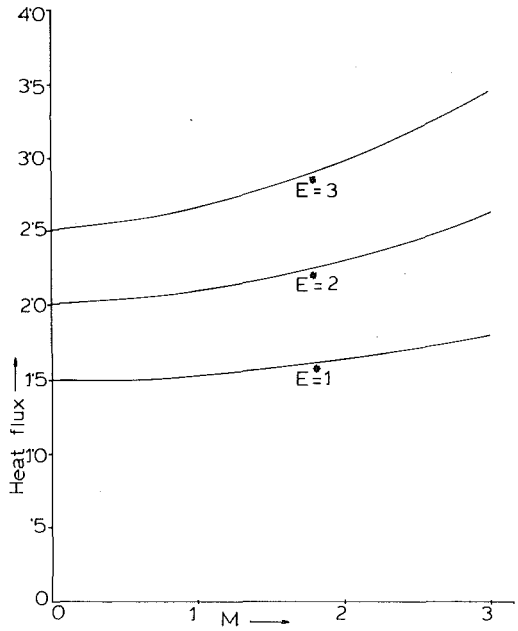


Fig. 2 Heat flux from the lower plate

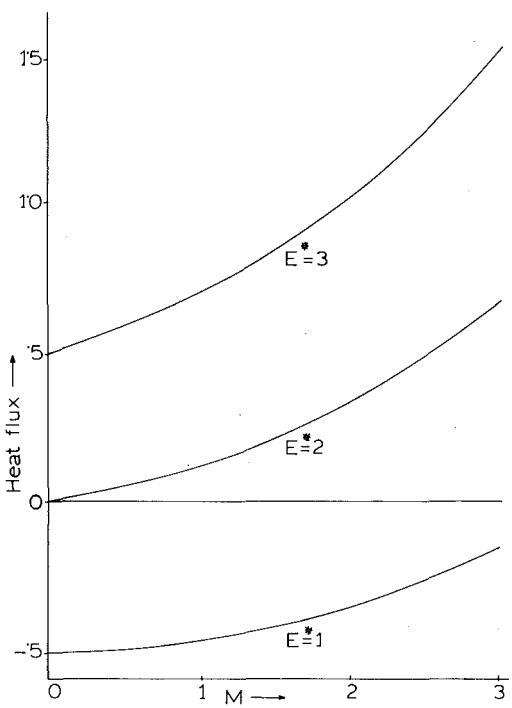


Fig. 3 Heat flux from the upper plate

where

$$X = r/d, \text{ and } E = R/S.$$

The variations of $(T - T_0/T_1 - T_0)$ against $\eta = z/d$ have been plotted graphically for $M = 1, 2, 3$ when $PEX^2 = 20$. From the graphs, it has been observed that the plane of maximum temperature is at the mid-distance between the plates. As Hartmann number goes on increasing the maximum temperature between the plates goes on increasing.

Srivastava [4] has studied the effects of nonlinear terms on the temperature distribution in non-Newtonian fluids. From the results obtained here, it has been observed that the effect of magnetic field in electrically conducting fluid and that of nonlinear terms in non-Newtonian fluids, on the temperature distribution are qualitatively similar.

The heat flux from the plates $\eta = 0$ and $\eta = 1$ are, respectively, given by

$$q_0 = -\frac{k}{d}\left(\frac{\partial T}{\partial \eta}\right)_{\eta=0}$$

and

$$q_1 = -\frac{k}{d}\left(\frac{\partial T}{\partial \eta}\right)_{\eta=1} \tag{28}$$

Neglecting edge effects, the heat transfer per unit time from a circular disk of radius "a" coinciding with the planes $\eta = 0$ and $\eta = 1$, respectively, are given by

$$Q_0 = \frac{1}{\pi a^2}\int_0^a 2\pi r q_0 dr$$

$$Q_1 = \frac{1}{\pi a^2}\int_0^a 2\pi r q_1 dr \tag{29}$$

For finite "a", (29) reduce to

$$Q_0 = \frac{(T_0 - T_1)k}{d}\left[1 + \frac{ME^*}{4\sinh M}(1 + \cosh M)\right],$$

$$Q_1 = \frac{(T_0 - T_1)k}{d}\left[1 - \frac{ME^*}{4\sinh M}(1 + \cosh M)\right], \tag{30}$$

where $E^* = PEa^2/2d^2$ is a dimensionless number. The variation of $(Q_0 d/k(T_0 - T_1))$ and $(Q_1 d/k(T_0 - T_1))$ for different values of M when $E^* = 1, 2, 3$, have been plotted. It has been observed that the magnitudes of the heat flux from both the plates increase with the increase of the Hartmann number, i.e., the heat flux increases with the increase of the strength of the magnetic field. Thus the magnitudes of the heat flux are more in the conducting fluid in presence of magnetic field than that in an ordinary Newtonian viscous fluid.

From the equation of energy for the MHD flow, we see that the term $\mathbf{J}^2/\sigma$ is added to the equation, which depends on the electric field, the magnetic field and the velocity. Also electric field depends on the velocity and the magnetic. Therefore, we can say that the current depends on the magnetic field when the velocity and the applied electric field are kept constant. Hence the temperature distribution in the medium increases with the increase of the Hartmann number, i.e., the magnetic field though the boundary mediums are kept at constant temperature.

The studies of these types of problems help in understanding many geophysical and technological problems. It has been observed that uses of conducting fluid in presence of a magnetic field, increase the temperature distribution in the medium. This temperature distribution depends on the strength of the magnetic field as well as the conductivity of the medium. By decreasing the strength of the magnetic field or the conductivity of the substance, we can cool the medium. Thus the results can be used for cooling the nuclear reactor.

References

1 Srivastava, A. C., and Sharma, S. K., "The Effect of a Transverse Magnetic Field on the Flow Between Two Infinite Disks—One Rotating and Other at Rest," *Bull. Acad. Pol. Sci. Sor. Sci. Tech.*, Vol. 9, (11), 1961, pp. 639–644.

2 Stephenson, C. J., "Magnetohydrodynamic Flow Between Rotating Coaxial Disks," *Journal of Fluid Mechanics*, Vol. 38, Part 2, 1969, pp. 335–352.

3 Srivastava, A. C., "Flow of Non-Newtonian Fluid at Small Reynolds Number Between Two Infinite Disks," *Quarterly Journal of Mathematics and Mechanics*, Vol. 14(13), 1961, pp. 353–358.

4 Srivastava, A. C., "Flow of Second-Order Fluids With Heat Transfer," *AIAA Journal*, Vol. 2 (4), 1964, pp. 749–750.

On the Stokes Second Problem of Slightly Non-Newtonian Fluids

N. Phan-Thien[1]

1 Introduction

The description of the Stokes second problem is trivially simple: a flow field is set up on the upper half plane ($y > 0$) by oscillating the boundary $y = 0$ in time so that the fluid velocity at the wall is given by

$$u(0, t) = Ue^{i\omega t}; \quad i^2 = -1 \tag{1}$$

If the fluid is Newtonian of kinematic viscosity $\nu = \eta/\rho$, the exact solution can be easily shown to be

$$u(y, t) = Ue^{-\sqrt{i\omega/\nu}\, y + i\omega t} \tag{2}$$

This solution indicates that, at a distant δ from the wall given by

$$\delta = 2\sqrt{\frac{2\nu}{\omega}} \tag{3}$$

the amplitude of the motion is about 13 percent of its maximum value U. That is, viscous effects extend over a distance $\sim\delta$ (boundary-layer thickness).

The solution to the corresponding non-Newtonian problem is not yet known for the simple reason that the nonlinearity in any reasonable constitutive equation adopted defeats exact analytical attempts. However, if the fluid is slightly non-Newtonian in the sense of Kazakia and Rivlin [1], then analytical progress is possible and we can make some simple qualitative statements about the boundary layer of such a fluid.

2 Analysis

Basically, we consider fluids obeying the following constitutive equation

$$\boldsymbol{\tau} = 2\eta\mathbf{D} + \epsilon \sum_{s=0}^{\infty} [\mathbf{B}_t(s)] \tag{4}$$

where $\boldsymbol{\tau}$ is the extra stress tensor, $\mathbf{D}$, the strain rate, η, the fluid Newtonian viscosity, $\epsilon\mathbf{F}$, the extra non-Newtonian stress ($\epsilon \ll 1$), and $\mathbf{B}_t(s)$ is a strain tensor [$\mathbf{B}_t(s) = \mathbf{0}$ is the rest history].

In flow fields where the velocity is a perturbation about the Newtonian velocity we have

$$\mathbf{u} = \mathbf{u}_0 + \epsilon\mathbf{u}_1 + \ldots \tag{5}$$

$$\mathbf{B}_t(s) = \mathbf{B}_t{}^{(0)}(s) + \epsilon\mathbf{B}_t{}^{(1)}(s) + \ldots \tag{6}$$

and thus

$$\sum_{s=0}^{\infty} [\mathbf{B}_t(s)] = \sum_{s=0}^{\infty} [\mathbf{B}_t{}^{(0)}(s)] + \partial \sum_{s=0}^{\infty} [\mathbf{B}_t{}^{(0)}(s)/\epsilon\mathbf{B}_t{}^{(1)}(s)] \tag{7}$$

In (7) $\partial\mathbf{F}$ is a term of order $O(\epsilon)$. Hence, the first-order non-Newtonian extra stresses are fully determined once the Newtonian velocity, $\mathbf{u}_0$, is specified.

Returning to the problem at hand, the relevant equation of motion is

$$\rho\frac{\partial u}{\partial t} = \eta\frac{\partial^2 u}{\partial y^2} + \epsilon\frac{\partial \tau}{\partial y} \tag{8}$$

where u is the x-component velocity which satisfies the boundary condition (1) and τ is the extra non-Newtonian shear stress.

Adopting the perturbation scheme (5), then it is easy to show that u_0 is the Newtonian velocity given by (2) and u_1 is governed by

$$\rho\frac{\partial u_1}{\partial t} = \eta\frac{\partial^2 u_1}{\partial y^2} + \frac{\partial \tau_0}{\partial y} \tag{9}$$

where

$$\tau_0 = \sum_{s=0}^{\infty} [\mathbf{B}_t{}^{(0)}(s)] \tag{10}$$

In view of the boundary conditions, u_1 takes the form

$$u_1 = \sum_{n=1}^{\infty} w_n(y)e^{in\omega t} \tag{11}$$

where $' \equiv \partial/\partial y$ and

$$w_n'' - \frac{in\omega}{\nu}w_n = -\frac{\omega}{2\eta\pi}\int_0^{2\pi/\omega} \tau_0'\, e^{-i\omega n t}dt \tag{12}$$

$$w_n(0) = 0.$$

To obtain specific solutions we adopt the following constitutive relations for τ_0:

$$\tau_0 = \tilde{\eta}\dot{k}_0 + \hat{\eta}\ddot{k}_0 + \bar{\eta}k_0{}^3 \tag{13}$$

and

$$\tau_0 = \int_0^{\infty} G(s)k_0(t - s)\, ds \tag{14}$$

In (13) $\tilde{\eta}, \hat{\eta}, \bar{\eta}$ are constants and k_0 is the Newtonian shear rate given by

$$k_0 = \frac{\partial u_0}{\partial y} = -U\sqrt{\frac{i\omega}{\nu}}\exp\left(-\sqrt{\frac{i\omega}{\nu}}\, y + i\omega t\right) \tag{15}$$

In (14), $G(s)$ is a "relaxation spectrum."

The status of the approximation (13) has been discussed at length in [1]. Equation (14) is only an empirical approximation (the integral Maxwell model) and does not enjoy any special status in the construction of a hierarchy of constitutive equations.

In a constant shear rate experiment, (13) and (14) predict that the fluid viscosity takes the form, respectively,

$$\mu(\dot{\gamma}) = \eta + \epsilon\bar{\eta}\dot{\gamma}^2 \tag{16}$$

and

$$\mu(\dot{\gamma}) = \eta + \epsilon\int_0^{\infty} G(s)\, ds \tag{17}$$

In an oscillatory shear rate experiment, (13) and (14) demand that the fluid complex viscosity be given by, respectively,

$$\mu^*(\omega) = \eta - \epsilon\hat{\eta}\omega^2 + i\epsilon\tilde{\eta}\omega \equiv \eta' - i\eta'' \tag{18}$$

and

$$\mu^*(\omega) = \eta + \epsilon\eta^* \equiv \eta' - i\eta'' \tag{19}$$

where

$$\eta^* = \int_0^{\infty} G(s)e^{-i\omega s}\, ds \tag{20}$$

Now, using (13), equation (12) becomes

$$w_n'' - \frac{in\omega}{\nu}w_n = \frac{U\omega^2}{\eta\nu}(\tilde{\eta} + i\omega\hat{\eta})e^{-\sqrt{i\omega/\nu}\, y}\delta_{1n} + 3\frac{U^3\omega^2}{\eta\nu^2}\bar{\eta}e^{-3\sqrt{i\omega/\nu}\, y}\delta_{3n} \tag{21}$$

from which it can be verified that

$$w_1 = -\frac{U\omega^2}{2\eta\nu}\sqrt{\frac{\nu}{i\omega}}(\tilde{\eta} + i\omega\hat{\eta})ye^{-\sqrt{i\omega/\nu}\, y} \tag{22}$$

and

[1] Lecturer, Department of Mechanical Engineering, University of Sydney, N.S.W., 2006, Australia.

