Transactions of the ASME

Journal of Applied Mechanics

Technical Editor, **L. B. FREUND**
Division of Engineering
Brown University
Providence, R.I. 02912

APPLIED MECHANICS DIVISION

Chairman, **M. M. CARROLL**
Secretary, **J. R. RICE**
Associate Editors,
T. BELYTSCHKO
D. B. BOGY
R. M. CHRISTENSEN
R. J. CLIFTON
T. L. GEERS
MAURICE HOLT
THOMAS J. R. HUGHES
W. D. IWAN
W. G. KNAUSS
F. A. LECKIE
FRANCIS C. MOON
J. W. NUNZIATO
R. H. PLAUT
R. T. SHIELD
J. G. SIMMONDS
K. R. SREENIVASAN
J. S. WALKER
Z. WARHAFT
G. A. WEMPNER
L. T. WHEELER
S. E. WIDNALL

BOARD ON COMMUNICATIONS
Chairman and Vice-President
K. N. REID, JR.

Members-at-Large
W. BEGELL
J. T. COKONIS
W. G. GOTTENBERG
F. LANDIS
J. R. LLOYD
R. E. NICKELL
J. ORTLOFF
C. PHILLIPS
R. E. REDER
F. W. SCHMIDT

President, **L. S. FLETCHER**
Exec. Dir.
PAUL ALLMENDINGER
Treasurer, **ROBERT A. BENNETT**

PUBLISHING STAFF
Mng. Dir., Pub., **J. J. FREY**
Dep. Mng. Dir., Pub.
JOS. SANSONE
Managing Editor,
CORNELIA MONAHAN
Production Editor, **REMO SALTA**
Prod. Asst., **MARISOL ANDINO**

The Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly for $100 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017. Second class postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to The Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.

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

PRICES: To members, $24.00, annually; to nonmembers, $100.00.
Add $6.00 for postage to countries outside the United States and Canada.

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

COPYRIGHT © **1986** by the American Society of Mechanical Engineers. Reprints from this publication may be made on condition that full credit be given the TRANSACTIONS OF THE ASME, JOURNAL OF APPLIED MECHANICS, and the author, and date of publication be stated.

INDEXED by the Engineering Information

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 53 • NUMBER 1 • MARCH 1986

TECHNICAL PAPERS

(Contents continued on Inside Back Cover)

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

CONTENTS (CONTINUED)

DESIGN DATA AND METHODS

BRIEF NOTES

DISCUSSION

BOOK REVIEWS

ERRATA

Impact With Friction

J. B. Keller
Departments of Mathematics
and Mechanical Engineering,
Stanford University,
Stanford, Calif. 94305

A theory of the impact or collision of two rigid bodies, taking account of friction, is presented. It determines how the direction of sliding varies during the impact, which must be known to calculate the direction of the frictional force and thence the frictional impulse. This is accomplished by analyzing the equations of motion of the bodies during the collision. The normal impulse is determined by using a coefficient of restitution. When the direction of sliding is constant throughout the collision, the theory agrees with that given by Whittaker, which is correct only in this case.

1 Introduction

In rigid body mechanics, a collision between two bodies is treated as instantaneous, with contact at a single point. Each body is assumed to exert an impulsive force on the other at the point of contact. In the absence of friction the impulse of this force can be calculated in terms of a coefficient of restitution. But in the presence of friction there is no satisfactory method of determining the impulse within the framework of rigid body mechanics.

To develop one we give up the idea that the collision is instantaneous, and assume instead that it has a duration t_c small compared to a typical time scale T of the motion before or after the collison. From the equations of motion of the two bodies during the collision we determine their relative tangential velocity at the point of contact. This slip velocity and the law of friction yield the time varying frictional force exerted by each body on the other. The integral of this force is the frictional impulse delivered during the collision.

When the direction of slip remains constant throughout the collision, the frictional impulse is in the slip direction and its magnitude is just the coefficient of friction times the magnitude of the normal impulse. This result provides the basis for Whittaker's (1904) method of treating impacts with friction. However, he did not note that it holds only when the slip direction stays constant during the collision, and is incorrect otherwise. Its incorrectness became evident when Kane (1984) applied it to a compound pendulum striking a fixed surface. He found that for certain parameter values, Whittaker's method led to an increase of energy. The explanation is that the slip velocity reverses its direction during the impact, so Whittaker's method is inapplicable.

Our calculation has the intrinsic difficulty associated with the motion of two rigid bodies, and the added difficulty that the normal force is an unknown function of time. To overcome the first difficulty, we take advantage of the fact that the positions of the bodies hardly change during the collision. Formally we do this by solving the problem asymptotically as t_c/T tends to zero. To overcome the second difficulty we use the normal impulse up to time t as the independent variable, instead of t itself. In this way we eliminate the normal force from the equations of motion.

The problem is formulated and simplified via asymptotic considerations in Section 2. Then the normal impulse is introduced as the new independent variable and the tangential impulse is expressed as an integral of the slip direction in Section 3. The slip direction is found in Section 4. The resulting theory is summarized in Section 5, and it is applied to some special cases in Section 6.

Some interesting problems of impact with friction are considered in the mechanics text of Kilmister and Reeve (1966). However, one of their analyses is marred by an error.[1]

2 Formulation

Let us consider two rigid bodies labelled $j=1$ and $j=2$ which collide at the time $t'=0$. Then their center of mass velocities U_j and their angular velocities Ω_j will change discontinuously from their values U_j^- and Ω_j^- just before the collision to their values $U_j^+$and Ω_j^+ just after the collision. The goal of the theory of collisions is to determine the jumps $[U_j]=U_j^+-U_j^-$ and $[\Omega_j]=\Omega_j^+-\Omega_j^-$.

To determine them we introduce the contact force $((-1)^j/t_c)F(t'/t_c)$ exerted on body j at the point of contact by the other body during the collision, which we assume to start at $t'=0$ and to end at $t'=t_c$. The factor $(-1)^j$ guarantees that the forces exerted by the bodies on each other are equal and opposite. The factor $1/t_c$ makes the magnitudes of the forces large enough to have a significant effect during the collision.

The contact force depends upon the velocities of the bodies during the collision, so it is necessary to analyze them. Therefore, we denote by M_j the mass of body j, by J_j its central moment of inertia tensor, by $F_j^e(t'/T)$ the external force applied at its center mass, and by $G_j^e(t'/T)$ the external torque on it. Here T is a characteristic time on which the external forces and torques vary. Then the equations of motion are

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 8, 1985; final revision, August 27, 1985.

---

[1] On pages 190 and 191 the authors treat the problem of a lamina falling vertically, without rotation, in a vertical plane and hitting a fixed horizontal plane. They solve it correctly by Newton's method, but incorrectly by Poisson's method. Their error is to omit v' from the first equation on page 191, which should be $m(V-v')=R'$, and to ignore the condition of zero normal velocity at the end of the compression phase, which is $v'-\omega'(a^2-b^2)^{1/2}=0$. When these errors are corrected, the result is the same as that found by Newton's method.

$$M_j \frac{dU_j}{dt'} = \frac{(-1)^j}{t_c} F\left(\frac{t'}{t_c}\right) + F_j^e\left(\frac{t'}{T}\right), \tag{2.1}$$

$$\frac{d(J_j\Omega_j)}{dt'} = \frac{(-1)^j}{t_c} R_j\left(\frac{t'}{T}\right)$$

$$\times F\left(\frac{t'}{t_c}\right) + G_j^e\left(\frac{t'}{T}\right), \quad j=1,2 \tag{2.2}$$

In (2.2) $R_j(t'/T)$ is the vector from the center of mass of body j to the point of contact.

The concept of impact or collision is based on the assumption that $\epsilon = t_c/T$ is small. To study the motion during the collision we introduce the new dimensionless time variable $t = t'/t_c$ and rewrite (2.1) and (2.2) as

$$M_j \frac{dU_j}{dt} = (-1)^j F(t) + \epsilon T F_j^e(\epsilon t), \tag{2.3}$$

$$\frac{d(J_j\Omega_j)}{dt} = (-1)^j R_j(\epsilon t) \times F(t) + \epsilon T G_j^e(\epsilon t). \tag{2.4}$$

The quantities $U_j(t, \epsilon)$, $\Omega_j(t, \epsilon)$, $J_j(\epsilon t, \epsilon)$ and $R_j(\epsilon t, \epsilon)$ depend upon ϵ as well as upon t. We assume that they have limits as ϵ tends to zero. Therefore, we let ϵ tend to zero in (2.3) and (2.4) to obtain the following equations for these limits:

$$M_j \frac{dU_j}{dt} = (-1)^j F(t), \tag{2.5}$$

$$J_j(0) \frac{d\Omega_j}{dt} = (-1)^j R_j(0) \times F(t). \tag{2.6}$$

We see that $J_j(0)$ and $R_j(0)$ are constant in (2.6) so we shall just write J_j and R_j. The external forces and torques do not occur in these equations, which hold only during the collision. Since $t = t'/t_c$ vanishes at $t' = 0$, the initial conditions are

$$U_j(0) = U_j^-, \quad \Omega_j(0) = \Omega_j^-. \tag{2.7}$$

Now $t = 1$ at $t' = t_c$, so the values U_j^+ and Ω_j^+ after the collision are given by

$$U_j^+ = U_j(1), \quad \Omega_j^+ = \Omega_j(1). \tag{2.8}$$

Then the jumps in U_j and Ω_j during the collision, obtained by integrating (2.5) and (2.6) from $t=0$ to $t=1$ and using (2.7) and (2.8), are

$$[U_j] = (-1)^j M_j^{-1} I \tag{2.9}$$

$$[\Omega_j] = (-1)^j J_j^{-1} (R_j \times I). \tag{2.10}$$

Here I is defined by

$$I = \int_0^1 F(t)dt, \tag{2.11}$$

3 Determination of I

To find I we denote by n the unit normal to the surfaces of the two bodies at the point of contact, pointing into body 2. It is independent of t in the limit $\epsilon = 0$. Then we write $N(t) = n \cdot F(t)$ for the normal component of F. The integral of $N(t)$ from 0 to t is the normal component of impulse exerted upon body 2 up to time t, which we denote by $\tau(t)$:

$$\tau(t) = \int_0^t N(s)ds. \tag{3.1}$$

In terms of τ we can formulate Poisson's version of the theory of impact. First we introduce $u(t)$, the relative velocity of the two bodies at the point of contact, defined by

$$u(t) = U_2 + \Omega_2 \times R_2 - (U_1 + \Omega_1 \times R_1). \tag{3.2}$$

Then we let t_0 be the time when $u_N(t) = n \cdot u(t)$ vanishes:

$$u_N(t_0) = 0 \tag{3.3}$$

The interval from $t=0$ to $t=t_0$ is the compressive phase of the collision, and during it the normal impulse on body 2 is $\tau(t_0)$. The subsequent interval, from $t=t_0$ to $t=1$, is the phase of restitution during which the normal impulse on body 2 is $\tau(1) - \tau(t_0)$. Poisson's hypothesis is that the latter impulse is e times the former, i.e., $\tau(1) - \tau(t_0) = e\tau(t_0)$, where e is the coefficient of restitution. Thus

$$\tau(1) = (1+e)\tau(t_0). \tag{3.4}$$

Since $n \cdot I = \tau(1)$, (3.4) yields the normal component of I.

The tangential part of the contact force is due to friction. It is proportional to the normal force N when the two bodies are sliding relative to one another, so we shall write

$$F(t) = N(t)[n + f(t)], \quad n \cdot f(t) = 0. \tag{3.5}$$

The tangential vector $f(t)$ is given by the law of friction, which holds for $N \geq 0$:

$$\begin{aligned} f &= -\mu \hat{u}_T \quad \text{if } u_T \neq 0, \\ |f| &\leq \mu \quad \text{if } u_T = 0. \end{aligned} \tag{3.6}$$

Here μ is the coefficient of sliding friction, u_T is the tangential part of $u(t)$, and $\hat{u}_T$ is a unit vector along u_T:

$$u_T = u - (u \cdot n)n, \quad \hat{u}_T = \frac{u_T}{|u_T|} \quad \text{if } u_T \neq 0. \tag{3.7}$$

Now use (3.5) for F with $f = -\mu\hat{u}_T$ in (2.11). Since $N(t)dt = d\tau$, as (3.1) shows, we use τ as the integration variable and write $\hat{u}_T$ as a function of τ. In this way we get, with $\tau_o = \tau(t_o)$,

$$I = \tau(1)n - \mu \int_0^{\tau(1)} \hat{u}_T(\tau)d\tau = (1+e)\tau_0 n - \mu \int_0^{(1+e)\tau_0} \hat{u}_T(\tau)d\tau. \tag{3.8}$$

Finally we substitute the result (3.8) for I into (2.9) and (2.10) to get

$$[U_j] = (-1)^j M_j^{-1} \left\{ (1+e)\tau_0 n - \mu \int_0^{(1+e)\tau_0} \hat{u}_T(\tau)d\tau \right\}, \tag{3.9}$$

$$[\Omega_j] = (-1)^j J_j^{-1} \left(R_j \times \left\{ (1+e)\tau_0 n - \mu \int_0^{(1+e)\tau_0} \hat{u}_T(\tau)d\tau \right\} \right). \tag{3.10}$$

4 The Slip Direction $\hat{u}_T(\tau)$

Our results (3.9) and (3.10) involve the slip direction $\hat{u}_T(\tau)$ and the normal impulse τ_0 imparted during the compressive phase. To find them we differentiate (3.2) and use (2.5) and (2.6) to get

$$\frac{du}{dt} = \sum_{j=1}^{2} \{ M_j^{-1} F + (J_j^{-1}[R_j \times F]) \times R_j \}. \tag{4.1}$$

Next we use (3.5) in (4.1) and divide by N, to obtain

$$\frac{1}{N(t)} \frac{du}{dt} = \sum_{j=1}^{2} \{ M_j^{-1}(n+f) + (J_j^{-1}[R_j \times (n+f)]) \times R_j \}. \tag{4.2}$$

From (3.1) we see that $N^{-1}d/dt = d/d\tau$, so we introduce τ as a new independent variable in place of t, and (4.2) becomes

$$\frac{du}{d\tau} = \sum_{j=1}^{2} \{ M_j^{-1}(n+f) + (J_j^{-1}[R_j \times (n+f)]) \times R_j \}. \tag{4.3}$$

The initial value of $u(\tau)$ is obtained by using (2.7) in (3.2) and noting that $t=0$ corresponds to $\tau=0$. In this way we get

$$u(0) = U_2^- + \Omega_2^- \times R_2 - U_1^- - \Omega_1^- \times R_1. \tag{4.4}$$

To solve (4.3) we split it into its normal and tangential parts by multiplying it by $n\cdot$ and $(1 - nn\cdot)$, respectively, which gives

$$\frac{du_N}{d\tau} = M_1^{-1} + M_2^{-1} + \sum_{j=1}^{2} n \cdot \{ (J_j^{-1}[R_j \times (n+f)]) \times R_j \}, \tag{4.5}$$

$$\frac{du_T}{d\tau}=\sum_{j=1}^{2}\{M_j^{-1}f+(1-nn\bullet)(J_j^{-1}[R_j\times(n+f)])\times R_j\}, \qquad (4.6)$$

The initial values $u_N(0)$ and $u_T(0)$ are the normal and tangential parts of the right side of (4.4).