Manuscript received by ASME Applied Mechanics Division, March, 1980.

$$w_3 = \frac{i\omega}{2\nu} U^3 \frac{\bar{\eta}}{\eta} (e^{-\sqrt{3i\omega/\nu}\, y} - e^{-3\sqrt{i\omega/\nu}\, y}) \tag{23}$$

all other w_n are zero.

Thus the fluid velocity, neglecting terms of order $O(\epsilon^2)$ is

$$u = u_0 + \epsilon U \left(-\frac{\omega}{2\eta} \sqrt{\frac{\omega}{\nu}} (\tilde{\eta}^2 + \omega^2 \hat{\eta}^2)^{1/2} y e^{-\sqrt{\omega/2\nu}\, y + i(\omega t - \sqrt{\omega/2\nu}\, y - \pi/4 + \theta)} + \frac{i\omega}{2\nu} U^2 \frac{\bar{\eta}}{\eta} (e^{-\sqrt{3i\omega/\nu}\, y} - e^{-3\sqrt{i\omega/\nu}\, y}) \right), \tag{24}$$

where

$$\theta = \tan^{-1}\left(\frac{\omega\hat{\eta}}{\tilde{\eta}}\right) = \tan^{-1}\left(\frac{\eta' - \eta}{\eta''}\right) \tag{25}$$

Thus whether the boundary layer (say, the distance over which u decreases to $1/e^2 \simeq 13$ percent of its maximum value U) is thicker than its Newtonian counterpart depends on θ, i.e., depends on the fluid properties. In fact, neglecting the second term inside the curly brackets of (24), the boundary layer of the fluid is thicker than its Newtonian value if $\tan\theta > 1$ or

$$\frac{\eta' - \eta}{\eta''} > 1 \tag{26}$$

On the other hand, if we adopt (14) as the constitutive equation, then (9) becomes

$$\rho \frac{\partial u_1}{\partial t} = \eta \frac{\partial^2 u_1}{\partial y^2} + U \frac{i\omega}{\nu} \eta^*(\omega) e^{-\sqrt{i\omega/\nu}\, y + i\omega t} \tag{27}$$

where $\eta^*(\omega)$ is defined by (20).

The solution to (27) is

$$u_1 = \frac{U\eta^*}{2\eta} \sqrt{\frac{i\omega}{\nu}}\, y e^{-\sqrt{i\omega/\nu}\, y + i\omega t} \tag{28}$$

and the fluid velocity up to $O(\epsilon)$ is

$$u = u_0 + \frac{\epsilon |\eta^*| U}{2\eta} \sqrt{\frac{\omega}{\nu}}\, y e^{-\sqrt{\omega,2\nu}\, y + i(\omega t - \sqrt{\omega/2\nu}\, y + \pi/4 - \phi)} \tag{29}$$

where

$$\phi = \tan^{-1} \frac{\eta''}{\eta' - \eta} \tag{30}$$

Again, the boundary layer of the fluid is thicker than its Newtonian value if $\tan\phi < 1$ or

$$\frac{\eta' - \eta}{\eta''} > 1$$

which is equivalent to condition (26).

Most polymeric liquids possess a positive η'' and $\eta > \eta'$. Thus, in this flow field, the boundary layer of a slightly non-Newtonian fluid is thinner than that of a Newtonian fluid having the same viscosity. This is true even if the fluid is fully Maxwellian in which case the velocity is given by

$$u = U \exp\left(-\sqrt{\frac{i\omega\rho}{\mu^*}}\, y + i\omega t\right) \tag{31}$$

The generalization of this statement to other flow fields must be taken with caution [2].

References

1 Kazakia, J. Y., and Rivlin, R. S., "The Influence of Vibration on Poiseuille Flow of a Non-Newtonian Fluid," *Rheology Acta*, Vol. 17, 1978, pp. 210–226.

2 Denn, M. M., "Boundary Layer Flows For a Class of Elastic Fluids," *Chem. Eng. Sci.*, Vol. 22, 1967, pp. 395–405.

Force of Extraction for a Cylinder Buried in Sand[1]

S. C. Cowin,[2] and L. E. Trent[3]

Lower bounds are derived for the force needed to extract a cylinder from a larger cylinder when the annular region between the two cylinders contains a granular material.

Introduction

In this Note, we solve for the lower bound on the force (F) that is necessary to extract a cylinder from a larger cylindrical container when the annular region between the two cylinders is filled with a granular material such as sand or grain. In our analysis, the extracted cylinder and the container are very rigid compared to the granular material. Also, the granular material is cohesionless (has no tensile strength) and is porous enough so there are no pneumatic effects during extraction. For the coefficients of friction, the coefficient for the granular material on itself is greater than those for the granular material on the surface of the extracted cylinder or on the surface of the container. The extraction force is the force necessary to initiate motion of the cylinder from the granular material.

Perhaps more interesting than the problem addressed is the method of solution. The method is unusual in that it employs for granular material a simple constitutive assumption introduced in 1895 by Janssen [1] and improved recently by Cowin [2]. The constitutive assumption is that the ratio of the horizontal stress exerted by a granular material on the vertical walls of its container (averaged over the container perimeter) to the vertical stress in the granular material (averaged over the cross-sectional area occupied by the granular material) is a constant. This constant is called the pressure ratio and is denoted by K. The values of K reported by Caughey, et al., [3] are 0.61 for wheat, 0.6 for shelled corn, 0.38 for soy beans, 0.4 for cement, 0.39 for sand, and 0.33 for pea gravel. A reassessment of the experimental determination of K-values is given in [4]. This constitutive assumption, the equations of equilibrium, and the frictional properties of the surface lead to a differential inequality for the vertical stress in the granular material. Solution of this differential inequality and subsequent integrations yield the extraction force.

Statement of the Problem

We let F denote the force necessary to extract the buried cylinder of weight W, buried length l, and cross-sectional perimeter L_*. The buried cylinder is contained in a larger cylinder of length greater than or equal to l and of cross-sectional perimeter L_0. The annular region between the two cylinders is filled with a granular material and has corss-sectional area A.

Fig. 1 illustrates the horizontal cross section of the two cylinders and the granular material. Fig. 2 gives a perspective view of the problem and illustrates the surface surcharge stress P. The coefficient of friction between the granular material and the cylinder to be extracted is denoted by μ_*, and the coefficient between the granular material and the larger cylinder is denoted by μ. The weight per unit volume of the granular material is γ. For generality, we assume that the pressure ratio K might be different for the two surfaces, because one will be generally concave and the other generally convex. The pressure ratios for the smaller and larger cylinders will be denoted

[1] This work was performed under the auspices of the U. S. Department of Energy by the Lawrence Livermore Laboratory under Contract Number W-7405-Eng-48.

[2] Consultant to Lawrence Livermore Laboratory. Permanent Address: School of Engineering, Tulane University, 6823 St. Charles Avenue, New Orleans, La. 70118. Mem. ASME.

[3] University of California, Lawrence Livermore Laboratory, Livermore, Calif. 94550.

Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, May, 1980.

$$w_3 = \frac{i\omega}{2\nu} U^3 \frac{\bar{\eta}}{\eta} (e^{-\sqrt{3i\omega/\nu}\, y} - e^{-3\sqrt{i\omega/\nu}\, y}) \tag{23}$$

all other w_n are zero.

Thus the fluid velocity, neglecting terms of order $O(\epsilon^2)$ is

$$u = u_0 + \epsilon U \left(-\frac{\omega}{2\eta} \sqrt{\frac{\omega}{\nu}} (\tilde{\eta}^2 + \omega^2 \hat{\eta}^2)^{1/2} y e^{-\sqrt{\omega/2\nu}\, y + i(\omega t - \sqrt{\omega/2\nu}\, y - \pi/4 + \theta)} + \frac{i\omega}{2\nu} U^2 \frac{\bar{\eta}}{\eta} (e^{-\sqrt{3i\omega/\nu}\, y} - e^{-3\sqrt{i\omega/\nu}\, y}) \right), \tag{24}$$

where

$$\theta = \tan^{-1}\left(\frac{\omega\hat{\eta}}{\tilde{\eta}}\right) = \tan^{-1}\left(\frac{\eta' - \eta}{\eta''}\right) \tag{25}$$

Thus whether the boundary layer (say, the distance over which u decreases to $1/e^2 \simeq 13$ percent of its maximum value U) is thicker than its Newtonian counterpart depends on θ, i.e., depends on the fluid properties. In fact, neglecting the second term inside the curly brackets of (24), the boundary layer of the fluid is thicker than its Newtonian value if $\tan\theta > 1$ or

$$\frac{\eta' - \eta}{\eta''} > 1 \tag{26}$$

On the other hand, if we adopt (14) as the constitutive equation, then (9) becomes

$$\rho \frac{\partial u_1}{\partial t} = \eta \frac{\partial^2 u_1}{\partial y^2} + U \frac{i\omega}{\nu} \eta^*(\omega) e^{-\sqrt{i\omega/\nu}\, y + i\omega t} \tag{27}$$

where $\eta^*(\omega)$ is defined by (20).

The solution to (27) is

$$u_1 = \frac{U\eta^*}{2\eta} \sqrt{\frac{i\omega}{\nu}}\, y e^{-\sqrt{i\omega/\nu}\, y + i\omega t} \tag{28}$$

and the fluid velocity up to $O(\epsilon)$ is

$$u = u_0 + \frac{\epsilon |\eta^*| U}{2\eta} \sqrt{\frac{\omega}{\nu}}\, y e^{-\sqrt{\omega,2\nu}\, y + i(\omega t - \sqrt{\omega/2\nu}\, y + \pi/4 - \phi)} \tag{29}$$

where

$$\phi = \tan^{-1} \frac{\eta''}{\eta' - \eta} \tag{30}$$

Again, the boundary layer of the fluid is thicker than its Newtonian value if $\tan\phi < 1$ or

$$\frac{\eta' - \eta}{\eta''} > 1$$

which is equivalent to condition (26).

Most polymeric liquids possess a positive η'' and $\eta > \eta'$. Thus, in this flow field, the boundary layer of a slightly non-Newtonian fluid is thinner than that of a Newtonian fluid having the same viscosity. This is true even if the fluid is fully Maxwellian in which case the velocity is given by

$$u = U \exp\left(-\sqrt{\frac{i\omega\rho}{\mu^*}}\, y + i\omega t\right) \tag{31}$$

The generalization of this statement to other flow fields must be taken with caution [2].

References

1 Kazakia, J. Y., and Rivlin, R. S., "The Influence of Vibration on Poiseuille Flow of a Non-Newtonian Fluid," *Rheology Acta*, Vol. 17, 1978, pp. 210–226.

2 Denn, M. M., "Boundary Layer Flows For a Class of Elastic Fluids," *Chem. Eng. Sci.*, Vol. 22, 1967, pp. 395–405.

Force of Extraction for a Cylinder Buried in Sand[1]

S. C. Cowin,[2] and L. E. Trent[3]

Lower bounds are derived for the force needed to extract a cylinder from a larger cylinder when the annular region between the two cylinders contains a granular material.

Introduction

In this Note, we solve for the lower bound on the force (F) that is necessary to extract a cylinder from a larger cylindrical container when the annular region between the two cylinders is filled with a granular material such as sand or grain. In our analysis, the extracted cylinder and the container are very rigid compared to the granular material. Also, the granular material is cohesionless (has no tensile strength) and is porous enough so there are no pneumatic effects during extraction. For the coefficients of friction, the coefficient for the granular material on itself is greater than those for the granular material on the surface of the extracted cylinder or on the surface of the container. The extraction force is the force necessary to initiate motion of the cylinder from the granular material.

Perhaps more interesting than the problem addressed is the method of solution. The method is unusual in that it employs for granular material a simple constitutive assumption introduced in 1895 by Janssen [1] and improved recently by Cowin [2]. The constitutive assumption is that the ratio of the horizontal stress exerted by a granular material on the vertical walls of its container (averaged over the container perimeter) to the vertical stress in the granular material (averaged over the cross-sectional area occupied by the granular material) is a constant. This constant is called the pressure ratio and is denoted by K. The values of K reported by Caughey, et al., [3] are 0.61 for wheat, 0.6 for shelled corn, 0.38 for soy beans, 0.4 for cement, 0.39 for sand, and 0.33 for pea gravel. A reassessment of the experimental determination of K-values is given in [4]. This constitutive assumption, the equations of equilibrium, and the frictional properties of the surface lead to a differential inequality for the vertical stress in the granular material. Solution of this differential inequality and subsequent integrations yield the extraction force.

Statement of the Problem

We let F denote the force necessary to extract the buried cylinder of weight W, buried length l, and cross-sectional perimeter L_*. The buried cylinder is contained in a larger cylinder of length greater than or equal to l and of cross-sectional perimeter L_0. The annular region between the two cylinders is filled with a granular material and has corss-sectional area A.

Fig. 1 illustrates the horizontal cross section of the two cylinders and the granular material. Fig. 2 gives a perspective view of the problem and illustrates the surface surcharge stress P. The coefficient of friction between the granular material and the cylinder to be extracted is denoted by μ_*, and the coefficient between the granular material and the larger cylinder is denoted by μ. The weight per unit volume of the granular material is γ. For generality, we assume that the pressure ratio K might be different for the two surfaces, because one will be generally concave and the other generally convex. The pressure ratios for the smaller and larger cylinders will be denoted

[1] This work was performed under the auspices of the U. S. Department of Energy by the Lawrence Livermore Laboratory under Contract Number W-7405-Eng-48.

[2] Consultant to Lawrence Livermore Laboratory. Permanent Address: School of Engineering, Tulane University, 6823 St. Charles Avenue, New Orleans, La. 70118. Mem. ASME.

[3] University of California, Lawrence Livermore Laboratory, Livermore, Calif. 94550.

Manuscript received by ASME Applied Mechanics Division, March, 1980; final revision, May, 1980.

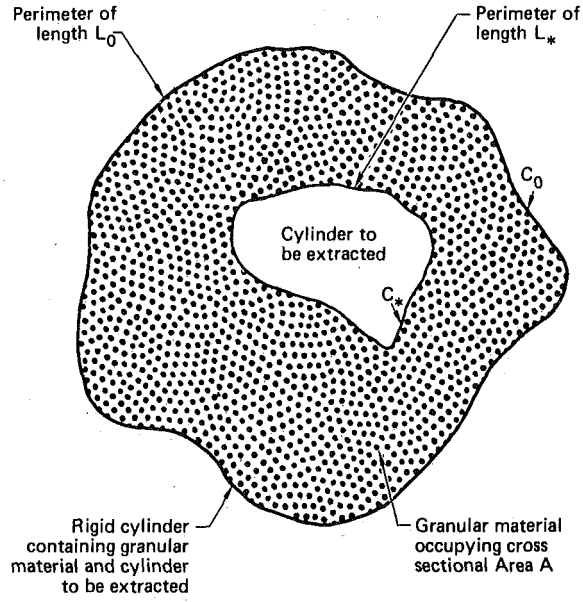


Fig. 1 Horizontal cross section of the two cylinders and the granular material

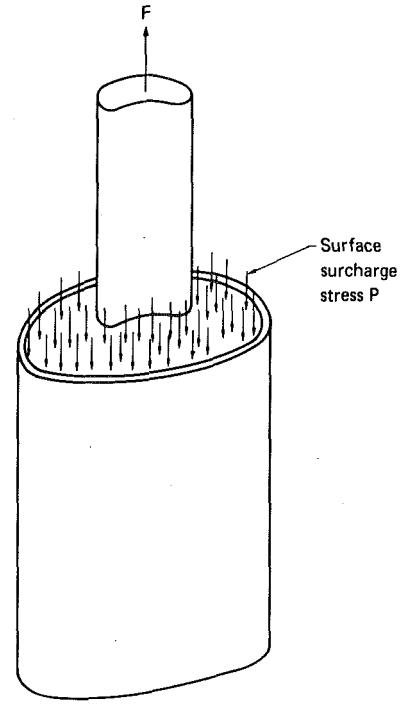


Fig. 2 Perspective view of the problem

by K_* and K_0, respectively. We let C_* denote the curve coincident with the perimeter of the cylinder to be extracted and C_0 the curve coincident with the inner surface of the container.

The problem is to determine the force F in terms of the quantities introduced in the previous paragraph. The solution can be presented more concisely if two-dimensionless parameters α and β are introduced. They are defined by

$$\alpha \equiv \frac{\mu_0 K_0 L_0}{\mu_* K_* L_*} \quad \text{and} \quad \beta \equiv \mu_* K_* L_* \frac{l}{A}. \tag{1}$$

The principal results of this Note are the following lower bounds on the force of extraction F. If slip occurs at the surface C_0 of the containing cylinder ($\alpha < 1$), the lower bound is

$$F \geq W + \frac{Al\gamma}{1-\alpha}\left(1 - \alpha\left\{\frac{1-\exp[-\beta(1-\alpha)]}{\beta(1-\alpha)}\right\}\right) + \frac{\alpha AP}{1-\alpha}\{1-\exp[-\beta(1-\alpha)]\}. \tag{2}$$

If slip occurs simultaneously on both surfaces ($\alpha = 1$) or if slip occurs at the surface C_* of the cylinder being extracted ($\alpha > 1$), the lower bound is

$$F \geq W + \frac{Al\gamma}{1+\alpha}\left(1 - \left\{\frac{1-\exp[-\beta(1+\alpha)]}{\beta(1+\alpha)}\right\}\right) + \frac{AP}{1+\alpha}\{1-\exp[-\beta(1+\alpha)]\}. \tag{3}$$

Equilibrium Analysis

In this section, we derive an appropriate form of the stress equation of equilibrium. We let T_{ij} denote the components of the stress tensor and n_i denote the components of the two-dimensional unit vector normal to the inner or outer cylindrical surface. The components t_i of the stress vector that act on one of the cylindrical surfaces are then related to the stress-tensor components T_{ij} by $t_i = T_{ij}n_j$, where the summation convention over repeated indices is to be employed. The component of the stress tensor acting on the cylindrical surfaces in the direction normal to the surfaces is denoted by T_{nn} and related to t_i and T_{ij} by $T_{nn} = t_i n_i = T_{ij}n_i n_j$. In our analysis, we are concerned with the components T_{nn} and T_{zz} of the stress tensor and the component t_z of the stress vector acting on the vertical cylindrical surfaces.

The analysis involves three different averages of the stress components. Some components of the stress tensor will be averaged over the cross-sectional area A of the granular material and others will be averaged over the inner and outer perimeters C_* and C_0, respectively, of the cross-sectional area. The average of the function $f(x, y, z)$ over the cross-sectional area is defined by

$$\bar{f}(z) = \frac{1}{A}\int_A \int f(x, y, z)dxdy,$$

and the perimeter averages are defined by

$$f^*(z) = \frac{1}{L_*}\int_{C_*} f(x, y, z)ds$$

and

$$f_0(z) = \frac{1}{L_0}\int_{C_0} f(x, y, z)ds.$$

In the notation we have introduced, the stress equation of equilibrium necessary for the present analysis is

$$L_0 t_z{}^0 - L_* t_z{}^* + A\frac{d\overline{T}_{zz}}{dz} + A\gamma = 0. \tag{4}$$

This equation may be derived either by the differential slice method used in strength of materials and by Janssen [1] or by integration of the stress equations of equilibrium as done by Cowin [2].

Constitutive Assumption

Following Janssen [1] and Cowin [2], the constitutive assumption used for the granular material is that the ratio of the boundary average of the normal stress T_{nn} to the cross-sectional average of the vertical stress T_{zz} is a constant. One cylinder surface is concave and the other is convex. To account for this geometric difference, two pressure ratios are introduced, one for the inner and one for the outer perimeter,

$$K_* = \frac{T_{nn}{}^*}{\overline{T}_{zz}} \quad \text{and} \quad K_0 = \frac{T_{nn}{}^0}{\overline{T}_{zz}}, \tag{5}$$

respectively. In Janssen's original work, T_{nn} and T_{zz} were assumed to be uniform over the perimeter and cross section, respectively; hence, from Janssen's viewpoint, there would be no distinction between K_* and K_0. There are no measurements of the difference between K_* and K_0, and there may be no actual difference.

Solution When Slip Occurs at the Surface of the Extracted Cylinder

When slip is about to occur on a surface, the full frictional force is being applied. Thus, at the surface of the cylinder to be extracted, we assume that the average boundary shear stress t_z^* equals the normal stress T_{nn} times the firction coefficient μ_*:

$$t_z^* = -\mu_* T_{nn}^*, \tag{6}$$

where both sides of (6) are positive numbers. There are two considerations that lead to the negative sign in (6). First, the average boundry shear stress is positive upward on the inner cylindrical surface of the sand, as it should be if the extracted cylinder is slipping relative to the granular material. Second, because T_{nn} must be compressive, the sign convention that tensile stress is positive is being employed in this Note.

We assume that slip is not occurring at the outer cylindrical surface. The average boundary shear stress t_z^0 acting on that surface is required to satisfy the "no-slip" inequalities, e.g.,

$$\mu_0 T_{nn}^0 \le t_z^0 \le -\mu_0 T_{nn}^0. \tag{7}$$

When the definitions in (5) for the pressure ratios K_* and K_0 are employed in formulas (6) and (7), we find that

$$t_z^* = -\mu_* K_* \overline{T}_{zz} \tag{8}$$

and

$$\mu_0 K_0 \overline{T}_{zz} \le t_z^0 \le -\mu_0 K_0 \overline{T}_{zz}. \tag{9}$$

A differential inequality in $\overline{T}_{zz}$ is obtained by placing (8) and the equilibrium condition (4) into the left inequality of (9) and then employing definition (1). Thus

$$\frac{d\overline{T}_{zz}}{dz} + \gamma \le -\frac{\overline{T}_{zz}}{l}\beta(1+\alpha). \tag{10}$$

The right inequality of (9) leads to a contradiction unless $\alpha < 1$, a condition corresponding to slip at the outer cylindrical surface rather than at the inner.

The solution to the differential inequality (10) subject to the boundary condition $\overline{T}_{zz} = -P$ at $z = 0$ is

$$-\overline{T}_{zz} \ge \frac{\gamma l}{\beta(1+\alpha)} + \left[P - \frac{\gamma l}{\beta(1+\alpha)}\right] \exp\left[-\beta(1+\alpha)\frac{z}{l}\right]. \tag{11}$$

The solution of the differential inequality (10) subject to this boundary condition is straightforward and is given in the Appendix of the paper by Cowin [2]. From (1) and (11), an inequality involving the average boundary shear stress on the surface of the cylinder to be extracted is determined,

$$t_z^* \ge \frac{\mu_* K_* \gamma l}{\beta(1+\alpha)} + \mu_* K_* \left[P - \frac{\gamma l}{\beta(1+\alpha)}\right] \exp\left[-\beta(1+\alpha)\frac{z}{l}\right]. \tag{12}$$

This inequality is the key to construction of the lower bound on the force of extraction F (determined in the next paragraph).

We consider now a free-body diagram of the cylinder to be extracted. The force of extraction F must equal the weight W plus the shear force acting on the sides of the cylinder. Thus

$$F = W + \int_0^l t_z^* L_* dz. \tag{13}$$

Because the granular material in which the cylinder rests is cohesionless, the granular material across the container bottom offers no resistance to the upward motion of the cylinder. The final result, (3), follows when (12) and (13) are combined and the indicated integration accomplished. This result also holds in the case where slipping occurs on both cylindrical surfaces bounding the granular material.

Solution When Slip Occurs at the Inner Surface of the Container

For slip to occur at the outer perimeter of the annular region containing the granular material, α must be less than one. In this case the average stress t_z^0 for the outer-perimeter is given by

$$t_z^0 = -\mu_0 K_0 \overline{T}_{zz}; \tag{14}$$

and the average boundary stress t_z^* on the inner perimeter must satisfy the inequalities

$$\mu_* K_* T_{zz} \le t_z^* \le -\mu_* K_* \overline{T}_{zz}. \tag{15}$$

A differential inequality in T_{zz} is obtained by placing (14) and the equilibrium condition (4) into the right inequality of (15). Thus

$$\frac{d\overline{T}_{zz}}{dz} + \gamma \le -\beta(1+\alpha)\frac{\overline{T}_{zz}}{l}. \tag{16}$$

Similar treatment of the left inequality in (15) does not produce any new results. The solution to (16) subject to the boundary condition $\overline{T}_{zz} = -P$ at $z = 0$ is

$$-\overline{T}_{zz} \ge \frac{\gamma l}{\beta(1-\alpha)} + \left[P - \frac{\gamma l}{\beta(1-\alpha)}\right] \exp\left[-\beta(1-\alpha)\frac{z}{l}\right]. \tag{17}$$

Hence from (14),

$$t_z^0 \ge \frac{\mu_0 K_0 \gamma l}{\beta(1-\alpha)} + \mu_0 K_0 \left[P - \frac{\gamma l}{\beta(1-\alpha)}\right] \exp\left[-\beta(1-\alpha)\frac{z}{l}\right]. \tag{18}$$

This inequality is necessary in the construction of the lower bound on the extraction force (given in the next paragraph).