When $u_T(\tau)\neq 0$, (3.6) yields $f=-\mu\hat{u}_T$. Upon using this in (4.5) and integrating from 0 to τ we get

$$u_N(\tau)=u_N(0)+(M_1^{-1}+M_2^{-1})\tau+\sum_{j=1}^{2}n$$

$$\bullet\left\{\left(J_j^{-1}\left[R_j\times\left(n\tau-\mu\int_0^{\tau}\hat{u}_T(\tau')d\tau'\right)\right]\right)\times R_j\right\}. \qquad (4.7)$$

By using (4.7) in (3.3) we obtain the following equation for $\tau_0\equiv\tau(t_0)$:

$$(M_1^{-1}+M_2^{-1})\tau_0+\sum_{j=1}^{2}n\bullet\left\{\left(J_j^{-1}\left[R_j\right.\right.\right.$$

$$\left.\left.\left.\times\left(n\tau_0-\mu\int_0^{\tau_0}\hat{u}_T(\tau')d\tau'\right)\right]\right)\times R_j\right\}=-u_N(0). \qquad (4.8)$$

Next we set $f=-\mu\hat{u}_T$ in (4.6) which becomes

$$\frac{du_T}{d\tau}=-\mu(M_1^{-1}+M_2^{-1})\hat{u}_T(\tau)$$

$$+\sum_{j=1}^{2}(1-nn\bullet)\{(J_j^{-1}[R_j\times(n-\mu\hat{u}_T)])\times R_j\} \qquad (4.9)$$

From (4.4) we get the initial condition

$$u_T(0)=(1-nn\bullet)(U_2^-+\Omega_2^-\times R_2-U_1^--\Omega_1^-\times R_1). \qquad (4.10)$$

These equations, (4.9) and (4.10), determine $\hat{u}_T(\tau)$ for $0\leq\tau\leq(1+e)\tau_0$.

To solve (4.9) we introduce the constant vector a and the two-by-two constant matrix B defined by

$$a=\sum_{j=1}^{2}(1-nn\bullet)\{J_j^{-1}[R_j\times n])\times R_j\}, \qquad (4.11)$$

$$B\hat{u}_T=-\mu(M_1^{-1}+M_2^{-1})\hat{u}_T$$

$$+\mu\sum_{j=1}^{2}(1-nn\bullet)\{R_j\times(J_j^{-1}[R_j\times\hat{u}_T]). \qquad (4.12)$$

Then (4.9) takes the form

$$\frac{du_T}{d\tau}=a+B\frac{u_T}{|u_T|}. \qquad (4.13)$$

Next we write $u_T=(\rho\cos\theta,\ \rho\sin\theta)$ where $\rho=|u_T|$ and θ is an angle in the tangent plane at the point of impact. Then we can rewrite (4.13) as

$$\frac{d\rho}{d\tau}=a_1\cos\theta+a_2\sin\theta+B_{11}\cos^2\theta$$

$$+(B_{12}+B_{21})\sin\theta\cos\theta+B_{22}\sin^2\theta\equiv g(\theta), \qquad (4.14)$$

$$\rho\frac{d\theta}{d\tau}=-a_1\sin\theta+a_2\cos\theta-(B_{11}-B_{22})\sin\theta\cos\theta$$

$$-B_{12}\sin^2\theta+B_{21}\cos^2\theta\equiv h(\theta) \qquad (4.15)$$

By eliminating ρ from these two equations we get an equation for θ which has the solution

$$\tau=\rho(0)\int_{\theta(0)}^{\theta}\frac{1}{h(\theta')}e^{\int_{\theta(0)}^{\theta'}g(\theta'')/h(\theta'')d\theta''}d\theta'. \qquad (4.16)$$

Then we find

$$\rho(\tau)=\rho(0)e^{\int_{\theta(0)}^{\theta(\tau)}g(\theta')/h(\theta')d\theta'} \qquad (4.17)$$

The solutions (4.16) and (4.17) determine $\theta(\tau)$ and $\rho(\tau)$, and thus $u_T(\tau)$. They are valid provided that $h[\theta(0)]\neq 0$, because the integrals diverge in that case. However, then $\theta(\tau)=\theta(0)$ and ρ increases linearly in τ, as we shall see in section 6.

These results hold provided that $u_T(\tau)$ does not vanish. If $u_T(\tau^*)=0$ for some τ^* satisfying $\tau(1)>\tau^*\geq 0$, we much check to see if $u_T(\tau)$ will remain zero for $\tau\geq\tau^*$. To do so we set $du_T/d\tau=0$ in (4.6) and solve for f. If the solution f satisfies (3.6), i.e., if $|f|\leq\mu$, then $u_T(\tau)$ will remain zero and f will remain constant for $\tau\geq\tau^*$. Then in the preceding equations (3.8)-(3.10) we must replace

$$\mu\int_0^{(1+e)\tau_0}\hat{u}_T(\tau)d\tau \text{ by } \mu\int_0^{\tau^*}\hat{u}_T(\tau)d\tau+[(1+e)\tau_0-\tau^*]f. \qquad (4.18)$$

A corresponding change must be made in (4.8) if $\tau^*<\tau_0$.

5 Summary of the Theory

We can now summarize our theory as follows:

(1) Calculate the slip velocity $u_T(\tau)$ from (4.16) and (4.17) for $0\leq\tau\leq(1+e)\tau_0$, with τ_0 the solution of (4.8).

(2) Evaluate the jumps $[U_j]$ and $[\Omega_j]$ from (3.9) and (3.10).

(3) If $u_T(\tau^*)=0$, check to see whether or not $u_T(\tau)$ remains zero for $\tau>\tau^*$ by the procedure in the preceding paragraph. If u_T does remain zero make the replacement (4.18) in the expressions for the jumps, and a corresponding change in (4.8) if $\tau^*<\tau^0$.

The theory involves two material constants: the coefficient of the restitution e and the coefficient of friction μ. The main difficulty in using this theory arises in calculating u_T. We shall now illustrate how to use the theory in some special cases.

6 Applications

***(a)* Constant Slip Direction.** Suppose that (4.9) and (4.10) yield

$$\frac{du_T(0)}{d\tau}\text{ parallel to }u_T(0). \qquad (6.1)$$

Then $u_T(\tau)$ remains parallel to $u_T(0)$ and the slip direction $\hat{u}_T(\tau)$ stays constant until $u_T(\tau)$ vanishes at $\tau=\tau^*$. Thus the solution of (4.9) is

$$u_T(\tau)=u_T(0)+\tau\frac{du_T(0)}{d\tau},\quad 0\leq\tau\leq\tau^*. \qquad (6.2)$$

Then (4.7) also yields the simple result

$$u_N(\tau)=u_N(0)+\tau\frac{du_N(0)}{d\tau},\quad 0\leq\tau\leq\tau^*. \qquad (6.3)$$

Now we can use (6.3) in (3.3) provided that $\tau^*\geq\tau_0$, and solve for τ_0 with the result

$$\tau_0=-\frac{u_N(0)}{du_N(0)/d\tau},\quad \tau_0\leq\tau^*. \qquad (6.4)$$

To find τ^* we use (6.2) to write $u_T(\tau^*)=0$, multiply by $u_T(0)$ and solve to get

$$\tau^*=-u_T^2(0)/u_T(0)\bullet du_T(0)/d\tau. \qquad (6.5)$$

If $\tau^*>(1+e)\tau_0$, so that u_T does not vanish during the collision, the jumps (3.9) and (3.10) simplify to

$$[U_j]=-(-1)^jM_j^{-1}(1+e)\frac{u_N(0)}{du_N(0)/d\tau}(n-\mu\hat{u}_T(0)), \qquad (6.6)$$

$$[\Omega_j]=-(-1)^j(1+e)\frac{u_N(0)}{du_N(0)/d\tau}J_j^{-1}[R_j\times(n-\mu\hat{u}_T(0))]. \qquad (6.7)$$

The right side of (6.6) is $(-1)^jM_j^{-1}I$, so we see that the magnitude of the frictional impulse is μ times that of the normal impulse, in agreement with Whittaker's theory.

(*b*) Slip Direction Changes Once. Let us now suppose that (6.1)-(6.5) hold, but that

$$\tau_0 \leq \tau^* < (1+e)\tau_0. \tag{6.8}$$

Then for $\tau > \tau^*$, $du_T/d\tau$ and $\hat{u}_T$ will have constant values $du_T(\tau^*+)/d\tau$ and $\hat{u}_T(\tau^*+)$. They can be obtained from (4.9) by using the fact that $\hat{u}_T(\tau^*+)$ is parallel to $du_T(\tau^*+)/d\tau$. Then (3.8) yields

$$I = (1+e)\tau_0 n - \mu\tau^*\hat{u}(0) - \mu[(1+e)\tau_0 - \tau^*]\hat{u}_T(\tau^*+). \tag{6.9}$$

The jumps $[U_j]$ and $[\Omega_j]$ are given by (2.9) and (2.10) with this value of I.

In the special case of a rigid pendulum striking a fixed surface, we have $\tau^* = \tau_0$ because both the normal and tangential relative velocities become zero at the same instant. Furthermore, $\hat{u}_T(\tau^*+) = -\hat{u}_T(0)$ because the tangential velocity reverses its direction at τ^*. Then (6.9) becomes

$$I = (1+e)\tau_0 n - (1-e)\mu\tau_0\hat{u}_T(0). \tag{6.10}$$

In this case the magnitude of the frictional impulse is $\mu(1-e)/(1+e)$ times that of the normal impulse. This is smaller than the ratio μ which holds when the slip direction remains constant throughout the collision.

Acknowledgments

Research supported by the Office of Naval Research, the Air Force Office of Scientific Research, the Army Research Office, and the National Science Foundation.

References

Kane, T. R., 1984, "A Dynamics Puzzle," *Stanford Mechanics Alumni Club Newsletter*, p. 6.

Kilmister, C. W., and Reeve, J. E., 1966, *Rational Mechanics*, Longsmans, London.

Whittaker, E. T., 1904, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Cambridge, p. 232.

E. H. Dowell
Dean of Engineering.
Duke University,
Durham, NC 27706

C. Pezeshki
Graduate Student.
Department of Mechanical Engineering
and Materials Science,
Duke University,
Durham, NC 27706

On the Understanding of Chaos in Duffings Equation Including a Comparison With Experiment

The dynamics of a buckled beam are studied for both the initial value problem and forced external excitation. The principal focus is on chaotic oscillations due to forced excitation. In particular, a discussion of their relationship to the initial value problem and a comparison of results from a theoretical model with those from a physical experiment are presented.

Introduction, Background and Motivation

In the present paper the following equation is studied

$$\ddot{A}+\gamma\dot{A}-\frac{A}{2}(1-A^2)=F(\tau) \qquad (1)$$

This is the particular form of Duffings equation (with a negative linear stiffness) studied by Moon [1]. It is known this equation has solutions with chaotic oscillations under certain conditions. Here we extend the earlier work on Duffings equation and provide an improved understanding of why the chaotic oscillations occur by first considering the initial value problem when $F\equiv 0$. These results are of substantial interest in their own right as well as leading to additional understanding of why chaotic oscillations occur. The present theoretical results are also compared to the physical experiments of Moon. The opportunity to compare the present theoretical results with the experimental data of Moon is also an important motivation for this work.

A physical model is helpful in the interpretation of Duffings equation. Following previous authors [1-4], we interpret equation (1) as describing a one mode oscillation of a buckled beam under the action of a prescribed lateral external force, $F(\tau)$. Other physical systems may also be described by this equation, but they will not be discussed here. As may be seen from equation (1), when $F\equiv 0$ there are three static equilibrium solutions: $A=0$, $+1$ and -1. It is easily shown that the first of these (an unbuckled beam) is dynamically unstable and the latter two are stable with respect to *infinitesimal* disturbances. It is of great help in understanding the onset of chaos to consider next the stability of the static equilibria, $A=+1$ or -1, with respect to *finite* disturbances. This is done in the following section of the paper and then chaos due to a harmonic force, $F(\tau)$, is considered directly in subsequent sections.

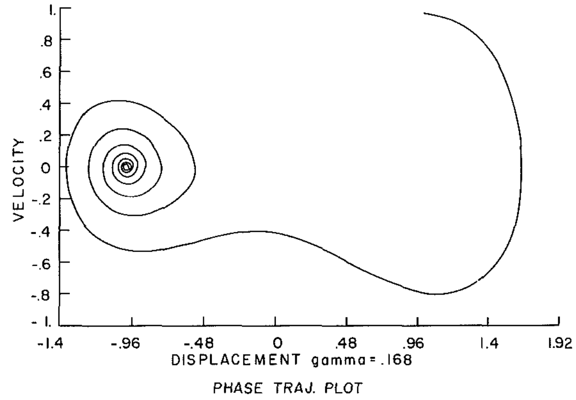


Fig. 1 Phase plane trajectory

The Initial Value Problem for the Homogeneous Duffings Equation ($F\equiv 0$)

It is helpful to think first in physical terms. Consider the buckled beam at rest in one of its stable (with respect to *infintesimal* disturbances) static equilibria, say $A=+1$. With prescribed initial conditions,

$$A(\tau\equiv 0)=A_0$$
$$\dot{A}(\tau\equiv 0)=\dot{A}_0$$

consider the transient solution and the final steady state solution as $\tau\rightarrow\infty$. Obviously $A(\tau\rightarrow\infty)\rightarrow +1$ or -1. The question is which of these two solutions is the correct one for given A_0, $\dot{A}_0$. As shall be seen, the answer is in a certain sense unknowable (or to use a more technical term, *uncertain*). Once the reason for this is understood, the occurence of chaos for certain $F\neq 0$ becomes more understandable, perhaps even expected.

It is possible to construct a diagram (which is called a *shell plot* because of its appearance) that summarizes compactly

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami, Fla., November 17-21, 1985 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 11, 1984; final revision, April 30, 1985. Paper No. 85-WA/APM-27.