We now consider a free-body diagram consisting of the cylinder to be extracted and the annular volume of granular material surrounding the cylinder. The force of extraction F must be equal to the weight W of the cylinder plus the weight $A\gamma l$ of the annular volume of the granular material plus shear force acting on the outer perimeter of the annular granular-material volume; thus

$$F = W + A\gamma l + \int_0^l t_z^0 L_0 dz. \tag{19}$$

The granular material across the container bottom offers no resistance to the upward motion of the cylinder, because the material in which the cylinder rests is cohesionless. This is because cohesionless granular materials can sustain no tensile stresses. The final result (2) follows when (18) and (19) are combined and the indicated integration accomplished.

References

1 Janssen, H. A., "Versuche über Getreidedruck in Silozellen," *Zeitschrift Verein Deutscher Ingenieure*, Vol. 39, August 31, 1895, pp. 1045–1049.

2 Cowin, S. C., "The Theory of Static Loads in Bins," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 409–412.

3 Caughey, R. A., Tooles, C. W., and Scheer, A. C., "Lateral and Vertical Pressure of Granular Material in Deep Bins," *Bulletin No. 173*, Iowa Engineering Experiment Station, Iowa State College, Ames, Iowa, November 14, 1951.

4 Sundaram, V., and Cowin, S. C., "A Reassessment of Static Bin Pressure Experiments," *Powder Technology*, Vol. 22, 1979, pp. 23–32.

Fracture Initiation From Singular Points of Rigid Inclusions

E. E. Gdoutos[1]

The objective of this work is to study the general problem of frac-

[1] Professor of Applied Mechanics, Democritus University of Thrace, Xanthi, Greece. Mem. ASME.

Manuscript received by ASME Applied Mechanics, February, 1980; final revision, June, 1980.

Solution When Slip Occurs at the Surface of the Extracted Cylinder

When slip is about to occur on a surface, the full frictional force is being applied. Thus, at the surface of the cylinder to be extracted, we assume that the average boundary shear stress t_z^* equals the normal stress T_{nn} times the firction coefficient μ_*:

$$t_z^* = -\mu_* T_{nn}^*, \tag{6}$$

where both sides of (6) are positive numbers. There are two considerations that lead to the negative sign in (6). First, the average boundry shear stress is positive upward on the inner cylindrical surface of the sand, as it should be if the extracted cylinder is slipping relative to the granular material. Second, because T_{nn} must be compressive, the sign convention that tensile stress is positive is being employed in this Note.

We assume that slip is not occurring at the outer cylindrical surface. The average boundary shear stress t_z^0 acting on that surface is required to satisfy the "no-slip" inequalities, e.g.,

$$\mu_0 T_{nn}^0 \le t_z^0 \le -\mu_0 T_{nn}^0. \tag{7}$$

When the definitions in (5) for the pressure ratios K_* and K_0 are employed in formulas (6) and (7), we find that

$$t_z^* = -\mu_* K_* \overline{T}_{zz} \tag{8}$$

and

$$\mu_0 K_0 \overline{T}_{zz} \le t_z^0 \le -\mu_0 K_0 \overline{T}_{zz}. \tag{9}$$

A differential inequality in $\overline{T}_{zz}$ is obtained by placing (8) and the equilibrium condition (4) into the left inequality of (9) and then employing definition (1). Thus

$$\frac{d\overline{T}_{zz}}{dz} + \gamma \le -\frac{\overline{T}_{zz}}{l}\beta(1+\alpha). \tag{10}$$

The right inequality of (9) leads to a contradiction unless $\alpha < 1$, a condition corresponding to slip at the outer cylindrical surface rather than at the inner.

The solution to the differential inequality (10) subject to the boundary condition $\overline{T}_{zz} = -P$ at $z = 0$ is

$$-\overline{T}_{zz} \ge \frac{\gamma l}{\beta(1+\alpha)} + \left[P - \frac{\gamma l}{\beta(1+\alpha)}\right] \exp\left[-\beta(1+\alpha)\frac{z}{l}\right]. \tag{11}$$

The solution of the differential inequality (10) subject to this boundary condition is straightforward and is given in the Appendix of the paper by Cowin [2]. From (1) and (11), an inequality involving the average boundary shear stress on the surface of the cylinder to be extracted is determined,

$$t_z^* \ge \frac{\mu_* K_* \gamma l}{\beta(1+\alpha)} + \mu_* K_* \left[P - \frac{\gamma l}{\beta(1+\alpha)}\right] \exp\left[-\beta(1+\alpha)\frac{z}{l}\right]. \tag{12}$$

This inequality is the key to construction of the lower bound on the force of extraction F (determined in the next paragraph).

We consider now a free-body diagram of the cylinder to be extracted. The force of extraction F must equal the weight W plus the shear force acting on the sides of the cylinder. Thus

$$F = W + \int_0^l t_z^* L_* dz. \tag{13}$$

Because the granular material in which the cylinder rests is cohesionless, the granular material across the container bottom offers no resistance to the upward motion of the cylinder. The final result, (3), follows when (12) and (13) are combined and the indicated integration accomplished. This result also holds in the case where slipping occurs on both cylindrical surfaces bounding the granular material.

Solution When Slip Occurs at the Inner Surface of the Container

For slip to occur at the outer perimeter of the annular region containing the granular material, α must be less than one. In this case the average stress t_z^0 for the outer-perimeter is given by

$$t_z^0 = -\mu_0 K_0 \overline{T}_{zz}; \tag{14}$$

and the average boundary stress t_z^* on the inner perimeter must satisfy the inequalities

$$\mu_* K_* T_{zz} \le t_z^* \le -\mu_* K_* \overline{T}_{zz}. \tag{15}$$

A differential inequality in T_{zz} is obtained by placing (14) and the equilibrium condition (4) into the right inequality of (15). Thus

$$\frac{d\overline{T}_{zz}}{dz} + \gamma \le -\beta(1+\alpha)\frac{\overline{T}_{zz}}{l}. \tag{16}$$

Similar treatment of the left inequality in (15) does not produce any new results. The solution to (16) subject to the boundary condition $\overline{T}_{zz} = -P$ at $z = 0$ is

$$-\overline{T}_{zz} \ge \frac{\gamma l}{\beta(1-\alpha)} + \left[P - \frac{\gamma l}{\beta(1-\alpha)}\right] \exp\left[-\beta(1-\alpha)\frac{z}{l}\right]. \tag{17}$$

Hence from (14),

$$t_z^0 \ge \frac{\mu_0 K_0 \gamma l}{\beta(1-\alpha)} + \mu_0 K_0 \left[P - \frac{\gamma l}{\beta(1-\alpha)}\right] \exp\left[-\beta(1-\alpha)\frac{z}{l}\right]. \tag{18}$$

This inequality is necessary in the construction of the lower bound on the extraction force (given in the next paragraph).

We now consider a free-body diagram consisting of the cylinder to be extracted and the annular volume of granular material surrounding the cylinder. The force of extraction F must be equal to the weight W of the cylinder plus the weight $A\gamma l$ of the annular volume of the granular material plus shear force acting on the outer perimeter of the annular granular-material volume; thus

$$F = W + A\gamma l + \int_0^l t_z^0 L_0 dz. \tag{19}$$

The granular material across the container bottom offers no resistance to the upward motion of the cylinder, because the material in which the cylinder rests is cohesionless. This is because cohesionless granular materials can sustain no tensile stresses. The final result (2) follows when (18) and (19) are combined and the indicated integration accomplished.

References

1 Janssen, H. A., "Versuche über Getreidedruck in Silozellen," *Zeitschrift Verein Deutscher Ingenieure*, Vol. 39, August 31, 1895, pp. 1045–1049.

2 Cowin, S. C., "The Theory of Static Loads in Bins," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 409–412.

3 Caughey, R. A., Tooles, C. W., and Scheer, A. C., "Lateral and Vertical Pressure of Granular Material in Deep Bins," *Bulletin No. 173*, Iowa Engineering Experiment Station, Iowa State College, Ames, Iowa, November 14, 1951.

4 Sundaram, V., and Cowin, S. C., "A Reassessment of Static Bin Pressure Experiments," *Powder Technology*, Vol. 22, 1979, pp. 23–32.

Fracture Initiation From Singular Points of Rigid Inclusions

E. E. Gdoutos[1]

The objective of this work is to study the general problem of frac-

[1] Professor of Applied Mechanics, Democritus University of Thrace, Xanthi, Greece. Mem. ASME.

Manuscript received by ASME Applied Mechanics, February, 1980; final revision, June, 1980.

ture initiation from singular points of the boundary of a rigid inclusion embedded in a matrix.

Consider a rigid inclusion with a singular corner 0 perfectly bonded to an infinite isotropic elastic plate which is subjected to a system of stresses at infinity. A reference frame of Cartesian coordinates is attached to the point 0 with the x-axis along the tangent of the boundary of the inclusion at 0. For this situation Panasyuk, et al. [1], gave the following expressions for the polar components of stress σ_r, σ_θ, $\tau_{r\theta}$ in the vicinity of the point 0:

$$\sigma_r = \frac{1}{4\sqrt{2r}}\left[k_1\left[5\cos\frac{\theta}{2} + (2\kappa+1)\cos\frac{3\theta}{2}\right] - k_2\left[5\sin\frac{\theta}{2} + (2\kappa-1)\sin\frac{3\theta}{2}\right]\right]$$

$$\sigma_\theta = \frac{1}{4\sqrt{2r}}\left[k_1\left[3\cos\frac{\theta}{2} - (2\kappa+1)\cos\frac{3\theta}{2}\right] - k_2\left[3\sin\frac{\theta}{2} - (2\kappa-1)\sin\frac{3\theta}{2}\right]\right]$$

$$\tau_{r\theta} = \frac{1}{4\sqrt{2r}}\left[k_1\left[\sin\frac{\theta}{2} - (2\kappa+1)\sin\frac{3\theta}{2}\right] + k_2\left[\cos\frac{\theta}{2} - (2\kappa-1)\cos\frac{3\theta}{2}\right]\right] \quad (1)$$

where the coefficients k_1 and k_2 depend on the loading conditions, the material of the plate and the geometrical shape of the inclusion at the singular corner. In the foregoing relations $\kappa = (3-\nu)/(1+\nu)$ or $\kappa = 3 - 4\nu$ for plane stress or plane strain conditions, respectively, with ν representing the Poisson's ratio of the material of the plate.

For the determination now of the failure mode of the composite plate we make the hypothesis that the fracture of the plate starts from the singular corner 0. On the basis of this hypothesis we will now analyze the fracture behavior of the composite plate using the maximum hoop stress and the minimum strain-energy density models as they were developed by Erdogan and Sih [2] and Sih [3], respectively.

According to the maximum hoop stress model fracture initiation is controlled by the maximum hoop stress σ_θ in the vicinity of the singular corner. Thus we have the following relations:

$$\frac{\partial\sigma_\theta}{\partial\theta} = 0 \quad \frac{\partial^2\sigma_\theta}{\partial\theta^2} < 0 \quad (2)$$

Introducing the value of σ_θ from the second of relations (1) into the above first equation, the following equation is obtained:

$$(1+\kappa)\,k_1\tan^3\frac{\theta}{2} + (3\kappa-1)\,k_2\tan^2\frac{\theta}{2} - (1+3\kappa)\,k_1\tan\frac{\theta}{2} - (\kappa-1)\,k_2 = 0 \quad (3)$$

For any given set of values k_1 and k_2 and the material constant κ solution of equation (3) enables the determination of the critical angle θ_1 which the fracture path will follow with respect to the x-axis of the inclusion at 0.

The critical stress of fracture σ_{1cr} is determined by calculating the value of σ_θ for $\theta = \theta_1$ from the second of relations (1). For this reason a critical value c of the radial distance r at which the stress θ_{1cr} is measured should be introduced. Thus the critical stress σ_{1cr} is determined from the relation

$$\sigma_\alpha = \sigma_\theta(\sigma_{1cr}, c, \theta_1) \quad (4)$$

where σ_α is the allowable stress of the material of the matrix.