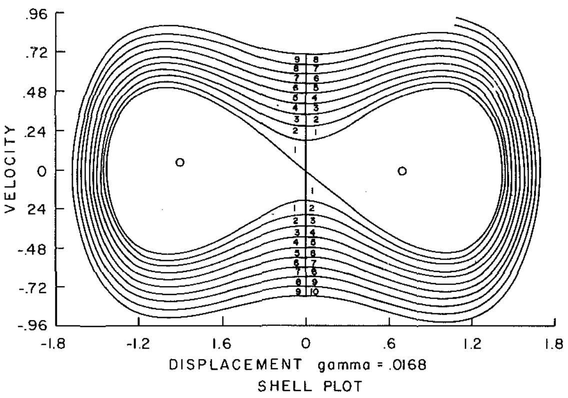


Fig. 2 Shell plot

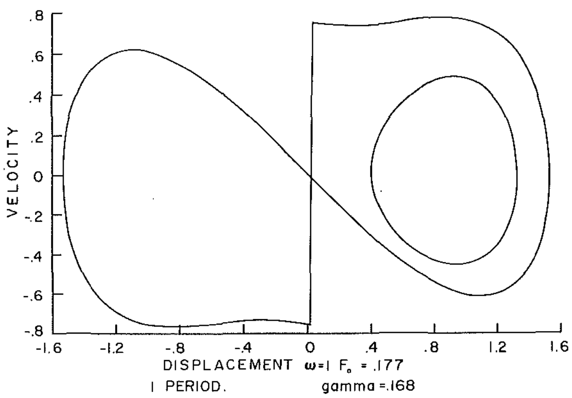


Fig. 3(*a*) Phase plane trajectory (1 period motion)

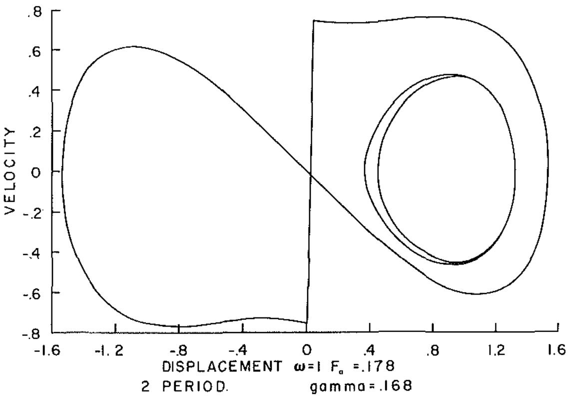


Fig. 3(*b*) Phase plane trajectory (2 period motion)

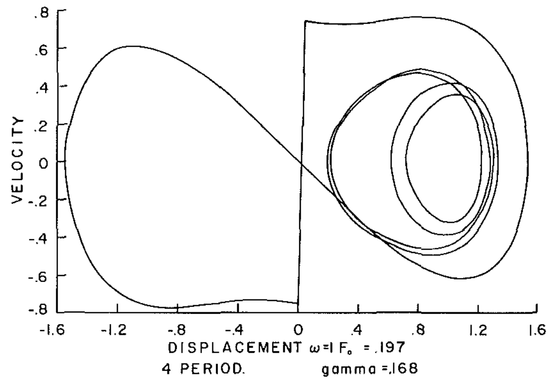


Fig. 3(c) Phase plane trajectory (4 period motion)

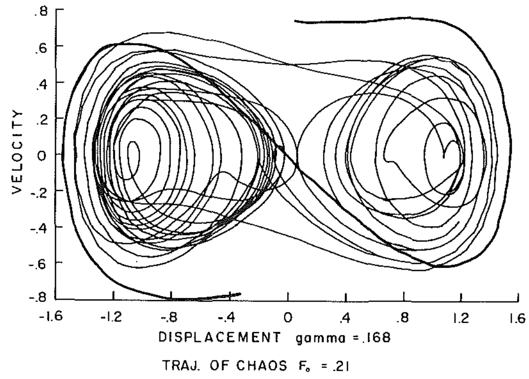


Fig. 3(*d*) Phase plane trajectory (chaotic motion)

which of the two static equilibria solutions will be reached for a given set of initial conditions, A_0, $\dot{A}_0$.

To anticipate the form of the shell plot, consider a specific example of initial conditions and the subsequent solution trajectory. This is shown as a phase plane trajectory in Fig. 1 for $A_0 = 1$ and $\dot{A}_0 = 1$. Because $\dot{A}_0 > 0$, A increases for small time but then decreases for larger time because of the nonlinear restoring stiffness. Indeed, A subsequently becomes negative (the beam moves from one buckled configuration, $A = +1$, to the other, $A = -1$, and beyond). The damping term, $\gamma\dot{A}$, leads to dissipation of energy; thus the beam does not continue to oscillate between and about the two static buckled equilibria, but instead spirals into one of them as $\tau \rightarrow \infty$.

From such phase plane trajectories, one can construct the shell plot, which shows the final state of the system as $\tau \rightarrow \infty$, $A = +1$ or -1, for given initial conditions, A_0 and $\dot{A}_0$. This is shown (partially) in Fig. 2. Here $\dot{A}_0$ is plotted versus A_0 and various regions are identified with integral values, 0, 1, 2, 3, 4, . . . Note there are two disjoint regions associated with each integer value. Consider first the integer zero (0) regions. For definiteness consider the region where $A_0 \geq 0$. If the system starts with A_0, $\dot{A}_0$ within the zero region, the solution spirals into $A = +1$ as $\tau \rightarrow 0$ and crosses the $A = 0$ axis zero times. Consider now the 1 region. A solution begun there moves clockwise and crosses the $A = 0$ axis one time and enters the zero region for $A < 0$. Once there, it spirals into $A = -1$ as $\tau \rightarrow \infty$. To firmly establish the pattern, finally consider the 2 region. For initial conditions in the 2 region, the phase plane trajectory moves clockwise, crosses the $A = 0$ axis the first time and moves into the 1 region for $A < 0$. It then continues to move clockwise and crosses the $A = 0$ axis a second (and final time) and moves into the 0 region for $A > 0$. Once there it spirals into $A = +1$ as $\tau \rightarrow \infty$.

The pattern is now clear. For $A_0 > 0$, initial conditions in an even integer region reach a final state of $A = +1$. Those in an odd integer region reach a final state of $A = -1$. The integer number corresponds to the number of crossings of the $A = 0$ axis during the completion of the motion (phase plane trajectory). For $A_0 < 0$, a similar sequence of events occurs. For initial conditions which lie precisely on a shell boundary, the final configuration would be A, $\dot{A} \rightarrow 0$ as $\tau \rightarrow \infty$. In practice this will never occur, of course, because the shell boundary curves are of vanishing thickness.

It is interesting to note that a shell plot of any finite extent can be constructed from a single artfully chosen phase plane trajectory, once the zero region is known. The latter region is readily determined by direct calculation.

Now comes the central point. If there is *sufficient* uncertainty in the values of the initial conditions, A_0, $\dot{A}_0$, it is clear from an examination of the shell plot that the final system state, $A = +1$ or -1, is unpredictable, unknowable or uncertain. This point is made all the more powerful by noting that as the damping becomes even smaller the width of each region of the shell plot (excluding the 0 region) becomes even smaller and vanishes as $\gamma \rightarrow 0$. Hence for any (finite) uncertainty in A_0 or $\dot{A}_0$ the final system state is unpredictable as $\gamma \rightarrow 0$.

Two additional points are worthy of note in concluding this discussion. First, the boundary contours of the shell plot are curves of essentially constant total (kinetic plus potential) energy. Secondly as $\gamma \rightarrow 0$, the boundary curves for the two zero regions correspond to the separatrix of the undamped system.

Although not concerned with chaos, the reader will find Ref. 5 on the initial value problem of interest.

The Continuous Oscillation Problem for the Inhomogenous Duffings Equation ($F \neq 0$)

Here a simple harmonic external force is considered,

$$F = F_0 \sin \omega t \tag{2}$$

where F_0 is the force amplitude and ω its frequency of excitation. This is not the only force-time history which might be studied. It is, perhaps, the simplest periodic force.

As the reader may note, the initial value problem previously studied can be also thought of as an external force problem. For example, an initial velocity, $\dot{A}_0$, corresponds to an impulsive force,

$$F(\tau) = \dot{A}_0 \delta(\tau) \tag{3}$$

This suggests that a study of continual impulses, periodically or randomly spaced in time, would be of interest. Nevertheless, only a simple harmonic force will be considered here.

The response of the system will be considered first for fixed frequency, $\omega = 1$, and increasing force amplitude, F_0. The frequency is normalized by the small amplitude natural frequency about the beam buckled equilibrium. For F_0 sufficiently small, it is expected that the response of the system will be a simple harmonic oscillation about one or the other of the two static equilibria, $A = +1$, or -1. For definiteness the initial conditions, $A_0 = 1$, $\dot{A}_0 = 0$ are chosen so that for small F_0 the harmonic response oscillation is about $A = +1$. It is anticipated that, as F_0 increases and the response phase plane trajectory approaches the zero region boundary of the shell plot, interesting response behavior will occur.

Note that for small F_0 the phase plane trajectory is an ellipse indicating a simple harmonic response oscillation. As F_0 increases additional harmonics beyond the fundamental are detected and the phase plane trajectory is distorted from a simple ellipse. See Fig. 3(*a*) for the result for $F_0 = 0.177$. Also shown for reference is the boundary for the zero region from the previously discussed shell plot.

For $0 \leq F_0 \leq .177$ the response is termed 1 period motion. By that is meant, as the force oscillates through one period, the response also oscillates through one period. For $F_0 = .178$, however, as the force oscillates through one period the response oscillates through only one half a period. For the response to go through one period, the force must oscillate through two periods. Thus this is called 2 period motion. See Fig. 3(*b*). This change from 1 to 2 period response as $F_0 = 0.177 \rightarrow 0.178$ appears to be a bifurcation.

At a higher F_0, 4 period motion occurs. See Fig. 3(*c*) for example, and at yet higher F_0, 8 and 16 period motion occurs. Holmes [4] has suggested that 32, etc. period motions occur as F_0 increases further. This may well happen but this behavior has not been observed by the present authors. Possibly this is because the range of F_0 over which the higher period motions occur is very small. This period doubling behavior has been previously described and discussed by Feigenbaum [6] in a more general context.

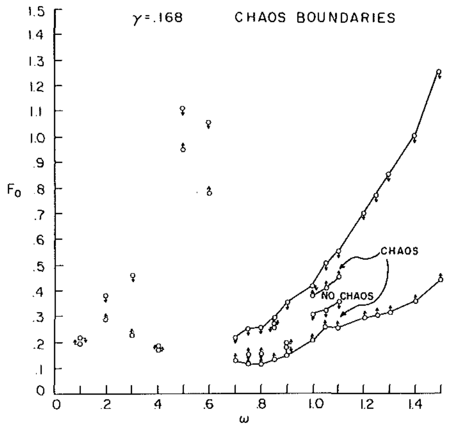


Fig. 4(*a*) Force amplitude versus frequency of excitation

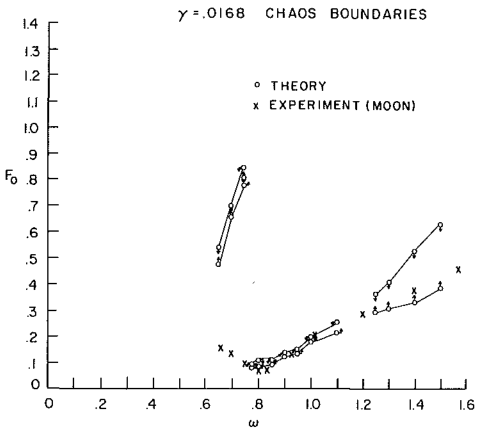


Fig. 4(*b*) Force amplitude versus frequency of excitation

For $F_0 \geq 0.205$ chaos is observed, i.e., no periodicity is apparent. See Fig. 3(*d*) which gives results for $F_0 = 0.21$. As Holmes has indicated, for yet higher F_0 the chaos no longer appears and periodicity returns.

It is clear that for F_0 just below the value where chaos first appears the periodic response phase plane trajectory approaches and slightly penetrates the boundary of the zero region shell plot. See Fig. 3(*c*). Moreover it is clear that for this frequency, $\omega = 1$, chaos occurs when the motion is no longer about only one of the static equilibria points, say $A = +1$, but instead encircles both, $A = +1$ and -1. This is called snap-buckling. These observations suggest that the onset of chaos can be associated with periodic motions which penetrate the zero region boundary and thus lead subsequently to motion about both static equilibria points. Moon in an earlier paper [1] suggested a more restricted notion of this sort when he took as an empirical criterion for the onset of chaos that the periodic response maximum velocity (in his calculation he assumed one period motion) must exceed the maximum velocity of the system separatrix. Recall it has been shown here that the zero region boundary of the shell plot corresponds to the system separatrix as $\gamma \rightarrow 0$.

It is speculated, though it remains to be shown, that as $\gamma \rightarrow 0$ *any* penetration of the zero region boundary by the phase plane trajectory leads to chaos. For small, but finite, γ the phase plane trajectory may (slightly) penetrate the zero region boundary before chaos occurs. Hence the penetration of the zero region boundary by a phase plane trajectory at a certain force level may provide a lower bound criterion for the onset of chaos, at least for ω near 1.

For excitation frequencies well away from resonance, in particular for $\omega << 1$, chaos was found to occur even without

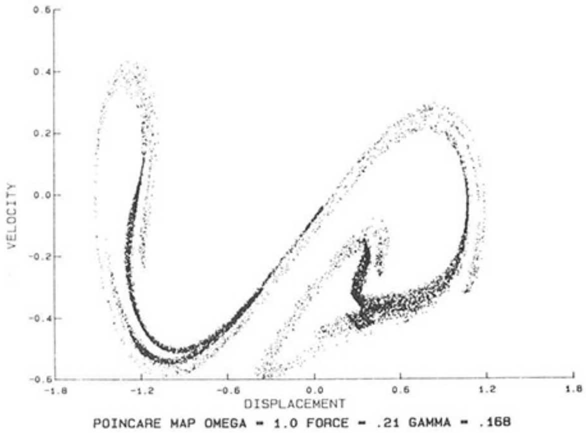


Fig. 5(*a*) Poincare map; $\omega = 1, \gamma = .168, F_0 = .21$

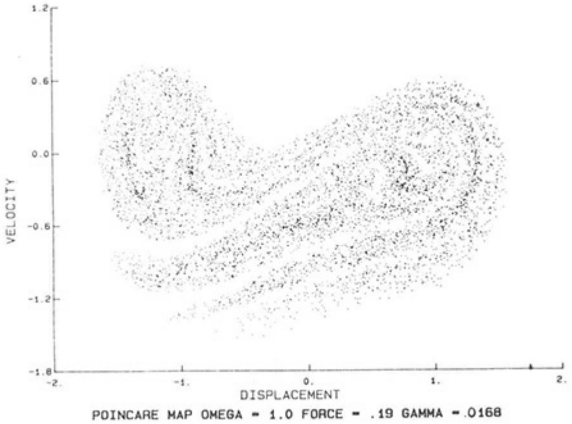


Fig. 5(*b*) Poincare map; $\omega = 1, \gamma = .0168, F_0 = .09$

snap-buckling. The minimum F_0 for the onset of chaos occurs for $\omega \cong 0.85$. This is the characteristic frequency for free vibration about both static equilibrium points.

Comparison of Theory and Experiment

Calculations similar to those described in the previous section have been carried out for several excitation frequencies, ω. From these a summary plot may be constructed of the force required to cause snap buckling, period doubling, or chaos versus the excitation frequency. Such a plot for the onset of chaos is shown in Fig. 4(*a*) and 4(*b*) for $\gamma = 0.168$ and 0.0168, respectively. The uncertainty in the data is less than a diameter of a circle.

Time integrations using the Runge-Kutta method were performed for frequency values ranging from 0.1 to 1.5 for varying force levels at damping levels of 0.168 and 0.0168. Principal lower and upper force chaos boundaries were found for a discrete series of frequencies by incrementing the first from zero until chaos was observed. See Fig. 4. Force increments of 0.01 and, where necessary, 0.001, were used. All results shown are for simulations started from the initial condition values of one for the displacement and zero for the velocity. Other initial conditions were tried, but no observable effect of initial conditions on chaos boundaries was detected. Of course, the time history details do depend upon initial conditions, particularly in the chaotic regime.