According to the minimum strain-energy-density model fracture is controlled by the minimum strain energy in the vicinity of the singular corner. The value of the strain-energy-density factor S, as it was defined in reference [3], is given by the following relation:

$$S = a_{11}k_1^2 + 2a_{12}k_1k_2 + a_{22}k_2^2 \quad (5)$$

with

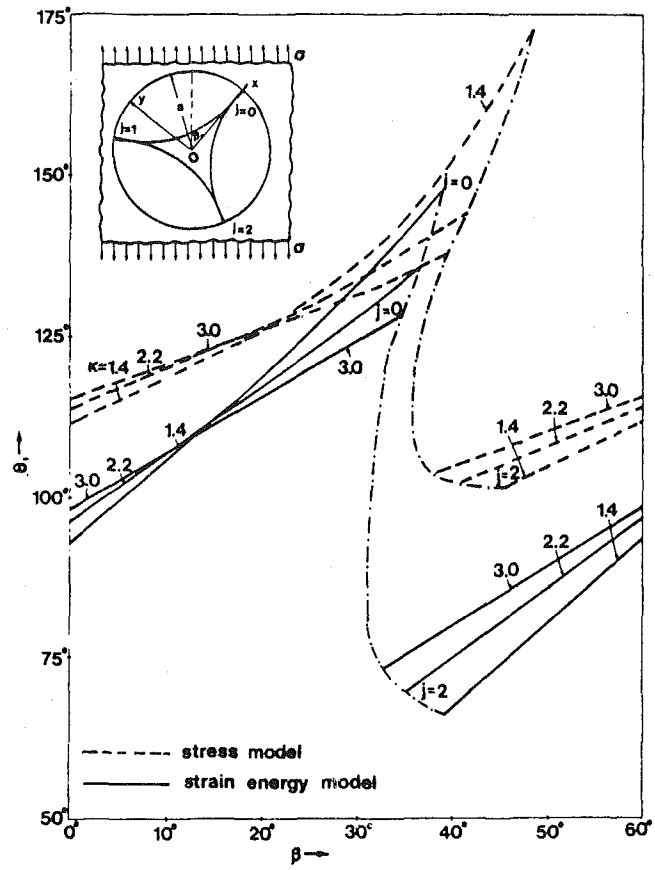


Fig. 1 Variation of the fracture angle θ_1 versus the angle β for tensile applied loads according to the maximum hoop stress and minimum strain-energy-density models. Regions where fracture initiation starts from either of the corners $j = 0$ or $j = 2$ are indicated in the figure.

$$16\,Ga_{11} = 2(\kappa-1)\cos^2\frac{\theta}{2} + \kappa^2 + (2\kappa+1)\cos^2\theta$$

$$16Ga_{12} = -\,[(\kappa-1) + 2\kappa\cos\theta]\sin\theta$$

$$16Ga_{22} = 2(\kappa-1)\sin^2\frac{\theta}{2} + \kappa^2 - (2\kappa-1)\cos^2\theta \quad (6)$$

Applying now the principles of the strain-energy-density model [3] we obtain the following relations:

$$[(\kappa-1) + 2(2\kappa+1)\cos\theta]\sin\theta k_1^2 + 2[(\kappa-1)\cos\theta + 2\kappa\cos 2\theta]k_1k_2 + [-(\kappa-1) - 2(2\kappa-1)\cos\theta]\sin\theta k_2^2 = 0 \quad (7)$$

$$[-(\kappa-1)\cos\theta - 2(2\kappa+1)\cos 2\theta]k_1^2 + 2[(\kappa-1) + 8\kappa\cos\theta]\sin\theta k_1k_2 + [(\kappa-1)\cos\theta + 2(2\kappa-1)\cos 2\theta]k_2^2 > 0 \quad (8)$$

Solution of equation (7) in combination with inequality (8) enables the determination of the value θ_1 of the angle θ which the fracture path of the composite plate starting from the singular corner at 0 will form with the x-axis.

If we introduce the value of θ_1 into relations (5) and (6) we obtain the critical value S_{min} of the strain-energy-density factor. It is assumed, that S_{min} represents a material constant and therefore experimental determination of S_{min} enables the subsequent calculation of the critical fracture stress σ_{cr} of the composite plate.

The foregoing developed maximum hoop stress and minimum strain-energy models were applied to the case of a composite plate reinforced by a hypocycloidal inclusion with three singular corners (Fig. 1). If β is the angle that the applied stress σ subtends with the x-axis then the coefficients k_1 and k_2 are given by [1]

$$k_1^{(j)} = \frac{\sqrt{2a}}{3\kappa}\,\sigma\left[\frac{\kappa-1}{2} + \cos\left(\frac{4\pi j}{3} - 2\beta\right)\right]$$
$$k_2^{(j)} = \frac{\sqrt{2a}}{3\kappa}\,\sigma\sin\left(\frac{4\pi j}{3} - 2\beta\right) \tag{9}$$

with $j = 0, 1, 2$ for the three corners of the inclusion (Fig. 1).

Introducing these values of $k_1^{(j)}$ and $k_2^{(j)}$ into equation (3) of the maximum hoop stress model and solving the resulting equation the values of angles $\theta_1^{(j)}$ for each particular corner of the inclusion were determined. For each such angle $\theta_1^{(j)}$ the corresponding critical stress $\theta_{cr}^{(j)}$ required for fracture initiation from the corner j was calculated. The critical stress σ_{cr} of the composite plate corresponds to the minimum of the aforementioned three stresses $\sigma_{cr}^{(j)}$ ($j = 0, 1, 2$). The same procedure was then followed for the case of the minimum strain-energy-density model. Fig. 1 presents the variation of the fracture angle θ_1 versus angle β according to the maximum hoop stress and the minimum strain-energy models. Comparing the values of θ_1 as they are determined by both models we observe that they are in satisfactory agreement.

References

1 Panasyuk, V. V., Berezhnitskii, L. T., and Trush, I. I., "Stress Distribution About Defects Such as Rigid Sharp-Angled Inclusions," *Problem y Prochnosti*, Vol. 7, 1972, pp. 3–9.

2 Erdogan, F, and Sih, G. C., "On the Crack Extension in Plates Under Plane Loading and Transverse Shear," ASME *Journal of Basic Engineering*, Vol. 85, 1963, pp. 519–525.

3 Sih, G. C., "Strain-Energy-Density Factor Applied to Mixed Mode Crack Problems," *International Journal of Fracture*, Vol. 10, 1974, pp. 305–321.

On The Nonbuckling of a Circular Ring Under a "Wrapping" Loading

T. J. Lardner[1]

Introduction

In a recent unpublished manuscript, Farris and Filippov [1] note the interesting result that a thin-walled circular cylinder wrapped by a high tension band will not buckle from the pressure load imposed on the cylinder by the wrapping. Furthermore, they found that this absence of buckling is independent of the material behavior of the cylinder. As a consequence, they conclude that this type of loading allows for the convenient investigation of the behavior of the material in the cylinder under high compressive loading, a situation important for the study of polymeric materials. The proof they offer for the result on nonbuckling is based on a specific set of equilibrium equations for circular cylindrical shells for which buckling solutions are sought under perturbations from the initial state.

The purpose of this Note is to present a different derivation for circular rings which relies directly on the exact form of the equilibrium equations.

We show that for circular rings (long circular cylinders) that the conclusion of Farris and Filippov can be demonstrated in a very simple way, that is, we show that circular rings with arbitrary material properties will not buckle under a wrapping load (in the absence of friction). The wrapping load can be visualized as placing a string around the ring (as in the nature of a capstan) and pulling tightly on both ends.

Since a wrapping load is an *intrinsic* geometric loading, we write the equations in an intrinsic form which exhibits the changes in geometry through changes in angles independent of the displacements. We also obtain *en passant* the buckling load for an inextensible ring with an active pressure loading.

Formulation

Consider a circular ring of radius a; the equilibrium equations of the ring can be derived following the formulation of Simmonds [2]. The equations expressed in terms of the axial force F, the shear force Q, and the bending moment M are

$$F' - (1+\theta')Q + aP_T(1+\Gamma) = 0$$
$$Q' + (1+\theta')F + aP_N(1+\Gamma) = 0 \tag{1}$$
$$M' + a(1+\Gamma)Q = 0$$

where primes indicate differentiation with respect to ϕ, the position on the undeformed ring, P_N and P_T are the loadings per unit length of the deformed ring in the normal and tangential directions of the deformed ring, $1 + \Gamma$ is the stretch ratio of the ring center line, Γ is the extensional strain, and θ is the change of the tangent angle of the ring.

Upon elimination between the three equilibrium equations in (1), we find

$$\left(\frac{1}{1+\theta'}\left(\frac{M'}{a(1+\Gamma)}\right)'\right)' + (1+\theta')\left(\frac{M'}{a(1+\Gamma)}\right) - a^2\left(\frac{P_N(1+\Gamma)}{(1+\theta')}\right)' + a^2P_T(1+\Gamma) = 0 \tag{2}$$

Equation (2) is a consequence of equilibrium alone. Of importance for our later consideration is the combination of the loading terms in (2).

Consider first the case of a nearly inextensible elastic ring for which $1 + \Gamma = 1$ and

$$M = aB\Delta\kappa = B\theta'$$

We obtain from (2)

$$\left(\frac{\theta'''}{1+\theta'}\right)' + (1+\theta')\theta'' + \frac{p_N\theta''}{(1+\theta')^2} - \frac{p_N'}{(1+\theta')} + p_T = 0 \tag{3}$$

where the nondimensional pressures are $p = a^2P/B$.

We consider now an "active" pressure loading for which

$$P_N = P, \quad \text{a constant}$$

For small θ, we have from (3)

$$\theta^{\mathrm{IV}} + \theta''(1+p) = 0 \tag{4}$$

It follows that periodic solutions of (4) for small θ of the form $\cos 2\phi$ are possible if $p = 3$, corresponding to the classical buckling load of an inextensible ring.[2]

Finally, if the loading on the ring is that of a pressure induced from a string with constant tension T wrapped around the ring, we have (in the absence of friction)

$$P_N = T\kappa; \quad P_T = 0 \tag{5}$$

or

$$P_N(1+\Gamma) = T(1+\theta')(1/a) \tag{6}$$

where κ is the curvature of the deformed ring. Equation (2) becomes

$$\left(\frac{1}{1+\theta'}\left(\frac{M'}{a(1+\Gamma)}\right)'\right)' + (1+\theta')\left(\frac{M'}{a(1+\Gamma)}\right) = 0 \tag{7}$$

The periodic solution in ϕ of (7) is $M = 0$. Hence no buckling occurs

[1] Professor, Department of Civil Engineering, University of Massachusetts, Amherst, Mass. 01003. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, April, 1980.

[2] For loadings of the form $\mathbf{P} = P_n$, we find $p = 4$.

$$k_1^{(j)} = \frac{\sqrt{2a}}{3\kappa}\,\sigma\left[\frac{\kappa-1}{2} + \cos\left(\frac{4\pi j}{3} - 2\beta\right)\right]$$
$$k_2^{(j)} = \frac{\sqrt{2a}}{3\kappa}\,\sigma\sin\left(\frac{4\pi j}{3} - 2\beta\right) \tag{9}$$

with $j = 0, 1, 2$ for the three corners of the inclusion (Fig. 1).

Introducing these values of $k_1^{(j)}$ and $k_2^{(j)}$ into equation (3) of the maximum hoop stress model and solving the resulting equation the values of angles $\theta_1^{(j)}$ for each particular corner of the inclusion were determined. For each such angle $\theta_1^{(j)}$ the corresponding critical stress $\theta_{cr}^{(j)}$ required for fracture initiation from the corner j was calculated. The critical stress σ_{cr} of the composite plate corresponds to the minimum of the aforementioned three stresses $\sigma_{cr}^{(j)}$ ($j = 0, 1, 2$). The same procedure was then followed for the case of the minimum strain-energy-density model. Fig. 1 presents the variation of the fracture angle θ_1 versus angle β according to the maximum hoop stress and the minimum strain-energy models. Comparing the values of θ_1 as they are determined by both models we observe that they are in satisfactory agreement.

References

1 Panasyuk, V. V., Berezhnitskii, L. T., and Trush, I. I., "Stress Distribution About Defects Such as Rigid Sharp-Angled Inclusions," *Problem y Prochnosti*, Vol. 7, 1972, pp. 3–9.

2 Erdogan, F, and Sih, G. C., "On the Crack Extension in Plates Under Plane Loading and Transverse Shear," ASME *Journal of Basic Engineering*, Vol. 85, 1963, pp. 519–525.

3 Sih, G. C., "Strain-Energy-Density Factor Applied to Mixed Mode Crack Problems," *International Journal of Fracture*, Vol. 10, 1974, pp. 305–321.

On The Nonbuckling of a Circular Ring Under a "Wrapping" Loading

T. J. Lardner[1]

Introduction

In a recent unpublished manuscript, Farris and Filippov [1] note the interesting result that a thin-walled circular cylinder wrapped by a high tension band will not buckle from the pressure load imposed on the cylinder by the wrapping. Furthermore, they found that this absence of buckling is independent of the material behavior of the cylinder. As a consequence, they conclude that this type of loading allows for the convenient investigation of the behavior of the material in the cylinder under high compressive loading, a situation important for the study of polymeric materials. The proof they offer for the result on nonbuckling is based on a specific set of equilibrium equations for circular cylindrical shells for which buckling solutions are sought under perturbations from the initial state.

The purpose of this Note is to present a different derivation for circular rings which relies directly on the exact form of the equilibrium equations.

We show that for circular rings (long circular cylinders) that the conclusion of Farris and Filippov can be demonstrated in a very simple way, that is, we show that circular rings with arbitrary material properties will not buckle under a wrapping load (in the absence of friction). The wrapping load can be visualized as placing a string around the ring (as in the nature of a capstan) and pulling tightly on both ends.

Since a wrapping load is an *intrinsic* geometric loading, we write the equations in an intrinsic form which exhibits the changes in geometry through changes in angles independent of the displacements. We also obtain *en passant* the buckling load for an inextensible ring with an active pressure loading.

Formulation

Consider a circular ring of radius a; the equilibrium equations of the ring can be derived following the formulation of Simmonds [2]. The equations expressed in terms of the axial force F, the shear force Q, and the bending moment M are

$$F' - (1+\theta')Q + aP_T(1+\Gamma) = 0$$
$$Q' + (1+\theta')F + aP_N(1+\Gamma) = 0 \tag{1}$$
$$M' + a(1+\Gamma)Q = 0$$

where primes indicate differentiation with respect to ϕ, the position on the undeformed ring, P_N and P_T are the loadings per unit length of the deformed ring in the normal and tangential directions of the deformed ring, $1 + \Gamma$ is the stretch ratio of the ring center line, Γ is the extensional strain, and θ is the change of the tangent angle of the ring.

Upon elimination between the three equilibrium equations in (1), we find

$$\left(\frac{1}{1+\theta'}\left(\frac{M'}{a(1+\Gamma)}\right)\right)' + (1+\theta')\left(\frac{M'}{a(1+\Gamma)}\right) - a^2\left(\frac{P_N(1+\Gamma)}{(1+\theta')}\right)' + a^2P_T(1+\Gamma) = 0 \tag{2}$$

Equation (2) is a consequence of equilibrium alone. Of importance for our later consideration is the combination of the loading terms in (2).

Consider first the case of a nearly inextensible elastic ring for which $1 + \Gamma = 1$ and

$$M = aB\Delta\kappa = B\theta'$$

We obtain from (2)

$$\left(\frac{\theta'''}{1+\theta'}\right)' + (1+\theta')\theta'' + \frac{p_N\theta''}{(1+\theta')^2} - \frac{p_N'}{(1+\theta')} + p_T = 0 \tag{3}$$

where the nondimensional pressures are $p = a^2P/B$.

We consider now an "active" pressure loading for which

$$P_N = P, \quad \text{a constant}$$

For small θ, we have from (3)

$$\theta^{\mathrm{IV}} + \theta''(1+p) = 0 \tag{4}$$

It follows that periodic solutions of (4) for small θ of the form $\cos 2\phi$ are possible if $p = 3$, corresponding to the classical buckling load of an inextensible ring.[2]

Finally, if the loading on the ring is that of a pressure induced from a string with constant tension T wrapped around the ring, we have (in the absence of friction)

$$P_N = T\kappa; \quad P_T = 0 \tag{5}$$

or

$$P_N(1+\Gamma) = T(1+\theta')(1/a) \tag{6}$$

where κ is the curvature of the deformed ring. Equation (2) becomes

$$\left(\frac{1}{1+\theta'}\left(\frac{M'}{a(1+\Gamma)}\right)\right)' + (1+\theta')\left(\frac{M'}{a(1+\Gamma)}\right) = 0 \tag{7}$$

The periodic solution in ϕ of (7) is $M = 0$. Hence no buckling occurs

[1] Professor, Department of Civil Engineering, University of Massachusetts, Amherst, Mass. 01003. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, April, 1980.

[2] For loadings of the form $\mathbf{P} = P_n$, we find $p = 4$.

and this result is independent of the material properties of the ring.

Conclusion

We have presented an additional derivation to confirm the interesting result that circular rings under wrapping type loads will not buckle, a result first noted by Farris and Filippov [1].

References

1 Farris, R., and Filippov, A., "Experimental Method of Determining the Mechanical Characteristics of Materials Under Uniaxial Compression," unpublished manuscript, Department of Polymer Science and Engineering, University of Massachusetts, Amherst, Mass., 1979.

2 Simmonds, J. G., "Accurate Nonlinear Equations and a Perturbation Solution for the Free Vibrations of a Circular Elastic Ring," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, Mar. 1979, pp. 156–160.

Journal of Applied Mechanics

Discussion

On Laminar Dispersion for Flow Through Round Tubes[1]

W. N. Gill[2] and R. S. Subramanian.[3] Professor Yu's article[1] represents a timely contribution to dispersion theory. However, his comments and analysis in the section "Gill's Approach as a Special Case of the Present Method" are wrong because he has erroneously neglected the mixed derivatives

$$\frac{\partial^2 \psi_m}{\partial \tau \partial \zeta} (m = 0,1, \ldots)$$

in his equation (31). Therefore, we must disagree with him on his interpretation of our work on generalized dispersion theory [1, 2], and we will show that his solution and ours are formally identical if his incorrect equation (33) is replaced by our equation (8). Thus what Yu refers to as Gill's approach is *not* a special case of Yu's method.

Professor Yu compares his solution form in equation (7) with ours in his equation (28). The coefficient functions $f_m(\tau, \xi)$ in his equation (28) may be expanded naturally in the form of Yu's equation (7) as may be seen in [1, 2]. That is,

$$f_m(\tau, \xi) = \sum_{n=0}^{\infty} A_{nm}(\tau) \frac{J_0(\beta_n \xi)}{J_0(\beta_n)} \tag{1}$$

where

$$J_1(\beta_n) = 0 \tag{2}$$

for the type of problem investigated by Yu. Thus Yu's $\psi_n(\zeta, \tau)$ may be represented by

$$\psi_n(\zeta, \tau) = \sum_{m=1}^{\infty} A_{nm}(\tau) \frac{\partial^m \psi_0}{\partial \zeta^m} (n = 1, 2, 3 \ldots) \tag{3}$$

in terms of the Fourier coefficients $A_{nm}(\tau)$ of our functions $f_m(\tau, \xi)$ and the derivatives of ψ_0. Thus the two methods are formally equivalent.

Professor Yu's equation (31) for ψ_n contains mixed derivatives of the form

$$\frac{\partial^2 \psi_m}{\partial \tau \partial \zeta} (m = 0, 1, 2 \ldots)$$

which he neglects and consequently draws erroneous conclusions regarding our work. It is important to retain *all* the terms on the right-hand side of his equation (31), and to express the mixed derivatives of ψ_n in terms of purely axial derivatives. Gill, in Yu's reference [5], first showed how to deal with the mixed derivatives. As we have shown (2), (3), in the case of ψ_0, the procedure is to use equation (29) which involves *no assumptions,* and is a direct consequence of the solution form in equation (28). Differentiation of equation (29) with respect to ζ will provide an expression for $\partial^2\psi_0/\partial\tau\partial\zeta$ in terms of purely axial derivatives of ψ_0. The intimate coupling among the ψ_n's makes it more difficult to express $\partial^2\psi_m/\partial\tau\partial\zeta$ immediately in terms of the axial derivatives of ψ_0. However, we will show how this may be done.

Start with Yu's equation (31). Then differentiate his equation (9) with respect to ζ which gives $\partial^2\psi_n/\partial\zeta\partial\tau = \partial^2\psi_n/\partial\tau\partial\zeta$ on the left-hand side in terms of derivatives with respect to ζ only. Insert the resultant expression on the right-hand side for the mixed derivatives in Yu's (31) to obtain

$$\begin{aligned}\psi_n(\zeta, \tau) = &-\frac{1}{\beta_n{}^2}(1 - e^{-\beta_n{}^2\tau}) \left[c_{n0} \frac{\partial \psi_0}{\partial \zeta} + \sum_{m=1}^{\infty}{}' c_{nm} \frac{\partial \psi_m}{\partial \zeta}\right] \\ &+ \frac{1}{\beta_n{}^4}[1 - e^{-\beta_n{}^2\tau}(\beta_n{}^2\tau + 1)] \left[c_{n0}\left(-\sum_{m=0}^{\infty} c_{0m} \frac{\partial^2 \psi_m}{\partial \zeta^2}\right.\right. \\ &+ c_{nn} \frac{\partial^2 \psi_0}{\partial \zeta^2} + \sum_{m=1}^{\infty}{}' c_{nm} \left(-\beta_m{}^2 \frac{\partial \psi_m}{\partial \zeta} - \sum_{l=0}^{\infty} c_{ml} \frac{\partial^2 \psi_l}{\partial \zeta^2}\right. \\ &\left.\left.\left. + c_{nn} \frac{\partial^2 \psi_m}{\partial \zeta^2}\right)\right] + \ldots\end{aligned} \tag{4}$$

Note that $\beta_0 = 0$. Thus we have ψ_n in terms of derivatives with respect to ζ only. Now to evaluate $\partial\psi_m/\partial\zeta$ and $\partial^2\psi_m/\partial\zeta^2$, differentiate equation (4) once or twice. If we are interested only in determining $K_n(\tau)$ for $n = 1$ to 3, we can neglect all ζ derivatives of ψ_0 higher than the third and write

$$\frac{\partial \psi_n}{\partial \zeta} \approx A_n \frac{\partial^2 \psi_0}{\partial \zeta^2} + B_n \frac{\partial^3 \psi_0}{\partial \zeta^3}, \quad n = 1, 2 \ldots \tag{5}$$

$$\frac{\partial^2 \psi_n}{\partial \zeta^2} \approx A_n \frac{\partial^3 \psi_0}{\partial \zeta^3} \tag{6}$$

where $A_n = a_n c_{n0}$

$$B_n = b_n \left[c_{n0} c_{nn} - \sum_{m=1}^{\infty}{}' c_{nm} c_{m0}\right] + \sum_{m=1}^{\infty}{}' a_m c_{m0} c_{nm} (a_n - \beta_m{}^2 b_n)$$

$$a_n = -\frac{1}{\beta_n{}^2}(1 - e^{-\beta_n{}^2\tau})$$

and

$$b_n = \frac{1}{\beta_n{}^4}[1 - e^{-\beta_n{}^2\tau}(\beta_n{}^2\tau + 1)]$$

Consequently, inserting equations (5) and (6) in (4) we get for $n \geq 1$,

$$\psi_n = A_n \frac{\partial \psi_0}{\partial \zeta} + B_n \frac{\partial^2 \psi_0}{\partial \zeta^2} + \ldots \tag{7}$$

Now, inserting (7) in Yu's equation (9) we get

$$\frac{\partial \psi_0}{\partial \tau} = \frac{1}{Pe^2} \frac{\partial \psi_0}{\partial \zeta^2} - \sum_{j=1}^{\infty} c_{0j} \left(A_j \frac{\partial^2 \psi_0}{\partial \zeta^2} + B_j \frac{\partial^3 \psi_0}{\partial \zeta^3}\right) \tag{8}$$

Thus we are able to derive the proper version of Yu's equation (33), presumably by following the procedure he used, but without neglecting the mixed derivatives.

[1] By J. S. Yu, and published in the December, 1979, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 750–756.

[2] Chemical Engineering Department, State University of New York at Buffalo, Amherst, N. Y. 14260.

[3] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, Calif. 91103; on Sabbatical leave from Clarkson College of Technology, Potsdam, N. Y. 13676.

DISCUSSION

In view of the foregoing, we must disagree with Professor Yu's comments presented immediately after his equation (34) regarding our work. We are aware of the fact that for symmetric initial conditions, the two-term dispersion equation will predict only symmetric concentration profiles (see reference [3] for instance). However, the inclusion of terms of order $n > 2$ on the right-hand side of equation (29) *will* introduce asymmetry in the axial profiles. This is clear from an examination of the axial moments of such higher-order models. Thus Professor Yu's statement expressing doubt about the utility of equation (29) with higher-order terms in predicting the mean concentration at small values of time is quite unjustified. Our statements regarding the validity of our approach are supported by the recent work of DeGance and Johns [4, 5] who have shown that nth-order dispersion approximations obtained by truncating our generalized dispersion equation predict *precisely* the first $(n + 1)$ coefficients in a uniformly convergent modified Hermite expansion of the local concentration field.

In conclusion, we believe Professor Yu's article constitutes a useful contribution to the theory of solute dispersion. The only reservation we have expressed concerns his comments regarding the validity of our approach which we believe is formally equivalent to his, the differences being only a matter of the routes used.

References

1 Gill, W. N., and Sankarasubramanian, R., "Exact Analysis of Unsteady Convective Diffusion," *Proceedings of the Royal Society*, London, Series A, Vol. 316, 1970, pp. 341–350.

2 Gill, W. N., and Sankarasubramanian, R., "Dispersion of a Nonuniform Slug in Time-Dependent Flow," *Proceedings of the Royal Society*, London, Series A, Vol. 322, 1971, pp. 101–117.

3 Jayaraj, K., and Subramanian, R. S., "On Relaxation Phenomena in Field-Flow Fractionation," *Separation Science and Technology*, Vol. 13 (9), 1978, pp. 791–817.