The types of chaos found in the simulations varied from frequency to frequency. However, the form of the Poincare map for a given set of frequencies tended to be of the same type when the lower force steady state periodic phase plane portraits were shaped the same and possessed the same periodicity and when the corresponding upper force portraits were also the same across the frequency band. The lower and upper plane portraits were not identical. Such identification or association of results at several frequencies has led us to connect some data points in Fig. 4(*a*) and 4(*b*) by straight lines. Such results suggest that chaos lies in fragmented pockets in the force-frequency plane. These pockets can also have smaller pockets of force-frequency combinations within them that can lead to periodic phase plane orbits.

It is apparent from the situations simulated that chaos can assume many forms. Some of the entries into the chaotic region in the force-frequency plane for certain frequency values were precipitated by beam snap-through; others were not. At certain frequencies, the system went into chaos even before the beam snapped through. Period doubling was observed at some frequencies, e.g., $\omega = 0.9$ and 1.0, but not at others. Chaos appeared at all frequencies simulated, though this does not preclude the possibility of finding frequencies that are chaos-free. A simple boundary cannot be drawn in the force-frequency plane above which there is guaranteed chaos; in fact, the simulations point to the opposite.

The simulations run at high damping levels gave the same qualitative answer as the ones run at low damping. The major difference is that the width of the chaotic band in the force-frequency plane for the low damping case is much narrower than its counterpart for higher damping. The limit of zero damping may be pathological. Another difference is that the higher damping case allows a much richer selection of equilibrium periodic phase plane orbits. As the damping is decreased, the Poincare maps also lose their ordered structure.

The correlation between data obtained from simulation and the data obtained by Moon from his physical experiment also appears to be generally good. See Fig. 4(*b*). The principal difference is that at $\omega \cong 0.65$, the simulation predicts chaos at much higher force levels, $F_0 \neq 0.45$-0.55, than those observed by Moon in his physical experiment, $F_0 \cong .17$. It is worthy of note that the simulation predicts that snap-through of the beam occurs at $F_0 \cong .12$ and it is possible that this snap-through was identified as chaos in the physical experiment. At higher frequencies, snap-through and chaos occur at force levels which are much closer together. Of course other factors may enter in including the effects of higher beam modes.

For brevity, we have not shown the large number of phase-plane portraits and Poincare maps that have been calculated. The authors would be pleased to make these available to other investigators who may wish to extend the present study. In Fig. 5(*a*) and 5(*b*), two representative Poincare maps are shown for $\omega = 1$ and the two damping values used in this study.

Conclusions and Future Work

Among the conclusions reached based upon the present work are the following:

(1) The initial value problem for a second order homogeneous system is a key to the understanding of higher order systems, including the inhomogeneous second order system.

(2) Chaos is not difficult to find by numerical simulation, however a Feigenbaum (period doubling) sequence may be difficult to find for some parameter conditions.

(3) A comparison between theoretical results for Duffing's equation and (physical) experiments for a buckled beam shows generally good agreement.

(4) Future theoretical studies should consider

- investigating the limit as damping approaches zero; setting the damping identically zero may lead to pathological results
- multimode convergence studies (based upon the results from panel flutter calculations [7, 8], it is expected good convergence will occur)

(5) Future experimental work should attempt to study
  - various damping levels
  - determination of period doubling conditions
  - identification of entire pockets of chaos

Acknowledgment

This work was supported, in part, by NSF Grant MEA-8315193 and AFOSR Grant 83-0346. Drs. Elbert Marsh and Anthony Amos, respectively, are the technical monitors. The authors would like to thank Mr. Michael D'Antonio for his help with the computations. They would also like to thank Professor Francis C. Moon for helpful discussions.

References

1 Moon, F. C., "Experiments on Chaotic Motions of a Forced Nonlinear Oscillator: Strange Attractors," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 638–644.

2 Holmes, P. J., and Moon, F. C., "Strange Attractors and Chaos in Nonlinear Mechanics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 108, 1983, pp. 1021–1032.

3 Tseng, W.-Y., and Dugundji, J., "Nonlinear Vibrations of a Buckled Beam Under Harmonic Excitation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 38, 1971, pp. 467–476.

4 Holmes, P. J., "A Nonlinear Oscillator with a Strange Attractor," *Phil. Trans. of Royal Society*, London, Vol. 292, 1979, pp. 419–448.

5 Reiss, E. L., and Matkowsky, B. J., "Nonlinear Dynamic Buckling of a Compressed Elastic Column," *Quart. Appl. Math.*, Vol. 29, 1971, pp. 245–260.

6 Feigenbaum, M. J., "Quantitative Universality for a Class of Nonlinear Transformations," *J. Stat. Physics*, Vol. 19, 1978, pp. 25–52.

7 Dowell, E. H., "Flutter of a Buckled Plate as an Example of Chaotic Motion of a Deterministic Autonomous System," *J. Sound Vibration*, Vol. 85, 1982, pp. 333–344.

8 Dowell, E. H., "Observation and Evolution of Chaos in an Autonomous System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, 1984, pp. 333–344.

M. Ahmadian
Assistant Professor,
Mechanical Engineering Department,
Clemson University,
Clemson, S.C. 29634-0921
Assoc. Mem. ASME

D. J. Inman
Associate Professor,
Mechanical and Aerospace
Engineering Department,
State University of New York at Buffalo,
Buffalo, N.Y. 14260
Assoc. Mem. ASME

Some Stability Results for General Linear Lumped-Parameter Dynamic Systems

A technique is presented for stability of equilibrium of general, linear, lumped-parameter dynamic systems. Liapunov functions are used to develop stability conditions that are direct in terms of the mass, damping, and stiffness matrices. The significance of what is presented here is twofold. First, this technique can be applied to general asymmetric systems. Second, it offers direct conditions that can easily be programmed on a digital computer to handle high-order systems. Many previously developed results, such as the KTC theorem and its extensions, are mentioned. Next, it is shown that the present study may provide broader applications because general systems are included and a more convenient approach is offered. Examples are used to illustrate the validity and applications of the presented results.

Introduction

Stability of the equilibrium of linear, lumped-parameter dynamic systems described by the vector differential equation

$$M\ddot{\mathbf{x}}+C\dot{\mathbf{x}}+K\mathbf{x}=0 \qquad (1)$$

has received extensive attention for many decades. Here, M, C, and K are $n \times n$ real matrices referred to as mass, damping, and stiffness matrices, and the vector $\mathbf{x}$ is a real $n \times 1$ vector of generalized coordinates. Routh in his famous essay of 1877 [1] solved the problem of determining necessary and sufficient condition for the asymptotic stability of the general form of the foregoing systems, where no restriction is placed on the coefficient matrices. However, his conditions require the knowledge of the coefficients in the characteristic polynomial of the system and the evaluation of certain determinants, which may be rather difficult to apply when the order of the system is at all large. Another approach is the equivalent result of Liapunov [2, 3]. This method also involves solving a $2n$th order matrix equation and the evaluation again of some determinants. As a result, alternative methods such as those which provide simpler conditions directly in terms of the coefficient matrices prove to be more attractive.

The development of such stability conditions dates back to the time of Lord Rayleigh [4]. He proved that for the systems with positive definite mass and stiffness matrices if the damping matrix is positive definite, the equilibrium is asymptotically stable. The equilibrium becomes unstable if either the damping or the stiffness matrix is not nonnegative (i.e., C or K has at least one negative eigenvalue). In later studies it was shown that, under certain conditions, the equilibrium can still be asymptotically stable when C is only positive semidefinite (denoted $C \geqq 0$). Moran [5], illustrated that the necessary and sufficient conditions for the system (1) to be asymptotically stable are that none of the eigenvectors of the corresponding conservative system lies in the null space of the damping matrix. In another study by Walker and Schmitendorf [6] a controllability approach was taken to prove that the equilibrium is asymptotically stable if and only if the $n^2 \times n$ matrix

$$\begin{bmatrix} C \\ C(M^{-1}K) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ C(M^{-1}K)^{n-1} \end{bmatrix} \qquad (2)$$

has rank n.

If gyroscopic forces (e.g., Coriolis, Lorentz forces) are added to a symmetric system, then the equation of motion modifies to the more general form

$$M\ddot{\mathbf{x}}+(C+G)\dot{\mathbf{x}}+K\mathbf{x}=0 \qquad (3)$$

where G is a skew symmetric matrix (i.e. $G = -G^T$). In this case the well-known Kelvin-Tait-Chetaev (KTC) theorem can be applied. According to this theorem if C is positive definite (denoted $C > 0$), then the system (3) is asymptotically stable (unstable) if the corresponding nondissipative nongyroscopic system, which can be characterized here by $C = G = 0$, is

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami Beach, Fla., November 17-22, 1985, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September, 1984; final revision, February, 1985. Paper No. 85-WA/APM-20.

stable (unstable) [7]. In a later study, Zajac introduced the notion of pervasive-damping,[1] associated with positive semidefinite damping matrix, and proved that for a pervasively damped system, the equilibrium is still asymptotically stable if the corresponding nondissipative nongyroscopic system is stable [8, 9]. This extends the application of the KTC theorem, particularly in space mechanics. Other studies, such as the ones by Müller [13] and Hughes and Gardner [14] have suggested an "observability" approach for determining the stability of a gyroscopic system with positive semidefinite damping matrix.

Finally, considering the effects of constraint damping[2] and/or circulatory forces[3] modify the equation of motion to its most general form, described by

$$M\ddot{\mathbf{x}} + (C+G)\dot{\mathbf{x}} + (K+E)\mathbf{x} = 0 \tag{4}$$

where E is a skew symmetric matrix. Obviously, this class of systems does not fall within the scope of the *KTC* theorem and its extensions. Mingori [18] generalized the *KTC* theorem to include these systems. He demonstrated that if the matrices in equation (4) satisfy the following three conditions:

$$EC^{-1}M = MC^{-1}E \tag{5a}$$

$$EC^{-1}G = GC^{-1}E \tag{5b}$$

$$EC^{-1}K = KC^{-1}E \tag{5c}$$

and if a matrix H is defined as

$$H = EC^{-1}MC^{-1}E - GC^{-1}E + K \tag{6}$$

then the equilibrium of (4) will be asymptotically stable if all of the eigenvalues of H are negative. The equilibrium is unstable if at least one of the eigenvalues of H is positive. Although this approach includes a broader class of systems than the *KTC* theorem, it has limited applications due to the restrictive nature of the conditions in equation (5). However, if one is willing to sacrifice generality for simplicity, it is possible to develop less restrictive conditions for a subclass of general systems. According to [19], if the asymmetric matrices $(C+G)$ and $(K+E)$ are symmetrizable and possess positive eigenvalues, then the stability is determined by stability of the matrix

$$P = M^{-1}(C+G)S + S(C+G)^T M^{-T} \tag{7}$$

where

$$S = S^T = R(K+E) \tag{8}$$

and R is an arbitrary positive definite matrix. Due to the assumption on symmetrizability of $(K+E)$, it is always possible to find a matrix R to satisfy the equation (8) [20]. However, to the best of our knowledge, no simple, effective method for finding R has been introduced yet. The methods available presently require long, tedious computations which prove to be inefficient for most cases.

The main object of the present study is to develop stability conditions that include general asymmetric systems, avoid restrictive conditions, and can be programmed on a digital computer to handle systems with many degrees of freedom.

Results

Assuming a nonsingular mass matrix (i.e., det $(M) \neq 0$), one can reformulate the general linear system presented in the equation (4) as

$$\ddot{\mathbf{x}} + A\dot{\mathbf{x}} + B\mathbf{x} = 0 \tag{9}$$

where

$$A = M^{-1}(C+G) \tag{10a}$$

and

$$B = M^{-1}(K+E) \tag{10b}$$

are both real asymmetric matrices. Consider the function V given by

$$V = \mathbf{x}^T(B+B^T)\mathbf{x} + [\mathbf{x}^T A^T + \dot{\mathbf{x}}^T][A\mathbf{x} + \dot{\mathbf{x}}] + \dot{\mathbf{x}}^T\dot{\mathbf{x}}. \tag{11}$$

Differentiating V and substituting

$$\ddot{\mathbf{x}} = -A\dot{\mathbf{x}} - B\mathbf{x} \tag{12}$$

into $\dot{V}$ results in

$$\dot{V} = -\mathbf{x}^T(A^T B + B^T A)\mathbf{x} - \mathbf{x}^T(B^T - B)\dot{\mathbf{x}} - \dot{\mathbf{x}}^T(B - B^T)\mathbf{x} - \dot{\mathbf{x}}^T(A + A^T)\dot{\mathbf{x}}$$

which can be presented as

$$\dot{V} = -\mathbf{z}^T Q \mathbf{z}$$

where

$$Q = \begin{bmatrix} Q_1 & Q_2{}^T \\ Q_2 & Q_3 \end{bmatrix}$$

$$Q_1 = Q_1^T = A^T B + B^T A \tag{13a}$$

$$Q_2 = Q_2^T = B - B^T \tag{13b}$$

$$Q_3 = Q_3^T = A + A^T \tag{13c}$$

and

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix}$$

Lemma 1. The matrix Q is positive definite if and only if Q_1 and the matrix

$$Q_3 - Q_2 Q_1{}^{-1} Q_2^T \tag{14}$$

are positive definite.

Lemma 2. The matrix Q is positive semidefinite if and only if Q_1 is positive definite and the matrix (14) is positive semidefinite.

The proof is established in the Appendix.

Theorem 1. For the system (9), the equilibrium is stable if the symmetric matrices

$$B + B^T \tag{15}$$

and Q_1 are positive definite and the symmetric matrix (14) is positive semidefinite.

Proof: By assumption, V is positive definite and $\dot{V}$ is nonpositive. Therefore, according to [2], (theorem 4.1, pp. 14), the equilibrium is stable.

Theorem 2. The equilibrium is asymptotically stable if the symmetric matrices (14), (15), and Q_1 are all positive definite.

Proof: Considering the assumptions, V is positive definite and $\dot{V}$ is nonpositive (i.e., $V > 0$ and $\dot{V} \leqq 0$). In addition, the function $\dot{V}$ cannot remain zero, unless $\mathbf{x}$ and $\dot{\mathbf{x}}$ are both zero. Consequently, the results in [21] suggests that every motion of the system (9) is either asymptotically stable (tends to $\mathbf{z} = 0$) or unstable (unbounded). But no motion can be unstable, since according to theorem 1, it is at least globally stable. Therefore, the equilibrium is asymptotically stable.

Theorem 3. The equilibrium is unstable if (15) is not

---

[1] Damping is pervasive if the dissipative function $R = 1/2\ \dot{\mathbf{x}}^T C\dot{\mathbf{x}}$ is nonnegative and can be identically zero only when the system is at the equilibrium state for $t > 0$ [10]. Roberson [11] and Connell [12] have proposed different methods for determining whether or not a system is pervasively damped when $C \geqq 0$.

[2] Constraint damping results when the linear approximation of the dissipation forces are derivable from a dissipative function of the form $R_c = 1/2\,\dot{\mathbf{x}}^T C\dot{\mathbf{x}} + \dot{\mathbf{x}}^T E\mathbf{x}$. This type of dissipation may happen in a dissipative, gravity-oriented satellite [15]. A complete discussion of this type of dissipation can be found in [15–18].

[3] Circulatory forces are linear generalized forces which can be expressed by the vector $\mathbf{f} = -E\mathbf{x}$, where E is a skew symmetric matrix. Such forces arise in contemporary mechanical, aeronautical, and missle engineering.

nonnegative (i.e., (15) is negative definite, negative semidefinite, or indefinite) and Q_1 and (14) are both positive definite.

Proof: The assumption on (15) implies that V may be negative. Furthermore, the assumption on Q_1 and (14) indicates that $\dot{V}$ is nonpositive and is identically zero if and only if $\mathbf{x}$ and $\dot{\mathbf{x}}$ are both zero. According to [21], every motion of (9) either is asymptotically stable or unstable. But any motion starting at a point $\mathbf{x}_0$, $\dot{\mathbf{x}}_0$ such that $V(\mathbf{x}_0, \dot{\mathbf{x}}_0) < 0$ becomes unbounded since $\dot{V}$ is nonpositive along the motion [2, pp. 19].

Alternatively, a different set of conditions may be obtained by choosing a different function V. For the function

$$V=\mathbf{x}^T(A^TB+B-B^TB)\mathbf{x}+(\mathbf{x}^TB^T+\dot{\mathbf{x}}^T)(B\mathbf{x}+\dot{\mathbf{x}}) \tag{16}$$

the first time derivative is

$$\begin{aligned}\dot{V}=&\dot{\mathbf{x}}^T(A^TB+B-B^TB)\mathbf{x}+\mathbf{x}^T(A^TB+B-B^TB)\dot{\mathbf{x}}\\&(\dot{\mathbf{x}}^TB^T+\ddot{\mathbf{x}}^T)(B\mathbf{x}+\dot{\mathbf{x}})+(\mathbf{x}^TB^T+\dot{\mathbf{x}}^T)(B\dot{\mathbf{x}}+\ddot{\mathbf{x}})\end{aligned} \tag{17}$$

Assuming the matrix

$$A^TB+B-B^TB \tag{18}$$

to be symmetric and substituting equation (12) into the equation (17) gives

$$\begin{aligned}\dot{V}=&\dot{\mathbf{x}}^T(A^TB+B-B^TB)\mathbf{x}+\mathbf{x}^T(B^TA+B^T-B^TB)\dot{\mathbf{x}}\\&+(\dot{\mathbf{x}}^TB^T-\dot{\mathbf{x}}^TA^T-\mathbf{x}^TB^T)(B\mathbf{x}+\dot{\mathbf{x}})\\&+(\mathbf{x}^TB^T+\dot{\mathbf{x}}^T)(B\dot{\mathbf{x}}-A\dot{\mathbf{x}}-B\mathbf{x})\end{aligned}$$

or

$$\dot{V}=-2\mathbf{x}^TB^TB\mathbf{x}-\dot{\mathbf{x}}^T(A+A^T-B-B^T)\dot{\mathbf{x}} \tag{19}$$

Based on the functions V and $\dot{\mathbf{V}}$, the following theorems can be stated for the system in the equation (9):

Theorem 4. The equilibrium is stable if the matrix (18) is symmetric and positive definite and the matrix

$$A+A^T-B-B^T \tag{20}$$

is nonnegative.

Theorem 5. The equilibrium is asymptotically stable if the matrices (18) and (20) are both positive definite.

Theorem 6. The equilibrium is unstable if (18) is not nonnegative and (20) is positive definite.

Validity of these theorems follow directly from the proof stated for Theorems 1-3.

It is worthwhile to mention that, although there are many systems where the matrix (18) is symmetric (one such system will be presented in example 1), it is possible to find systems where (18) is not symmetric. For such systems theorems 4-6 fail to provide any results and, therefore, theorems 1-3 must be used alternatively.

Examples

The purpose of the following examples is to illustrate the presented approach, thus a small number of degrees of freedom are used. However, the power of the method is hoped to be in defining the stability of systems with many degrees of freedom.

Example 1: Consider the system

$$\ddot{\mathbf{x}}+\begin{bmatrix}10 & -3.25\\ 1 & 4\end{bmatrix}\dot{\mathbf{x}}+\begin{bmatrix}4 & -2\\ 1 & 2\end{bmatrix}\mathbf{x}=0$$

the conditions (18) and (20) give

$$A^TB+B^T-B^TB=\begin{bmatrix}26 & -9\\ -9 & 4\end{bmatrix}$$

$$A+A^T-B-B^T=\begin{bmatrix}12 & -1.25\\ -1.25 & 4\end{bmatrix}$$

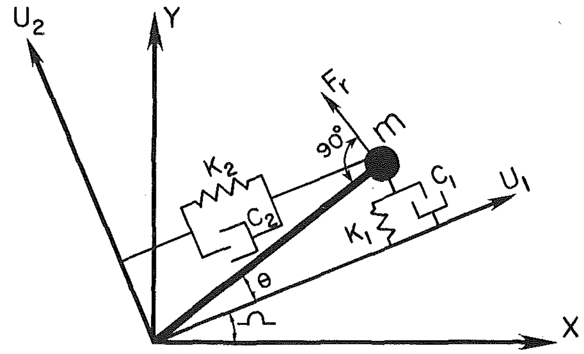


Fig. 1 Discretized model of a damped, flexible, rotating shaft

Since the foregoing matrices are both positive definite, theorem 5 indicates that the equilibrium is asymptotically stable. In addition, substituting A and B into equations (13), (14), and (15) results in

$$Q_1=\begin{bmatrix}78 & -35\\ -35 & 29\end{bmatrix},\quad B+B^T=\begin{bmatrix}8 & -3\\ -3 & 4\end{bmatrix}$$

and

$$Q_3-Q_2Q_1^{-1}Q_2^T=\begin{bmatrix}19.9248 & -2.2162\\ -2.2162 & 7.9720\end{bmatrix}$$

which are all positive definite. Therefore, therorem 2 also indicates that the equilibrium is asymptotically stable.

One way to verify the validity of these results is to form the associated lambda-matrix and obtain the latent roots of the system. Doing so results in

$$\lambda_{1,2}=-1.12385\pm1.2745\,i$$

$$\lambda_{3,4}=-\ .761503\pm2.3985\,i$$

where $i=\sqrt{-1}$, which is in agreement with the prediction of theorems 2 and 5.

Example 2: As another example, consider the discretized model of a damped flexible, rotating shaft presented in [22] and shown in Fig. 1. The mass is attached to a coordinate system which rotates at a constant angular velocity Ω. To make the problem more comprehensive, we assume the existence of a conservative force that is proportional to the radial distance of the mass from the origin and perpendicular to the radius vector. Such a force can be visualized as arising in a rotating fluid or in an electromagnetic field.

The governing equation of motion for the shaft is

$$\begin{bmatrix}m & 0\\ 0 & m\end{bmatrix}\ddot{u}+\begin{bmatrix}c_1 & -2m\Omega\\ 2m\Omega & c_2\end{bmatrix}\dot{u}+\begin{bmatrix}k_1-m\Omega^2 & F\\ -F & k_2-m\Omega^2\end{bmatrix}u=0 \tag{21}$$

where $u=[u_1, u_2]^T$ represents the displacements in the directions of the two rotating coordinate axes. Without loss of generality let $m=1$ and rewrite equation (21) as

$$\ddot{u}+\begin{bmatrix}c_1 & -2\Omega\\ 2\Omega & c_2\end{bmatrix}\dot{u}+\begin{bmatrix}k_1-\Omega^2 & F\\ -F & k_2-\Omega^2\end{bmatrix}u=0 \tag{22}$$

This allows one to use the results presented here to determine the way each parameter affects the stability of the given system. For instance, upon substituting A and B into the matrices (13-15), the conditions required by theorem 2 can be reduced to the following inequalities

$$(k_1-\Omega^2)>0 \tag{23a}$$

$$(k_2-\Omega^2)>0 \tag{23b}$$

$$p_1>0 \tag{23c}$$

$$p_1p_3-p_2^2>0 \tag{23d}$$

$$p_4>0 \tag{23e}$$

$$p_4p_6-p_5^2>0 \tag{23f}$$

where

$$p_1=2(c_1(k_1-\Omega^2)-2\Omega F)$$

$$p_2=(c_1F+2\Omega(k_2-\Omega^2)-2\Omega(k_1-\Omega^2)-c_2F)$$

$$p_3=2(c_2(k_2-\Omega^2)-2\Omega F)$$

$$p_4=c_1-4F^2(c_1(k_1-\Omega^2)-2\Omega F)/(p_1p_3-p_2^2)$$

$$p_5=-2F^2(c_1F+2\Omega(k_2-\Omega^2)-2\Omega(k_1-\Omega^2)-c_2F)$$

$$p_6=c_2-4F^2(c_2(k_2-\Omega^2)-2\Omega F)/(p_1p_3-p_2^2)$$

Now, one can easily investigate how each parameter affects the foregoing conditions and, therefore, the stability. This may be used also to determine the stability boundaries.

It is worthwhile to mention that the results in [18] fail to provide any answer to the preceding problems, since the commuting conditions, shown in (5), are not satisfied. Although the approach in [19] may provide an answer, the symmetrizability of the matrices A and B needs to be determined first. The approach used here eliminates that step.

Observations

The results presented here can be applied directly to the mass, damping, and stiffness matrices. Consequently, for systems with few degrees of freedom, an approach similar to that presented in example 2 can be used to determine the effect of parameter variations on the stability of system or to design a system. When dealing with high-order systems, one can show the effect of parameters on the stability by repeatedly changing the parameters and testing the conditions through an iterative routine. This requires eigensolution of the matrices Q_1, (14), and (15) if theorems 1-3 are used, or matrices (18) and (20) if theorems 4-6 are employed. However, since all these matrices are symmetric, many well-known classical methods, such as Jacobi's method, Given's method, Householder's method, QR method, can be used effectively [23-25].

It is known that the additional damping and stiffness matrices, resulting from velocity and displacement feedback, can stabilize an unstable system or in some cases destabilize an otherwise stable system. The approach presented here can be used to investigate the influence of the feedback gains on the stability of an actively controlled system. To illustrate this more clearly, consider the actively controlled system described by

$$\ddot{\mathbf{x}}+\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}\dot{\mathbf{x}}+\begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}\mathbf{x}=$$

$$\begin{bmatrix} 0 & g_1 \\ 0 & g_2 \end{bmatrix}\dot{\mathbf{x}}+\begin{bmatrix} 0 & h_1 \\ h_2 & 0 \end{bmatrix}\mathbf{x}$$

for this system, the conditions (18) and (20) can be reduced to

$$4(a_1+b_1)(a_4+b_4+g_2)$$

$$-(a_2+a_3-b_2-b_3+g_1-h_1-h_2)^2>0$$

$$a_1(b_2+h_2)+a_3b_4+b_3+h_2=$$

$$(a_2+g_1)b_1+(a_4+g_2)(b_3+h_2)+b_2+h_1$$

$$a_1b_1+a_3(b_3+h_2)+b_1-b_1^2-(b_3+h_2)^2>0$$

$$[a_1b_1+a_3b_3+a_3h_2+b_1-b_1^2-(b_3+h_2)^2]x$$

$$[(a_2+g_1)(b_2+h_1)+(a_4+g_2)b_4+b_4$$

$$-(b_2+h_1)^2-b_4]-[(a_2+g_1)b_1+(a_4+g_2)(b_3+h_2)$$

$$+(b_2+h_1)-b_1(b_2+h_1)-(b_3+h_2)b_4]^2>0,$$

which can be employed to find how the gain constants g_1-g_4 affect the stability of a given system (where a_1-a_4 and b_1-b_4 are known). The importance of feedback control in stability of dynamic systems and different ways of measuring their effects have been discussed in many studies, such as [26] and [27]. However, the significance of the approach presented here is that it extends the results in these studies to asymmetric systems.

Conclusion

For general linear lumped-parameter systems, which can be described by a second-order vector differential equation, different stability conditions were established. These conditions are directly in terms of the coefficient matrices, thus they can be used to determine the effect of system parameters on stability or to design a controller for an actively controlled system. Examples were used to illustrate the validity and the utility of the presented approach. A number of previously developed studies such as the *KTC* theorem and its extensions, the results on systems with constraint damping, and those on symmetrizable systems were mentioned. Although for systems with few degrees of freedom, the present approach is not necessarily advantageous over the Hurwitz stability criterion, it may prove to be more effective for systems with many degrees of freedom.

References

1 Routh, E. J., *A Treatise on the Stability of a Given State of Motion*, Macmillan and Co., London, 1977.

2 Hahn, W., *Theory and Application of Liapunov's Direct Method*, Prentice-Hall, Englewood Cliffs, N.J., 1963.

3 LaSalle, J., and Lefochetz, S., *Stability by Lyapunov's Direct Method*, Academic Press, New York, 1961.

4 Lord Rayleigh, *The Theory of Sound*, Vol. 1 Dover, New York, 1945.

5 Moran, T. J., "A Simple Alternative to the Routh-Hurwitz Criterion for Symmetric Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, Dec., 1970, pp. 1168-1170.

6 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, No. 4, 1973, pp. 1120-1121.

7 Chetaev, N. G., *The Stability of Motion* (translated by M. Nadler), Pergamon Press, p. 196.

8 Zajac, E. E., "The Kelvin-Tait-Chetaev Theorem and Extensions," *Journal of the Astronautical Sciences*, Vol. 11, No. 2, 1964, pp. 46-49.

9 Zajac, E. E., "Comments on Stability of Damped Mechanical Systems and a Further Extension," *AIAA Journal*, Vol. 3, 1965, pp. 1749-1750.

10 Huseyin, K., *Vibrations and Stability of Multiparameter-Systems*, Noordhoff International Publishing, The Netherlands, 1978.

11 Roberson, R. E., "Notes on the Thomason-Tait-Chetaev Stability Theorem," *Journal of the Astronautical Sciences*, Vol. 15, No. 6, 1968, pp. 319-324.

12 Connell, G. M., "Asymptotic Stability of Second-Order Linear Systems with Semidefinite Damping," *AIAA Journal*, Vol. 7, June 1969, pp. 1185-1187.

13 Müller, P. C., "Asymptotische Stabiliät von Linearen Mechanischen Systemen mit Postitiv Semidefiniter Dämpfungs Matrix," *ZAMM*, Vol. 52, 1971, pp. T197-T198.

14 Hughes, P. C., and Garder, L. T., "Asymptotic Stability of Linear Stationary Mechanical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, Mar. 1975, pp. 228-229.

15 Kane, T. R., and Barba, P. M., "Effects of Energy Dissipation on a Spinning Satellite," *AIAA Journal*, Vol. 4, 1966, pp. 1391-1394.

16 Likins, P. W., "Attitude Stability Criteria for Dual-Spin Spacecraft," *Journal of Spacecraft Rockets*, Vol. 4, 1967, pp. 1638-1643.

17 Likins, P. W., "Stability Analysis of Mechanical Systems with Constraint Damping," *AIAA Journal*, Vol. 5, 1967, pp. 2091-2094.

18 Mingori, D. L., "A Stability Theorem for Mechanical Systems with Constraint Damping," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, 1970, pp. 253-258.

19 Inman, D. J., "Dynamics of Asymmetric Non-Conservative Systems," ASME Journal of Applied Mechanics, Vol. 50, 1983, pp. 199–203.

20 Taussky, O., and Zassenhaus, H., "On the Similarity Transformation Between a Matrix and its Transpose," *Pacific Journal of Mathematics*, Vol. 9, 1969, pp. 839–896.

21 LaSalle, J. P., "Stability Theory for Ordinary Differential Equations, *Journal of Differential Equations*, Vol. 4, 1968, pp. 57–65.