4 DeGance, A. E., and Johns, L. E., "The Theory of Dispersion of Chemically Active Solutes in a Rectilinear Flow Field," *Applied Scientific Research*, Vol. 34, 1978, pp. 189–225.

5 DeGance, A. E., and Johns, L. E., "On the Dispersion Coefficients for Poiseuille Flow in a Circular Cylinder," *Applied Scientific Research*, Vol. 34, 1978, pp. 227–258.

Author's Closure

The author wishes to thank Professors Gill and Subramanian for pointing out the mistake that appeared in my paper in the section where equation (31) is made to compare with their general model of dispersion by diffusion approximation (henceforth the G-S model) and for pointing out that the general validity of the G-S model has been rigorously proved by DeGance and Johns (see Discussers' references [4, 5]) who used the technique of expanding the local concentration in Hermite polynomials in the axial coordinate. Neglecting the mixed derivatives in equation (31) gives an equation which, describing the change of the mean concentration, is correct only when terms involving $\partial^3\psi_0/\partial\zeta^3$ and higher can be ignored in the G-S model.

The general comparison between the G-S model and the present approach which uses the Green's function technique may be demonstrated as follows. The connection is made here by observing that equation (16) in my paper applies to $n = 0$ ($\beta_0 = 0$) as well. Let the last term on the RHS of equation (16) arising from the initial conditions be symbolically represented by $F_n(\zeta, \tau)$, i.e.,

$$F_n(\zeta, \tau) = \frac{Pe}{2\sqrt{\pi\tau}} e^{-\beta_n{}^2\tau} \int_{-\infty}^{\infty} \psi_n(\zeta', 0) \exp\left[-\frac{pe^2(\zeta - \zeta')^2}{4\tau}\right] d\zeta' \tag{1}$$

It can be shown that in the axial coordinate moving with the mean flow velocity, (ζ, τ),

$$\psi_0(\zeta, \tau) = -\sum_{m=1}^{\infty} c_{0m} \sum_{\nu=0}^{\infty} \frac{\tau^{\nu+1}}{\nu+1} \sum_{\mu=\nu}^{2\nu} \frac{(-1)^\mu 2^{\mu-\nu}}{Pe^{2(\mu-\nu)}(2\nu-1)!(\mu-\nu)!} \times \frac{\partial^{\mu+1}\psi_m}{\partial\zeta^{2(\mu-\nu)+1}\partial\tau^{2\nu-\mu}} + F_0, \tag{2}$$

and

$$\psi_n(\zeta, \tau) = -\sum_{m\neq n} c_{nm} \sum_{\nu=0}^{\infty} \beta_n{}^{-2(\nu+1)} \times \left[1 - e^{-\beta_n{}^2\tau} \sum_{\sigma=0}^{\nu} \frac{\nu!}{(\nu-\sigma)!} (\beta_n{}^2\tau)^{\nu-\sigma}\right] \sum_{\mu=\nu}^{2\nu} \frac{(-1)^\mu 2^{\mu-\nu}}{Pe^{2(\mu-\nu)}(\mu-\nu)!} \times \sum_{s=0}^{2\nu-\mu} \frac{c_{nn}{}^s}{s!(2\nu-\mu-s)!} \frac{\partial^{\mu+1}\psi_m}{\partial\zeta^{2(\mu-\nu)+s+1}\partial\tau^{2\nu-\mu-s}} + F_n, \quad n > 0, \tag{3}$$

where the velocity profile factors are defined by equation (10) of my paper and $c_{00} = 0$ has been used in equation (2).

Now as it has been so importantly pointed out by Professors Gill and Subramanian, the mixed derivatives in the foregoing equations can all be eliminated by repeated applications of equation (9) in my paper. Thus symbolically equations (2) and (3) may be written in the form

$$\psi_0(\zeta, \tau) = \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} A_{0m}{}^{(k)}(\tau) \frac{\partial^k\psi_m}{\partial\zeta^k} + F_0, \tag{4}$$

and

$$\psi_n(\zeta, \tau) = \sum_{m\neq n}^{\infty} \sum_{k=1}^{\infty} A_{nm}{}^{(k)}(\tau) \frac{\partial^k\psi_m}{\partial\zeta^k} + F_n, \quad n > 0. \tag{5}$$

We now form

$$F_n = R_n F_0, \tag{6}$$

where

$$R_n = F_n/F_0$$

is given by the initial condition and F_0 in equation (6) can be replaced by the expression given by equation (4). Thus we obtain

$$F_n = R_n\left(\psi_0 - \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} A_{0m}{}^{(k)} \frac{\partial^k\psi_m}{\partial\zeta^k}\right), \tag{7}$$

and by substituting equation (7) into equation (5) we get

$$\psi_n(\zeta, \tau) = R_n\psi_0 + \sum_{k=1}^{\infty} A_{n0}{}^{(k)} \frac{\partial^k\psi_0}{\partial\zeta^k} + \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} B_{nm}{}^{(k)} \frac{\partial^k\psi_m}{\partial\zeta^k}, \quad n \geq 1, \tag{8}$$

where

$$B_{nm}{}^{(k)} = A_{nm}{}^{(k)}(1 - \delta_{nm}) - R_n A_{0m}{}^{(k)} \tag{9}$$

with δ_{nm} being the Kronecker delta. Equation (8) is the desired equation relating $\psi_n(n > 0)$ to the transverse mean concentration ψ_0 in reference to the context of the G-S model. It is apparent that the effective use of equation (8) is through successive approximations. If the zeroth approximation for ψ_n is

$$\psi_n{}^{(0)} = R_n\psi_0 + \sum_{k=1}^{\infty} A_{n0}{}^{(k)} \frac{\partial^k\psi_0}{\partial\zeta^k}, \tag{10}$$

then the first approximation is given by

$$\psi_n{}^{(1)} = \psi_n{}^{(0)} + \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} B_{nm}{}^{(k)} \frac{\partial^k\psi_n{}^{(0)}}{\partial\zeta^k}, \tag{11}$$

and in general the pth approximation has the form

$$\psi_n{}^{(p)} = \psi_n{}^{(0)} + \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} B_{nm}{}^{(k)} \frac{\partial^k\psi_n{}^{(p-1)}}{\partial\zeta^k}. \tag{12}$$

The procedure thus reduces $\psi_n(n > 0)$ to a form completely expressed in terms of ψ_0 and its axial derivatives and a close representation for ψ_n can always be obtained by increasing the order of approximation. It can be shown by deduction that, like in the formulations of DeGance and Johns, the pth approximation of ψ_n provides the exact coefficients of $\partial^k\psi_0/\partial\zeta^k$, $0 < k < p$. Once the factors $A_{nm}(\tau)$ defined in

equations (4) and (5) in relation to equations (2) and (3) are algebraically spelled out, the present procedure can be used for the determination of the dispersion coefficients in the G-S model to any desired order.

Amplitude-Frequency Characteristics of Large-Amplitude Vibrations of Sandwich Plates[1]

Yi-Yuan Yu.[2] The writer is glad to see that his earlier work [1] on the nonlinear vibrations of sandwich plates was cited and used by the author as a basis of comparison and that the author's results based on the approximate method of Berger came out to be close to the writer's. However, special attention should be called to the author's opening statement and closing paragraph, where the writer's work [1] was referred to specifically, but not appropriately.

The author's opening statement that the writer has "treated nonlinear vibrations of sandwich plates applying von Karman equations" is in fact incorrect. As is well known, the von Karman equations do not include the effect of transverse shear deformation. In contrast, the writer's equations [1] not only were his own but also were derived to include the transverse shear effect particularly, as was one of the main purposes of the paper.

In his closing paragraph the author attempted to compare his or Berger's method with the writer's "method." As Berger emphasized in his original paper [2], no satisfactory physical justification could be given for his approximation, and any justification would have to be based on comparison with other available exact solutions (which should be called, more properly, "better" solutions, as theories of plates and shells and solutions based on them are inherently not exact). Since the author was apparently using the writer's earlier results as a basis of comparison and justification, it does not seem to be appropriate to claim that the method used by the author has an advantage over the writer's. Indeed, the writer's paper [1] was not meant to present a method of solution. But, as has been stated clearly, aside from deriving the equations of motion, it was to show the frequency-amplitude relations as well as the importance of transverse shear for nonlinear vibrations of sandwich plates, all for the first time.

References

1 Yu, Yi-Yuan, "Nonlinear Flexural Vibrations of Sandwich Plates," *Journal of the Acoustical Society of America*, Vol. 34, No. 9, Sept. 1962, pp. 1176–1183.

2 Berger, H. M., "A New Approach to the Analysis of Large Deflections of Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 22, Dec. 1955, pp. 465–472.

[1] By B. M. Karmakar, and published in the March, 1979, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 230–231.

[2] Executive Engineer, Research and Engineering, Rockwell International, Energy Systems Groups, Canoga Park, Calif. Mem. ASME

equations (4) and (5) in relation to equations (2) and (3) are algebraically spelled out, the present procedure can be used for the determination of the dispersion coefficients in the G-S model to any desired order.

Amplitude-Frequency Characteristics of Large-Amplitude Vibrations of Sandwich Plates[1]

Yi-Yuan Yu.[2] The writer is glad to see that his earlier work [1] on the nonlinear vibrations of sandwich plates was cited and used by the author as a basis of comparison and that the author's results based on the approximate method of Berger came out to be close to the writer's. However, special attention should be called to the author's opening statement and closing paragraph, where the writer's work [1] was referred to specifically, but not appropriately.

The author's opening statement that the writer has "treated nonlinear vibrations of sandwich plates applying von Karman equations" is in fact incorrect. As is well known, the von Karman equations do not include the effect of transverse shear deformation. In contrast, the writer's equations [1] not only were his own but also were derived to include the transverse shear effect particularly, as was one of the main purposes of the paper.

In his closing paragraph the author attempted to compare his or Berger's method with the writer's "method." As Berger emphasized in his original paper [2], no satisfactory physical justification could be given for his approximation, and any justification would have to be based on comparison with other available exact solutions (which should be called, more properly, "better" solutions, as theories of plates and shells and solutions based on them are inherently not exact). Since the author was apparently using the writer's earlier results as a basis of comparison and justification, it does not seem to be appropriate to claim that the method used by the author has an advantage over the writer's. Indeed, the writer's paper [1] was not meant to present a method of solution. But, as has been stated clearly, aside from deriving the equations of motion, it was to show the frequency-amplitude relations as well as the importance of transverse shear for nonlinear vibrations of sandwich plates, all for the first time.

References

1 Yu, Yi-Yuan, "Nonlinear Flexural Vibrations of Sandwich Plates," *Journal of the Acoustical Society of America*, Vol. 34, No. 9, Sept. 1962, pp. 1176–1183.

2 Berger, H. M., "A New Approach to the Analysis of Large Deflections of Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 22, Dec. 1955, pp. 465–472.

[1] By B. M. Karmakar, and published in the March, 1979, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 230–231.

[2] Executive Engineer, Research and Engineering, Rockwell International, Energy Systems Groups, Canoga Park, Calif. Mem. ASME

Journal of Applied Mechanics

Book Reviews

Fracture Mechanics Applied to Brittle Materials. Edited by S. W. Freiman. American Society for Testing and Materials, Philadelphia, Pa. 1979. ASTM Special Technical Publication 678. Pages 232. Price $25.

REVIEWED BY A. S. KOBAYASHI[1]

This state-of-the-science publication on fracture test techniques for glass, ceramics, rocks, and cementitious composites is a compilation of 14 papers which were presented at the Eleventh National Symposium on Fracture Mechanics, which was held June 12–14, 1978, at Virginia Polytechnic Institute and State University, Blacksburg, Va. These papers can be grouped into three categories: papers which relate to the theoretical and experimental details of test techniques; papers which describe the applications of these test techniques; and a paper which explains the use of fracture data in life time prediction of structural components.

Although Griffith fracture theory was motivated by fracture of glass, its modern application to brittle materials is complicated by the unforgiving nature of such materials. Additional complication is introduced in the high temperature test environment of some structural ceramics. These unique test requirements promote the use of double torsion, notch bend, and short bar specimens which are discussed in detail by Fuller, Pletka, et al., Bansal, et al., and Barker. A unique ceramic fracture specimen which uses an indentation generated surface flaw in a bend specimen is discussed by Petrovic, et al., Marian and Evans.

As for material characterization, dynamic fracture response of a birefringent polymer is discussed by Fourney, et al., and Mecholsky, et al., reported on ceramic fracture toughness obtained through fractographic analysis. Fracture toughness of alumina and Westerly granite are reported by Buresch and Schmidt, et al. The complex fracture behavior of cementitious composites is discussed by Naaman, et al.

The last paper by Wiederhorn, et al., presents a fracture mechanics approach for improving the reliability of structural ceramics.