22 Shieh, R. C., and Masur, E. F., "Some General Principal of Dynamic Instability of Solid Bodies," *ZAMP*, Vol. 19, 1968, pp. 927–941.

23 Blum, E. K., *Numerical Analysis and Computational Theory and Practice* (1st ed.), Addision-Wesley, 1972.

24 Havanessian, S. A., and Pipes, L. A., *Digital Computer Methods in Engineering* (1st ed.), McGraw Hill, 1969.

25 Shoup, T. E., *A Practical Guide to Computer Methods for Engineers* (1st ed.), Prentice-Hall, 1979.

26 C.S. Draper Lab., "Active Control of Space Structures," Final Report, Repoort No. R-1454, Feb. 1981.

27 Canavin, J. R., "The Control of Spacecraft Vibrations Using Mutlivariable Output Feedback," Paper 78-1419, AIAA/AAS Astrodynamics Conference, Palo Alto, Calif., Aug. 7–9, 1978.

APPENDIX

Lemma 1. The $2n \times 2n$ real symmetric matrix

$$Q = \begin{bmatrix} Q_1 & Q_2{}^T \\ Q_2 & Q_3 \end{bmatrix} \qquad \text{(A1)}$$

is positive definite if and only if both Q_1 and $Q_3 - Q_2 Q_1{}^{-1} Q_2{}^T$ are positive definite.

Proof: The matrix Q is positive definite if and only if

$$V = \mathbf{z}^T Q \mathbf{z}$$

is positive definite for all $\mathbf{z} \neq 0$. The vector z can be presented as

$$\mathbf{z} = Ty,$$

where

$$T = \begin{bmatrix} I & -Q^{-1} Q_2{}^T \\ 0 & I \end{bmatrix}$$

and

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Therefore, V can be written as

$$V = y^T \begin{bmatrix} I & 0 \\ -Q_2 Q_1{}^{-1} & I \end{bmatrix} \begin{bmatrix} Q_1 & Q_2{}^T \\ Q_2 & Q_3 \end{bmatrix} \begin{bmatrix} I & -Q_1{}^{-1} Q_2{}^T \\ 0 & I \end{bmatrix} y$$

or

$$V = y_1{}^T Q_1 y_1 + y_2{}^T (Q_3 - Q_2 Q_1{}^{-1} Q_2{}^T) y_2.$$

The function V, and as a result the matrix Q, is positive definite if and only if Q_1 and $Q_3 - Q_2\, Q_1{}^{-1}\, Q_2{}^T$ are both positive definite.

Lemma 2. The $2n \times 2n$ matrix Q, presented in the equation (A1), is positive semidefinite if and only if Q_1 is positive definite and $Q_3 - Q_2\, Q_1{}^{-1}\, Q_2{}^T$ is positive semidefinite.

Proof: Same as Lemma 1 except the term "positive definite" is replaced by "positive semi-definite" for the matrix $Q_3 - Q_2\, Q_1{}^{-1}\, Q_2{}^T$.

J. M. McCarthy
Assistant Professor.
Department of Mechanical Engineering
and Applied Mechanics,
University of Pennsylvania,
Philadelphia, PA 19104

B. Ravani
Assistant Professor.
Department of Mechanical Engineering,
University of Wisconsin,
Madison, WI 53706

Differential Kinematics of Spherical and Spatial Motions Using Kinematic Mapping

This paper develops the basic framework for studying differential kinematics of spherical and spatial motions using a mapping of spatial kinematics. Relationships are derived relating the intrinsic properties of the image curves corresponding to a mapping of spherical and spatial kinematics and the instantaneous invariants of the corresponding spherical and spatial motions. In addition, in the case of spherical motions, the equations for the inflection cone and cubic cone of stationary geodesic curvature, important in spherical mechanism synthesis, are derived in terms of the curvature and torsion of corresponding image curves. Similar relationships defining the polhodes of spherical motions and their curvature at the reference instant are recast as well. A simple example involving a special spherical four-bar motion is also presented.

1 Introduction

The set of rigid rotations of the three dimensional Euclidean space E^3, and the set of rigid displacements (rotations and translations) of E^3 both form differential manifolds known as Lie groups, see Herman 1966. Continuous curves in these manifolds are known, respectively, as spherical and spatial rigid motions. By choosing coordinate systems for the fixed and moving spaces, we may represent a spatial motion as a parameterized set of rotation matrices $A(t)$ and postion vectors $\mathbf{d}(t)$; $T(t)$: $(A(t), \mathbf{d}(t))$. Ravani and Roth 1984 showed that by using Study's parameters (Study 1891), a geometric mapping can be defined that carries a spatial motion $T(t)$ into a special curve in a projective space of three "dual" dimensions. They went further and showed that the real part of the same mapping carries $A(t)$ into a curve in a three dimensional real projective space; see also Bottema and Roth 1979. Using such a mapping, a motion can be studied by studying its image curve in the image space of the mapping. Ravani and Roth applied the mapping[1] to the study of spatial motions and mechanisms and in this way generalized a technique which has been highly developed in the study of planar rigid motion; see Blaschke and Muller 1956, Bottema and Roth 1979, DeSa and Roth 1981 *a*,*b*, and Ravani and Roth 1983.

In this paper, we develop the basic framework for applying this geometric mapping to the study of local properties of spherical and spatial motions. The study of local properties of rigid body motion is a highly developed branch of kinematics known as Instantaneous Kinematics. Important tools in this area are the canonical coordinate systems and unit angular velocity parameterization; see Veldkamp 1967, Kirson 1975, Roth and Yang 1977, Bottema and Roth 1979. The use of the canonical coordinate system and the unit angular velocity parameterization results in a set of unique constants that characterize the motion–the so-called Instantaneous Invariants. It is clear that the shape of the image curve of a mapping of a rigid motion also provides a set of constants which characterize a motion. In fact Blaschke and Muller 1956 have used this idea to study the differential kinematics of planar motions. In this paper, we generalize this technique to the study of differential kinematics of spherical and spatial motions. We will use the mapping introduced by Ravani and Roth 1984 and relate the differential properties of the image curves of spherical and spatial motions to the instantaneous invariants characterizing such motions. In the case of spherical motions, we go further by relating the local or differential properties of point paths of such motions to the local properties of their corresponding image curves. In addition, the inflection cone and cubic cone of stationary geodesic curvature important in spherical mechanism synthesis are derived in terms of the curvature and torsion of the corresponding image curve. The relationships defining the polhodes of spherical motions and their curvature at the reference instant in terms of the local properties of the corresponding image curves are also derived. Finally, we examine the motion of a special spherical four-bar mechanism for which an image curve is known and obtain its instantaneous properties throughout the motion.

---

[1] It should be pointed out that many other mappings can also be defined. In this paper, we are only referring to the mapping defined in Ravani and Roth (1984).

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami, Fla., November 17–21, 1985 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper shoud be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publicaton of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 27, 1984; final revision, May 21, 1985. Paper No. 85-WA/APM-26.

2 Euler Parameters of Orthogonal Matrices and the Mapping

A 3×3 orthogonal matrix characterizing a spherical displacement can be written in terms of it Euler parameters:

$$A = \Omega^{-2} \begin{vmatrix} (c_0^2 + c_1^2 - c_2^2 - c_3^2) & 2(c_1 c_2 - c_0 c_3) & 2(c_1 c_3 + c_0 c_2) \\ 2(c_2 c_1 + c_0 c_3) & (c_0^2 - c_1^2 + c_2^2 - c_3^2) & 2(c_2 c_3 - c_0 c_1) \\ 2(c_3 c_1 - c_0 c_2) & 2(c_3 c_2 + c_0 c_1) & (c_0^2 - c_1^2 - c_2^2 + c_3^2) \end{vmatrix} \tag{1}$$

where

$$\Omega^2 = c_0^2 + c_1^2 + c_2^2 + c_3^2 \tag{2}$$

and c_i ($i = 0,1,2,3$) are the Euler parameters of the orthogonal matrix A. Since the matrix A depends on three independent parameters, namely the ratios c_i/c_0 ($i = 1,2,3$), we can map it (Ravani and Roth 1984) into points of a three-dimensional projective space Σ'. A spherical motion characterized by a parameterized set of orthogonal matrices $A(t)$ is then mapped into a curve in the image space Σ'. In view of the fact that the Euler parameters, c_i, can always be normalized such that

$$c_0^2 + c_1^2 + c_2^2 + c_3^2 = 1, \tag{3}$$

the image of $A(t)$ can be considered as a curve, K, on a hypersphere of unit radius in a four dimensional affine space.

The differential geometry of curves on the surface of a hypersphere has been studied by McCarthy 1983; see also Flanders 1963 and O'Neill 1966. Such results can therefore be used to relate the local properties of spherical motions, that is the distribution of velocities, accelerations, change of accelerations, etc., of each point in the body, to geometric parameters which characterize the shape of the image curve, K, of the motion.

A spatial displacement is characterized by a 3×3 orthogonal matrix A and the position vector $\mathbf{d}$. These two quantities can be combined to construct an orthogonal 3×3 dual matrix $\hat{A}$ characterizing a spatial displacement. The dual matrix $\hat{A}$ is driven by

$$\hat{A} = A + \epsilon D A \tag{4}$$

where

$$D = \begin{vmatrix} 0 & -d_z & d_y \\ d_z & 0 & -d_x \\ -d_y & d_x & 0 \end{vmatrix}$$

and ϵ is the dual unit defined such that $\epsilon^2 = 0$. The dual matrix $\hat{A}$ can be written in a form identical to equation (1) using dual Euler parameters $c_i = c_i + \epsilon c_i^*$ ($i = 0,1,2,3$) where the dual parts c_i^* of the parameters are defined by the relations

$$\begin{aligned} c_0^* &= \frac{1}{2}(d_x c_1 + d_y c_2 + d_z c_3) \\ c_1^* &= \frac{1}{2}(-d_y c_0 + d_z c_2 - d_y c_3) \\ c_2^* &= \frac{1}{2}(-d_y c_0 - d_z c_1 + d_x c_3) \\ c_3^* &= \frac{1}{2}(-d_z c_0 + d_y c_1 - d_x c_2) \end{aligned} \tag{5}$$

and

$$\hat{\Omega}^2 = \hat{c}_0^2 + \hat{c}_1^2 + \hat{c}_2^2 + \hat{c}_3^2. \tag{6}$$

Since the 3×3 dual matrix $\hat{A}$ characterizing a spatial displacement depends on three independent dual parameters, namely the ratios $\hat{c}_i/\hat{c}_0$ ($i = 1,2,3$), we can map it into a point of a dual three-dimensional projective space $\hat{\Sigma}'$. A spatial motion characterized by a parameterized set of orthogonal matrices $\hat{A}(t)$ (with t being a real parameter) is then mapped into a special curve in the image space $\hat{\Sigma}'$. For more details on this mapping see Ravani and Roth 1984.

In the view of definition (5) of the dual parts of the Euler parameters, we can write

$$c_0 c_0^* + c_1 c_1^* + c_2 c_2^* + c_3 c_3^* = 0 \tag{7}$$

If normalized (real) Euler parameters are used, then this last equation implies that

$$\hat{\Omega}^2 = \hat{c}_0^2 + \hat{c}_1^2 + \hat{c}_2^2 + \hat{c}_3^2 = 1 \tag{8}$$

This means that the image of $\hat{A}(t)$ can be considered as a curve $\hat{K}$ on a dual hypersphere in an affine space of four dual dimensions. Ravani and Roth 1984 have shown that $\hat{\Omega}^2$ remains invariant under the group of transformations arising in the $\hat{\Sigma}'$ space from different choices of the moving and fixed frames in the real space. This, in the view of elliptic geometry of $\hat{\Sigma}'$ space, means that all choices of moving and fixed frames in the real space, result in mappings which differ by only $\hat{\Sigma}'$ displacements. The final conclusion is that the intrinsic properties of the image curve $\hat{K}$ of a motion are independent of the coordinate system and the parameterization used to define the motion.

Spherical Motion. A continuous rigid motion of Euclidean three-space which maintains a fixed point has the property that the trajectory of every point in the moving body lies on concentric spheres about the fixed point. This motion termed spherical motion is represented by a parameterized set of 3×3 orthogonal matrices $A(t)$. If $\mathbf{p}$: (x,y,z) is a point in the moving reference frame M then its trajectory $\alpha_p(t)$: $(X(t), Y(t), Z(t))$ is the set of points in the fixed frame F with which $\mathbf{p}$ coincides as $F(t)$ varies with t, given by the motion equation

$$\alpha_p(t) = A(t)\mathbf{p}$$

To study the instantaneous properties of the motion in the vicinity of a reference position, which we denote with the parameter value $t = 0$, we expand $A(t)$ in a Taylor series

$$A(t) = A_0 + A_1 t + A_2 \frac{t^2}{2} + \ldots, \tag{9a}$$

The subscript denotes the n^{th} derivatives evaluated at the reference position. The coefficient matrices A_i, $i = 0,1,2,\ldots$ are not unique to the motion since the reference frames M and F are chosen arbitrarily. Bottema and Roth show how a canonical set of reference frames may be chosen such that (9*a*) is unique up to the choice of the motion parameterization.

The motion parameter t defines the speed at which the body moves through a sequence of positions; it does not affect the shape of trajectories traced by the moving body. By choosing the standard unit angular velocity parameterization, we obtain a unique representation of the motion

$$A(\phi) = I + A_1 \phi + A_2 \frac{\phi^2}{2} + A_3 \frac{\phi^3}{6} + \ldots \tag{9b}$$

where ϕ is the new motion parameter, $\phi = 0$ denotes the reference position. I is the 3×3 identity matrix and A_i, $i = 1,2,3$ are given by

$$A_1=\left|\begin{array}{ccc}0&-1&0\\1&0&0\\0&0&0\end{array}\right|,\quad A_2=\left|\begin{array}{ccc}-1&0&e\\0&-1&0\\-e&0&0\end{array}\right|,$$

$$A_3=\left|\begin{array}{ccc}0&(1+e^2)&\gamma_y\\-(1+e^2)&0&\left(\frac{3}{2}e-\gamma_x\right)\\-\gamma_y&\left(\frac{3}{2}e+\gamma_x\right)&0\end{array}\right|.$$

The constants e, γ_x, γ_y define the spherical motion to third order. Each additional order of the expansion (9) adds three more constants. These constants are called the instantaneous invariants of the spherical motion.

The matrix function $A(t)$ is also prescribed by the four Euler functions $c_i(t)$, $i = 0,1,2,3$. The Taylor series expansion of $A(t)$ about $t = 0$ may be computed in terms of the series expansions of $c_i(t)$:

$$c_i(t)=c_{i0}+c_{i1}t+c_{i2}\frac{t^2}{2}+c_{i3}\frac{t^3}{6}+\ldots \tag{11}$$

We will relate the instantaneous invariants ϵ, γ_x and γ_y to constants c_{in}, $i = 0,1,2,3$, $n = 0,1,2,\ldots$ by determining the values of the c_{in} in the canonical coordinate system. In the computations which follow it is useful to note that

$$\begin{aligned}c_i(t)c_j(t)&=c_{i0}c_{j0}+(c_{i0}c_{j1}+c_{i1}c_{j0})t\\&+(c_{i0}c_{j2}+c_{i2}c_{j0}+2c_{i1}c_{j1})\frac{t^2}{2}\\&+(c_{i0}c_{j3}+c_{i3}c_{j0}+3c_{i1}c_{j2}+3c_{i2}c_{j1})\frac{t^3}{6}\end{aligned} \tag{12}$$

By choosing the reference frames F and M of the fixed and moving bodies so they coincide at the instant $t = 0$ we have

$$A_0=\left|\begin{array}{ccc}1&0&0\\0&1&0\\0&0&1\end{array}\right|$$

$$=\left|\begin{array}{ccc}c_{00}^2+c_{10}^2-c_{20}^2-c_{30}^2&2(c_{10}c_{20}-c_{00}c_{30})&2(c_{10}c_{30}+c_{00}c_{20})\\2(c_{20}c_{10}+c_{00}c_{30})&c_{00}^2-c_{10}^2+c_{20}^2-c_{30}^2&2(c_{20}c_{30}-c_{00}c_{10})\\2(c_{30}c_{10}-c_{00}c_{20})&2(c_{30}c_{20}+c_{00}c_{10})&c_{00}^2-c_{10}^2-c_{20}^2+c_{30}^2\end{array}\right| \tag{13}$$

from which we conclude that $c_{00} = 1$ and $c_{i0} = 0$, $i = 1,2,3$. The angular velocity matrix $\Omega(t) = \dot{A}A^T(t)$ of the motion at the instant $t = 0$ is given by

$$\Omega_0=A_1A_0^T=A_1$$

$$=\left|\begin{array}{ccc}2c_{01}&-2c_{31}&2c_{21}\\2c_{31}&2c_{01}&-2c_{11}\\-2c_{21}&2c_{11}&2c_{01}\end{array}\right|. \tag{14}$$

The angular velocity matrix is skew symmetric ($\Omega^T = -\Omega$); therefore $c_{01} = 0$. The vector $\mathbf{w} = (2c_{11}, 2c_{21}, 2c_{31})$ is the angular velocity vector of the motion at this instant and we orient the reference frames F and M such that the z-axes of both frames is directed along $\mathbf{w}$. This forces both c_{11} and c_{21} to be zero and we have

$$\Omega_0=A_1=\left|\begin{array}{ccc}0&-2c_{31}&0\\2c_{31}&0&0\\0&0&0\end{array}\right|. \tag{15}$$

The second order term A_2 is now computed to be

$$A_2=\left|\begin{array}{ccc}2c_{02}-2c_{31}^2&-2c_{32}&2c_{22}\\2c_{32}&2c_{02}-2c_{31}^2&-2c_{12}\\-2c_{22}&2c_{12}&2c_{02}+2c_{31}^2\end{array}\right| \tag{16}$$

This is simplifed by means of the restriction which (6) places on the functions $c_i(t)$ $i = 0,1,2,3$:

$$\begin{aligned}1=\sum_{i=0}^{3}\Big[&c_{i0}^2+(2c_{i0}c_{i1})t+(2c_{i0}c_{i2}+2c_{i1}^2)\frac{t^2}{2}\\&+(6c_{i1}c_{i2}+2c_{i0}c_{i3})\frac{t^3}{6}+\ldots\Big]\end{aligned} \tag{17}$$

Since

$$\sum_{i=0}^{3}c_{i0}^2=c_{00}^2+c_{10}^2+c_{20}^2+c_{30}^2=c_{00}^2=1 \tag{18}$$

the coefficient of every other term in (17) must equal zero. The first order term simply restates the requirement that $c_{01} = 0$. The second order term introduces the identity

$$2c_{02}+2c_{31}^2=0. \tag{19}$$

We further simplify (16) by rotating the reference frames F and M about $\mathbf{w}$ so that $c_{12} = 0$ and A_2 becomes

$$A_2=\left|\begin{array}{ccc}-4c_{31}^2&-2c_{32}&2c_{22}\\2c_{32}&-4c_{31}^2&0\\-2c_{22}&0&0\end{array}\right| \tag{20}$$

The angular acceleration matrix E at $t = 0$ is given by

$$E(0)=\frac{d}{dt}[\dot{A}A^T]=\ddot{A}A^T+\dot{A}\dot{A}^T=A_2+A_1A_1^T \tag{21}$$

which yields

$$E=\left|\begin{array}{ccc}0&-2c_{32}&2c_{22}\\2c_{32}&0&0\\-2c_{22}&0&0\end{array}\right| \tag{22}$$

In vector form we obtain $\dot{\mathbf{w}} = \mathbf{e} = (0, 2c_{22}, 2c_{32})$ thus the choice of the canonical coordinate frame is such that the angular acceleration vector lies in the y-z coordinate plane. We specify the sense of the y-axis to be such that the y-component of $\mathbf{e}$ is positive. The canonical frame is now completely prescribed.

The third order term A_3 is given by

$$A_3=\left|\begin{array}{ccc}(2c_{03}-6c_{31}c_{32})&-2(c_{33}+3c_{02}c_{31})&2c_{23}\\2(c_{33}+3c_{02}c_{31})&(2c_{03}-6c_{31}c_{32})&2(3c_{33}c_{31}-c_{13})\\-2c_{23}&2(3c_{22}c_{31}+c_{13})&(2c_{03}+6c_{31}c_{32})\end{array}\right| \tag{23}$$

The third order term of the normalization condition (17) yields

$$6c_{31}c_{32}+2c_{03}=0 \tag{24}$$

Using this relation as well as (19) we obtain

$$A_3 = \begin{vmatrix} -12c_{31}c_{32} & -2c_{33}+6c_{31}^3 & 2c_{23} \\ 2c_{33}-6c_{31}^3 & -12c_{31}c_{32} & 6c_{22}c_{31}-2c_{13} \\ -2c_{23} & 6c_{22}c_{31}+2c_{13} & 0 \end{vmatrix} \tag{25}$$

So far we have the constants c_{31}, c_{22}, c_{32}, c_{13}, c_{23} and c_{33} which define $A(t)$ to the third order. Each additional order of the expansion adds four constants, one of which may be eliminated using the normalization relation (17), thus only three of the four are independent.

The unit angular velocity parameterization is defined such that the magnitude of the angular velocity is unity throughout the motion. This is equivalent to requiring that

$$-\frac{1}{2}\,\mathrm{trace}\,(\Omega^2) = 1 \tag{26}$$

where Ω is the angular velocity matrix and the trace operator computes the sum of the diagonal terms. A series expansion of $\Omega(t)$ yields

$$\Omega(t) = \Omega_0 + Et + \left[\Gamma - \Omega_0^3 + \frac{1}{3}(\Omega_0 E - E\Omega_0)\right] t^2/2 + \ldots \tag{27}$$

where Ω_0 and E are given by (15) and (22) and Γ is

$$\Gamma = \begin{vmatrix} 0 & -2c_{33}+6c_{31}^3 & 2c_{23} \\ 2c_{33}-6c_{31}^3 & 0 & -2c_{13} \\ -2c_{23} & 2c_{13} & 0 \end{vmatrix} \tag{28}$$

The requirement of a unit angular velocity adds further constraints on the values of the constants c_{in}, one for each order of the expansion of Ω^2:

$$-\frac{1}{2}\,\mathrm{trace}\,(\Omega_0^2) = 4c_{31}^2 = 1 \tag{29}$$

$$-\frac{1}{2}\,\mathrm{trace}\,(\Omega_0 E + E\Omega_0) = 4c_{31}c_{32} = 0 \tag{30}$$

$$-\frac{1}{2}\,\mathrm{trace}\left((E^2 - \Omega_0^4) + \frac{1}{2}(\Omega_0\Gamma + \Gamma\Omega_0)\right) =$$
$$-2c_{31}\bullet(-2c_{33}+6c_{31}^3) + (2c_{22})^2 + (2c_{32})^2 + (2c_{31})^4 = 0 \tag{31}$$

These relations yield the results that $c_{31} = 1/2$, $c_{32} = 0$ and

$$c_{33} = -\left(2c_{22}^2 + \frac{1}{8}\right) \tag{32}$$

The coefficients matrices of $A(\phi)$ in the canonical coordinate system are thus obtained in terms of the coefficients of the Euler functions as:

$$A_0 = [I],\ A_1 \begin{vmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}$$

$$A_2 = \begin{vmatrix} -1 & 0 & 2c_{22} \\ 0 & -1 & 0 \\ -2c_{22} & 0 & 0 \end{vmatrix}$$

$$A_3 = \begin{vmatrix} 0 & (4c_{22}^2+1) & 2c_{23} \\ -(4c_{22}^2+1) & 0 & 3c_{22}-2c_{13} \\ -2c_{23} & 3c_{22}+2c_{13} & 0 \end{vmatrix}. \tag{33}$$

Comparing (33) to (10), we see immediately that

$$c_{22} = \frac{e}{2},\ c_{13} = \frac{\gamma_x}{2} \text{ and } c_{23} = \frac{\gamma_y}{2}.$$

In summary,

$$\begin{array}{llll} c_{00} = 1 & c_{10} = 0 & c_{20} = 0 & c_{30} = 0 \\ c_{01} = 0 & c_{11} = 0 & c_{21} = 0 & c_{31} = 1/2 \\ c_{02} = -1/4 & c_{12} = 0 & c_{22} = e/2 & c_{32} = 0 \\ c_{03} = 0 & c_{13} = \gamma_x/2 & c_{23} = \gamma_y/2 & c_{33} = -(e^2/2 + 1/8). \end{array} \tag{34}$$

Image Curves of Spherical Motions—Differential Properties. Using the Euler functions $c_i(t)$, $i = 0,1,2,3$ we obtain a curve $\mathbf{c}(t)$ on the unit hypersphere which is the image of the matrix function $A(t)$. In canonical coordinates $\mathbf{c}(t)$ takes the form to the third order

$$\mathbf{c}(\phi) = \left\{\begin{matrix} 0 \\ 0 \\ 0 \\ 1 \end{matrix}\right\} + \left\{\begin{matrix} 0 \\ 0 \\ 1/2 \\ 0 \end{matrix}\right\}\phi + \left\{\begin{matrix} 0 \\ e/2 \\ 0 \\ -1/4 \end{matrix}\right\}\phi^2/2$$
$$+ \left\{\begin{matrix} \gamma_x/2 \\ \gamma_y/2 \\ -\left(\dfrac{e^2}{2} + \dfrac{1}{8}\right) \\ 0 \end{matrix}\right\}\phi^3/6 \tag{35}$$

where ϕ represents the unit angular velocity parameter. The differential geometry of hyperspherical curves such as $\mathbf{c}(\phi)$ is developed in detail in McCarthy 1983. Two functions characterize the shape of these curves analogous to the curvature and torsion of curves in three dimensional space.

The geometric properties of curves on the hypersphere are studied by means of the Frenet reference frame (T,N,B,E) and the arc-length parameterization $\phi(s)$. The directions E and T of the Frenet frame are chosen along $\mathbf{c}(\phi)$ and $\dot{\mathbf{c}}(\phi)$. These directions are mutually perpendicular since $\mathbf{c}\bullet\mathbf{c} = 1$ implies $2\mathbf{c}\bullet\dot{\mathbf{c}} = 0$. The direction N is chosen along the component of dT/ds which does not lie in the E-T plane; in the same way B is chosen along the component of dN/ds not contained in the E-T-N subspace. The important result of these definitions is the Frenet equations for hyperspherical curves.

$$\frac{dE}{ds} = T$$
$$\frac{dT}{ds} = -E + \kappa N$$
$$\frac{dN}{ds} = -\kappa T + \tau B$$
$$\frac{dB}{ds} = -\tau N \tag{36}$$

The functions κ and τ measure how $\mathbf{c}$ bends out of the E-T plane and how it bends out of the E-T-N subspace. We will refer to these functions as the curvature and torsion of the curve $\mathbf{c}(\phi)$. The Frenet equations prove that κ and τ completely define the shape of hyperspherical curves.

Formulas for κ and τ in terms of $\mathbf{c}(\phi)$ and its derivatives with respect to ϕ are given in McCarthy 1983 as

$$\kappa^2 = \frac{{}^*[\dot{\mathbf{c}}\Lambda\ddot{\mathbf{c}}\Lambda\mathbf{c})\Lambda{}^*(\dot{\mathbf{c}}\Lambda\ddot{\mathbf{c}}\Lambda\mathbf{c})]}{(\dot{\mathbf{c}}\bullet\dot{\mathbf{c}})^3} \tag{37}$$

and

$$\tau = \frac{{}^*(\dot{\mathbf{c}}\Lambda\ddot{\mathbf{c}}\Lambda\mathbf{c}\Lambda\mathbf{c})}{(\dot{\mathbf{c}}\bullet\dot{\mathbf{c}})^3\kappa^2} \tag{38}$$

Equations (37) and (38) use the algebra of multi-vectors, Flanders 1963, in order to generalize the vector cross product to four dimensional space.

The dual of the rank three multi-vector $\dot{\mathbf{c}}\Lambda\ddot{\mathbf{c}}\Lambda\mathbf{c}$ is computed to be the four dimensional vector

$$^*(\dot{\mathbf{c}}\Lambda\ddot{\mathbf{c}}\Lambda\mathbf{c}) = \begin{Bmatrix} -e/4 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{39}$$

The combination of operations in the numerator of (37) is in essence the dot product of (39) with itself; thus we have

$$\kappa^2 = \frac{(e/4)^2}{(1/4)^3} = 4e^2 \tag{40}$$

The canonical coordinate frame is defined in such a way such that $e \geq 0$; thus we have $\kappa = 2e$. This is the curvature of $\mathbf{c}(\phi)$ at the reference instant $\phi = 0$.

The numerator of (38) is the determinant of the 4×4 matrix formed from the vectors $\mathbf{c}$, $\dot{\mathbf{c}}$, $\ddot{\mathbf{c}}$ and $\dddot{\mathbf{c}}$, thus τ is given by

$$\tau = -\frac{1}{2}\left(\frac{e}{2}\right)\left(\frac{\gamma_x}{2}\right)\Big/\left[\left(\frac{1}{4}\right)^3 (4e^2)\right] = -\frac{2\gamma_x}{e} \tag{41}$$

An expression for κ' is also given in McCarthy 1983 which in our case becomes

$$\frac{d\kappa}{ds} = \frac{(\dot{\mathbf{c}}\cdot\dot{\mathbf{c}})\{{}^*[(\dot{\mathbf{c}}\Lambda\ddot{\mathbf{c}}\Lambda\mathbf{c})\Lambda{}^*(\dot{\mathbf{c}}\Lambda\dddot{\mathbf{c}}\Lambda\mathbf{c})]\} - 3(\dot{\mathbf{c}}\cdot\ddot{\mathbf{c}})[{}^*[(\dot{\mathbf{c}}\Lambda\ddot{\mathbf{c}}\Lambda\mathbf{c})\Lambda{}^*(\dot{\mathbf{c}}\Lambda\ddot{\mathbf{c}}\Lambda\mathbf{c})]\}}{(\dot{\mathbf{c}}\cdot\dot{\mathbf{c}})^{1/2}(\dot{\mathbf{c}}\cdot\dot{\mathbf{c}})^4\kappa} \tag{42}$$

the dual of the multi-vector $\dot{\mathbf{c}}\Lambda\dddot{\mathbf{c}}\Lambda\mathbf{c}$ is the vector

$$^*(\dot{\mathbf{c}}\Lambda\dddot{\mathbf{c}}\Lambda\mathbf{c}) = \begin{Bmatrix} -\gamma_y/4 \\ \gamma_x/4 \\ 0 \\ 0 \end{Bmatrix} \tag{43}$$

thus (42) simplifies to be

$$\kappa' = \frac{\left(\frac{1}{4}\right)\left\{\left(\frac{e}{4}\right)\left(\frac{\gamma_y}{4}\right)\right\}}{\left(\frac{1}{4}\right)^4 2e\left(\frac{1}{4}\right)^{1/2}} = 4\gamma_y. \tag{44}$$

Equations (40), (44), and (41) give the curvature κ, rate of change of curvature κ' and torsion τ of the image curve $\mathbf{c}(\phi)$ of a spherical motion in terms of its instantaneous invariants e, γ_x and γ_y.