As is the recent ASTM practice, this STP contains a thorough summary by the editor, S. W. Freiman to which the JAM readers are referred to for further details short of reading each of the 14 papers.

Techniques of Finite Elements. By B. Irons and S. Ahmad. Halsted Press, A Division of John Wiley & Sons, Inc. 1979, Pages 529. Price, $72.50.

REVIEWED BY TED BELYTSCHKO[2]

This is certainly the most unusual book on finite elements. Presented in an interesting and witty manner with tremendous insight and intuition, it will intrigue some readers, dismay or infuriate others. The flavor of the book is best transmitted by some quotations. For example, the chapter *Nonlinearity* begins with "Medical Aspects: We have no idea how many nervous breakdowns are precipitated by the persistent failure of computer programs to solve nonlinear problems . . . We suspect the incidence of convulsive iteritis is steadily increasing . . . because of the incurable foolhardiness of certain finite-element users in attempting nearly impossible jobs."

The viewpoint is sometimes rather narrow: "Finite elements are already a billion dollar industry. We feel the time is past when we should confirm our answers by experiment, or compare them with other people's answers." Perhaps the last remark is tongue-in-cheek; it is certainly not shared by all finite-element users.

Another feature which sets this book apart from other finite-element books is the wealth of insights into finite-element concepts. Physical intuition is used with great ingenuity, and many concepts are presented with clarity and succinctness. The book is divided into 7 parts, 29 chapters. The first three parts are devoted to the presentation of the finite-element method to primarily linear problems: with emphasis on energy methods, shape functions, computer programming aspects, shortcomings and pitfalls and the patch test. Part 5 deals with symmetry, eigenvalue problems and, briefly, nonlinearities. Part 6, entitled "Speculations," is precisely that, though speculations are not limited to Part 6. Part 7 gives the theoretical details for the material in the preceding parts. A computer code, with program descriptions is also given; exercises and puzzles are scattered throughout, with solutions (or sketches thereof) in the back.

Although the book is intended as an introductory textbook, its suitability for beginning student is questionable. There is more finite-element "lore" than can be found in any other finite-element book; on the other hand, mundane details are often skimmed over, which sometimes leaves a beginner, or for that matter, a self-professed expert like the reviewer, baffled. However, for an advanced student or a teacher of the finite-element method, this book is invaluable. It is original, thought-provoking, and at the same time, very entertaining.

Stratified Flows. By Chia-Shun Yih. Academic Press, Inc. 1980. Pages xvii–418. Price $29.50

REVIEWED by R. R. LONG[3]

This book is a second edition, with some new material and with a new title, of "Dynamics of Nonhomogeneous Fluids" which appeared about 15 years ago. There is not a great deal of difference between the present and earlier version. The book has an improved appearance with glossy paper, for example, but in general it follows the same format of the earlier work except for a last chapter which extends the analogy between gravitational flows and other flows to the electromagnetic problem. Chapters 1 and 2, for example, still contain the same material, essentially, except for a considerable discussion on internal waves in basins of variable depth in Chapter 2. In Chapter 3 there is an added discussion of waves of permanent form including the work of Benjamin and some added material on shallow water

[1] Professor, University of Washington, Department of Mechanical Engineering, Seattle, Wash. 98195.

[2] Professor, Department of Civil Engineering, Northwestern University, Evanston, Ill. 60201.

[3] Professor of Fluid Mechanics, Department of Earth and Planetary Sciences, The Johns Hopkins University, Baltimore, Md. 21218.

theory in three dimensions and on edge waves in stratified fluids. Chapter 4 has some added material on the stability of stratified flows.

The amount of work done on stratified flows has greatly increased in the past 15 years and even more people will now find this book useful. The material reflects the author's interests rather closely. This is not a criticism, really, because Professor Yih is well recognized as the leading theoretician in this area and his book sets a standard for careful development and rigor which is and was needed to help to insure high standards in this area of research. At a recent meeting on stratified flows in Norway, Professor Yih was honored, and quite rightly so, as a pioneer in this area.

The new edition has a set of notes at the end of each chapter which contain chatty, historical, and reflective material which I rather liked and which adds a pleasant note to the book in general. I think my only comment on the adequacy and usefulness of the book is the lack of material on turbulence in stratified flows. At the Trondheim meeting there were, certainly, a large proportion of papers which attacked the problem of turbulence so that it is an important aspect of stratified flows. If I had been given the task of revising Professor Yih's excellent book I would have added a chapter on this subject.

Dislocations in Solids: The Elastic Theory. Vol. 1. Edited by F. R. N. Nabarro. North-Holland. 1978. Pages 350. Price $47.

REVIEWED BY T. MURA[4]

This is the first of five volumes devoted to the behavior of dislocations and their influence on the properties of solids. It contains five review papers on the fundamental theories of dislocations. The author, title of paper, and summary of contents of each of these papers are listed below.

J. Friedel, "Dislocations—An Introduction," Pages 3–32. Volterra's process constructing dislocation and disclination lines is explained. It is concluded that a rotational dislocation (disclination) is equivalent to a continuous distribution of infinitesimal translational dislocations. The reviewer thinks that this is only true for a discrete disclination. Continuously distributed disclinations cannot be equivalent to any distribution of translational dislocations. Interesting pictures of rotational dislocations are shown for molecular and liquid crystals. Brief discussion is given about the cores of translational dislocations and physical properties other than plasticity.

A. M. Kosevich, "Crystal Dislocations and the Theory of Elasticity," Pages 33–141. This chapter is well suited as a textbook for an introductory course on dislocations for graduate students in mechanics. Although the contents are not particularly new, the material is well organized, giving the important equations and examples. The article treats discrete dislocations, pileup of dislocations, continuous distribution of dislocations, dynamics of dislocation motion, the effective mass of a dislocation, dislocation damping, interaction with point defects, the helical dislocation, and the prismatic dislocation loop, among other topics.

J. W. Steeds and J. R. Willis, "Dislocations in Anisotropic Media," Pages 144–165. When J. W. Steeds wrote his book "Introduction to Anisotropic Elasticity Theory of Dislocations," (Oxford: Clarendon Press), I criticized in a book review that he made no mention of the series of works done by myself, Willis, Barnett, and his coworkers. However, this time Steeds has Willis as a coauthor. The present article exactly supplements the weak points in Steeds' last book. Willis performed the Fourier integrals involved in my formula (elastic distortion expressed by a line integral of Green's function along a dislocation segment). Since my formula reveals its value by Willis' work, I do not mind it being referred to as the Mura-Willis formula. The corresponding formulas of Lothe Brown, and Indenbom and Orlov are derived from the Mura-Willis formula by a simple geometrical consideration.

J. D. Eshelby, "Boundary Problems," Pages 168–221. The elastic field in a body whose outer boundary **S** is subjected to a traction **T** and a displacement **U** is the same as it would be if the material inside **S** formed part of an infinite medium provided **S**, now merely a surface marked out in the infinite medium, is covered with a layer of body force of surface density **T** and is the seat of a Smigliana dislocation whose variable discontinuity vector is equal to **U**. This theorem by Geffia leads to the integral equation for **U**. The article introduces many solutions obtained for a dislocation inside or outside of a circular inhomogeneity. The free boundary is a special case when the shear modulus of the inhomogeneity or that of the matrix becomes zero. A new solution is given for a screw dislocation in a cylinder of a general cross section by the use of the conformal mapping. The article also treats dislocations in a semi-infinite medium, plates, and disks. Elementary beam theory is proposed as an approximate method for an edge dislocation in a beam as well as the elementary torsion theory for a screw dislocation in a rod.

B. K. D. Gairola, "Nonlinear Elastic Problems," Pages 222–342. A nonlinear elastic theory is needed to investigate the strain field close to the dislocation core. Another example is the effect of dislocations on the macroscopic density of crystals. The linear theory of elasticity predicts a vanishing effect, although it has been known for a long time that dislocations lead to a positive volume expansion. The scattering of elastic waves by straight dislocations and kinks and the small-angle scattering of X-rays by dislocation lines and rings also need the nonlinear elastic theory. After a lengthy introduction of nonlinear elasticity, an example of an infinitely long and straight screw dislocation is shown. The solution calculated only up to the second-order is given from the work of Seeger and Mann for vanishing tractions on the core boundary and by the use of Signorini's method for nonvanishing tractions. These solutions are reexamined by the method of Seeger and Wesotowski and that of Green, Rivlin, and Shield. The article also contains the large deformation geometry of continuous distribution of dislocations.

Dislocations in Solids: Dislocations in Crystals. Vol. 2. Edited by F. R. N. Nabarro. North-Holland. 1979. Pages 562. Price $75.

REVIEWED BY T. MURA[5]

The aritcles in Vol. 2 shows how the simple models of a dislocation in an elastic continuum developed in Vol. 1 are modified when the dislocation is formed in a periodic structure. The author, title of paper, and summary of contents of each paper are listed below.

R. Bullough and V. K. Tewary, "Lattice Theories of Dislocations," pp. 1–65. The Peierls-Nabarro model is explained as a partially discrete model of a straight dislocation. According to this model, the strain immediately below the dislocation is too large to be consistent with the assumption of a linear Hooke's law in the elastic block. The corresponding tensile stress exceeds the theoretical tensile strength of many materials. The parametric modification by Foreman is introduced as well as other modifications by van der Merwe.

As a lattice model for a screw dislocation, Maradudin's equation of equilibrium for atoms and his solution are introduced. I don't think that most of readers understand this section unless they read the original paper of Maradudin. The article also treats dislocation-phonon interactions and computer simulation with a suitable interatomic potential.

S. Amelinckx, "Dissociations in Particular Structures," pp. 67–460. This is a richly illustrated chapter with about 400 pages. "As soon as we consider structures with more than one atom in the unit

---

[4] Professor, Department of Civil Engineering, The Technological Institute, Northwestern University, Evanston, Ill. 60201.

[5] Professor, The Technological Institute, Department of Civil Engineering, Northwestern University, Evanston, Ill. 60201.

cell, even simple structures such as the face and body-centered cubes, new considerations arise. In particular stacking faults of low energy surrounded by partial dislocations may be formed." Heidenreich and Shockley's partial dislocations are observed by transmission electron microscopy. More than 300 pictures and photographs are illustrated to discuss detail of the Heidenreich-Shockley model. Amelinckx published a book, "The Direct Observation of Dislocations," Academic Press, 1964. Comparing the present article with his previous book, we can see an advancement of electromicroscopy technique in the past 10 years. The article further discusses the geometry (atomic arrangement), force, and energy associated with dislocations in ordered alloys and covalent structures, dislocations in layer structures and ionic crystals.

J. W. Matthews, "Misfit Dislocations," pp. 461–545. If a pair of crystals with the same orientation but different lattice parameter are placed in perfect contact, the atoms near the interface adjust their position a little. This results in interfacial regions of good and bad register. The regions of bad register resemble crystal dislocations (interfacial or misfit dislocations). Many examples of misfit dislocations have been found in interfaces between face-centered cubic metal films. Misfit dislocations have also been observed in semiconducting devices on the boundary between two differently doped regions. Numerous examples of misfit dislocations have been seen at precipitates in alloys. The mathematical definition of misfit dislocations could be improved if the author employs the concept of surface dislocations used by Bullough and Bilby, *Proceedings of the Physical Society,* Vol. 69, Series B, 1956, p. 1276. Misfit dislocations surrounding an inclusion, for instance, can well be defined by the surface dislocations of Bullough and Bilby when Eshelby's concept of inclusion is additionally considered. The article treats observation of coherent interfaces, examples of misfit dislocations, mechanism for the generation of misfit dislocations, behavior of misfit dislocations during diffusion, and effect of misfit dislocations on interdiffusion. Quantitative tests of predictions and use of misfit strain to improve the perfection of crystals are also discussed.

Two-Phase Flows (Vieweg Tracts in Pure and Applied Physics). By Shih-I Pai. Heyden & Son Inc. 1979. Pages xii–359. Price $49.50

REVIEWED BY M. S. PLESSET[6]

This book covers topics in fluid dynamics which are of increasing interest and importance. The text very well illustrates the variety of rather distinct topics which are part of multiphase flow dynamics. As would be expected, the author covers the various aspects of this subject with exceptional clarity.

Some readers might be surprised that the author chose to include aeroelasticity and hydroelasticity in this text. Other readers might be surprised that the author has devoted his two last chapters to plasma theory and electromagneto-fluid dynamics. Among this same group of readers are some who would have preferred a more extended treatment of interfacial waves, including nonlinear effects; this subject is now of particular interest to many fluid dynamicists. Nevertheless, one must admit that the choice of topics is the author's prerogative, and one must further admit that the treatment of plasma theory, aeroelasticity, and hydroelasticity are indeed excellent.

This book can be strongly recommended to persons with any interest in continuum mechanics. It is a most lucid and readable text.

[6] Professor Emeritus of Engineering Science, California Institute of Technology, Pasadena, Calif. 91125