Spatial Motion. The spatial motion of a general rigid body in Euclidean three space is represented by a parameterized set of orthogonal matrices $A(t)$ together with a vector function $\mathbf{d}(t)$. Choosing a reference frame M in the moving body and another F in the fixed space defines the pair $T(t):(A(t),\mathbf{d}(t))$ such that is $\mathbf{p}:(x,y,z)$ is a point in M its trajectory $\alpha_p(t) = (X(t),Y(t),Z(t))$ in F is given by

$$\alpha_p(t) = A(t)\mathbf{p} + \mathbf{d}(t). \tag{45}$$

$A(t)$ defines the orientation of M relative to F and $\mathbf{d}(t)$ the position of its origin relative to the origin of F. A canonical pair of reference frames M and F and a special motion parameter ϕ may be chosen so that the Taylor series expansions of $A(\phi)$ and $\mathbf{d}(\phi)$ contain a unique set of constants known as the instantaneous invariants of the spatial motion. The orientation of these two frames and the special motion parameter are defined in exactly the same way as was done previously for spherical motion. Therefore, the series expansion of $A(\phi)$ for spatial motion is also given by (9) and (10). We now define the origin of F such that the series expansion

$$\mathbf{d}(\phi) = \mathbf{d}_0 + \mathbf{d}_1\phi + \mathbf{d}_2\frac{\phi^2}{2} + \mathbf{d}_3\frac{\phi^3}{6} + \ldots \tag{46}$$

takes a particulary simple form. First, we require that the origins of M and F coincide at the reference instant $t = \phi = 0$, thus $\mathbf{d}_0 = 0$. Now, as is described in Bottema and Roth 1979, the origin of F is restricted to lie on the line coinciding with the instantaneous screw axis (ISA) of the motion of M relative to F. The ISA is the unique line in the direction of the angular velocity vector $\mathbf{w}$ with the property that points in M lying on it have velocities in the direction parallel to $\mathbf{w}$. This defines $\mathbf{d}_1$ to be $\mathbf{d}_1 = (0,0,d_{z1})$ since the direction of $\mathbf{w}$ has been chosen as the Z axis of F.

Finally, on the ISA there is a unique point with the property that its accleration vector lies in the $Y-Z$ plane of the reference frame F. This point is the striction point at the reference position of the ruled surface generated by the instantaneous screw axis. Choosing this point as the origin of F we have $\mathbf{d}_2 = (0,d_{y2},d_{z2})$. Each additional order of expansion of $\mathbf{d}(\phi)$ adds more constants, for example $\mathbf{d}_3 = (d_{x3},d_{y3},d_{z3})$. Thus in the canonical coordinate system of spatial motion we have to the third order

$$\mathbf{d}_0 = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix},\ \mathbf{d}_1 = \begin{Bmatrix} 0 \\ 0 \\ d_{z1} \end{Bmatrix},\ \mathbf{d}_2 = \begin{Bmatrix} 0 \\ d_{y2} \\ d_{z2} \end{Bmatrix}$$

$$\text{and } \mathbf{d}_3 = \begin{Bmatrix} d_{x3} \\ d_{y3} \\ d_{z3} \end{Bmatrix} \tag{47}$$

The instantaneous invariants of spatial motion to the third order are the constants e, γ_x, γ_y and d_{x1}, d_{y2}, d_{z2}, d_{x3}, d_{y3}, d_{z3}; each additional order adds six more invariants, three from $A(\phi)$ and three from $\mathbf{d}(\phi)$.

The spatial motion $T(t)$: $(A(\phi), \mathbf{d}(\phi))$ is also prescribed by the four dual Euler functions $\hat{c}_i(\phi)$, $i = 0,1,2,3$. The series expansion of these functions about $t = \phi = 0$ is given by

$$\begin{aligned} \hat{c}_i(\phi) &= c_i(\phi) + \epsilon c_i^*(\phi) \\ &= \left(c_{i0} + c_{i1}\phi + c_{i2}\frac{\phi^2}{2} + c_{i3}\frac{\phi^3}{6} + \ldots\right) \\ &\quad + \epsilon\left(c_{i0}^* + c_{i1}^* + c_{i2}^*\frac{\phi^2}{2} + c_{i3}^*\frac{\phi^3}{6} + \ldots\right), \\ &\qquad i = 0,1,2,3 \end{aligned} \tag{48}$$

The real part of (48) depends only on the matrix $A(\phi)$, and the relationship between the invariants e, γ_x and γ_y of $A(\phi)$ and the constants c_{in}, $i = 0,1,2,3$, $n = 0,1,2,3$ has already been determined. The result is that the real part $\mathbf{c}(\phi)$ of the dual curve $\hat{\mathbf{c}}(\phi)$ representing the spatial motion is exactly equation (35) and what remains is the evaluation of the dual part $\mathbf{c}^*(\phi)$. This we compute from the defining equations (9). Since the series expansions of $\mathbf{c}(\phi)$ and $\mathbf{d}(\phi)$ in the canonical coordinate system are both known, equations (34) and (47), respectively, we obtain $\mathbf{c}^*(\phi)$ by direct computation to be to the third order

$$\mathbf{c}^*(\phi) = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ -\frac{d_{z1}}{2} \\ 0 \end{Bmatrix}\phi + \begin{Bmatrix} 0 \\ \frac{-d_{y2}}{2} \\ -\frac{d_{z2}}{2} \\ \frac{d_{z1}}{2} \end{Bmatrix}\phi^2/2 + \begin{Bmatrix} \frac{3d_{z1}e}{4} - \frac{3d_{y2}}{4} - \frac{d_{x3}}{2} \\ -\frac{d_{y3}}{2} \\ \frac{3d_{z1}}{8} - \frac{d_{z3}}{2} \\ \frac{3d_{z2}}{4} \end{Bmatrix}\phi^3/6 \tag{49}$$

The constants $d_{z1}, d_{y2}, d_{z2}, d_{x3}, d_{y3}, d_{z3}$ in (49), together with e, γ_x and γ_y are the instantaneous invariants of the spatial motion.

Image Curves of Spatial Motions – Differential Properties. The dual Euler functions $\hat{\mathbf{c}}_i(\phi) = c_i(\phi) + \epsilon c_i^*(\phi)$, $i = 0,1,2,3$ define a dual curve $\hat{\mathbf{c}}(\phi)$ which is the image of the spatial motion $T(t):(A(\phi), \mathbf{d}(\phi))$. In canonical coordinates $\hat{\mathbf{c}}(t)$ is to the third order

$$\hat{\mathbf{c}}(\phi) = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ \frac{1}{2} - \epsilon\frac{d_{z1}}{2} \\ 0 \end{Bmatrix}\phi + \begin{Bmatrix} 0 \\ \frac{e}{2} - \epsilon\frac{d_{y2}}{2} \\ -\epsilon\frac{d_{z2}}{2} \\ -\frac{1}{4} + \epsilon\frac{d_{z1}}{2} \end{Bmatrix}\frac{\phi^2}{2} + \begin{Bmatrix} \frac{\gamma_x}{2} + \epsilon\left(\frac{3}{4}ed_{z1} = \frac{3}{4}d_{y2} - \frac{d_{x3}}{2}\right) \\ \frac{\gamma_y}{2} - \epsilon\frac{d_{y3}}{2} \\ -\left(\frac{e^2}{2} + \frac{1}{8}\right) + \epsilon\left(\frac{3}{8}d_{z1} - \frac{d_{z3}}{2}\right) \\ \epsilon\frac{3}{4}d_{z2} \end{Bmatrix}\frac{\phi^3}{6} \tag{50}$$

where ϕ is the unit angular velocity parameter.

The curve $\hat{\mathbf{c}}(\phi)$ is a four dimensional vector function over the set of dual numbers. Vector functions similar to $\hat{\mathbf{c}}(\phi)$ but having only three dual number components have been used extensively to study ruled surfaces (see Dimentberg 1965 and Veldkamp 1976). In those works a complete dual vector calculus is developed and it is seen to be formally identical to the usual vector calculus over the set of real numbers. This may be attributed to the fact that algebraically dual numbers have all the properties of real numbers with the exception that division by pure dual numbers (those of the form $\hat{c} = 0 + \epsilon b$) is undefined. If we avoid this singular situation all computations using real numbers may be made using dual numbers as well. With this justification we move ahead with the "dualization" of equations (36), (37) and (38) to obtain results concerning the differential geometry of $\hat{\mathbf{c}}(\phi)$.

In order to study $\hat{\mathbf{c}}(\phi)$ we require the dual Frenet reference frame $(\hat{T}, \hat{N}, \hat{B}, \hat{E})$ and the dual arc-length parameterization. This will allow us to obtain the dual Frenet equations. We define the arc-length $\hat{s}(\phi)$ of $\hat{\mathbf{c}}(\phi)$ by the equation

$$\hat{s}(\phi) = \int_0^\phi (\dot{\hat{\mathbf{c}}} \cdot \dot{\hat{\mathbf{c}}})^{1/2} d\phi \tag{51}$$

the dot denotes differentiation with respect to ϕ. As long as the integrand of (51) is not a pure dual number, this function can be inverted to yield $\phi(\hat{s})$. In practice we compute derivatives with respect to $\hat{s}$ by noting $\hat{s}(\phi(\hat{s})) = \hat{s}$ therefore

$$\frac{d\hat{s}}{d\phi} \cdot \frac{d\phi}{d\hat{s}} = 1 \tag{52}$$

from which we obtain

$$\frac{d\phi}{d\hat{s}} = \frac{1}{\hat{v}} \tag{53}$$

where $\hat{v} = (\dot{\hat{\mathbf{c}}} \cdot \dot{\hat{\mathbf{c}}})^{1/2}$.

Finally we see that for any dual function $\hat{f}(\phi)$ we obtain $\hat{f}(\hat{s})$ such that

$$\frac{d\hat{f}}{d\hat{s}} = \frac{d\hat{f}}{d\phi}\frac{d\phi}{d\hat{s}} = \frac{1}{\hat{v}}\frac{d\hat{f}}{d\phi} \tag{54}$$

The dual Frenet reference frame for $\hat{\mathbf{c}}(\phi)$ is defined in a manner formally identical to the way the Frenet frame of hyperspherical curves is obtained. $\hat{E}$ is chosen in the direction $\hat{\mathbf{c}}$, and $\hat{T}$ in the direction $d\hat{\mathbf{c}}/d\hat{s}$. The direction $\hat{N}$ is chosen along the component of $d\hat{T}/d\hat{s}$ which does not lie in the $\hat{E}$-$\hat{T}$ plane; in the same way $\hat{B}$ is chosen along the component of $d\hat{N}/d\hat{s}$ orthogonal to the $\hat{E}$-$\hat{T}$-$\hat{N}$ subspace. The result of these definitions is the dual Frenet equations for dual hyperspherical curves:

$$\begin{aligned} \frac{d\hat{E}}{d\hat{s}} &= \hat{T} \\ \frac{d\hat{T}}{d\hat{s}} &= -\hat{E} + \hat{\kappa}\hat{N} \\ \frac{d\hat{N}}{d\hat{s}} &= -\hat{\kappa}\hat{T} + \hat{\tau}\hat{B} \\ \frac{d\hat{B}}{d\hat{s}} &= -\hat{\tau}\hat{N} \end{aligned} \tag{55}$$

The dual function $\hat{\kappa}$ and $\hat{\tau}$ together with the dual function $\hat{v}(\phi)$ characterize differential geometry of $\hat{\mathbf{c}}(\phi)$. From (53) we compute

$$\begin{aligned} \hat{v}_0 &= (\hat{\mathbf{c}}_1 \cdot \hat{\mathbf{c}}_1)^{1/2} = \frac{1}{2} - \epsilon\frac{d_{z1}}{2} \\ \hat{v}_1 &= (\hat{\mathbf{c}}_1 \cdot \hat{\mathbf{c}}_2)/\hat{v}_0 = -\epsilon\frac{d_{z2}}{2} \\ \hat{v}_2 &= (\hat{\mathbf{c}}_1 \cdot \hat{\mathbf{c}}_3 + \hat{\mathbf{c}}_2 \cdot \hat{\mathbf{c}}_2)/\hat{v}_0 = \frac{1}{2}\epsilon[d_{z1}e^2 - 2e\,d_{y2} - d_{z3}] \end{aligned} \tag{56}$$

Furthermore using (37) and (38) we compute

$$\hat{\kappa} = 2e + \epsilon 2(d_{y2} + 2d_{z1}\,e) \tag{57}$$

and

$$\hat{\tau} = -\frac{2\gamma_x}{e} - \epsilon\frac{1}{e^2}[2\gamma_x(d_{z1}e + d_{y2}) + 3d_{z1}e^2 - (3d_{y2} + 2d_{x3})e] \tag{58}$$

and finally (42) yields

$$\hat{\kappa}' = 4\gamma_y + \epsilon 2[4d_{z1}\gamma_y + 3d_{z2}e - d_{y3}]. \tag{59}$$

These computations were facilitated using the symbolic computation software MACSYMA with functions defined to manipulate dual numbers and wedge products.

In order to completely characterize a general motion to third order nine independent instantaneous invariants are required. The geometric parameters defining the curve $\hat{\mathbf{c}}(\phi)$ to third order also provide nine independent constants: the three dual parts of $\hat{v}(\phi)$ and the six constants which form the real and dual parts of $\hat{\kappa}$, $\hat{\tau}$ and $\hat{\kappa}'$. The equations (56)–(59) define the relationships between these two sets of constants.

Up to this point, we have derived the relationship between the instantaneous invariants and the intrinsic properties (curvature, torsion, rate of change of curvature) of image curves of spherical and spatial motions. These relationships allow determination of instantaneous invariants of a motion from its image curve without the need for transforming the motion to the canonical coordinate system.

In the next few sections, we take a different view point. We study the instantaneous kinematics of a motion directly in terms of the intrinsic properties of its image curve. We do this only for spherical motions.

Local Properties of Point Trajectories of Spherical Motions. Having obtained geometric parameters which characterize the local properties of the image curve $\mathbf{c}(\phi)$ (κ, κ' and τ), we substitute these back into the series expansion of the orthogonal matrix $A(\phi)$ and study geometric properties of point trajectories generated by spherical motion. The A_2 and A_3 terms in (9*b*) are now given by

$$A_2 = \left| \begin{matrix} -1 & 0 & \kappa/2 \\ 0 & -1 & 0 \\ -\kappa/2 & 0 & 0 \end{matrix} \right|$$

and

$$A_3 = \left| \begin{matrix} 0 & (1+\kappa^2/4) & \kappa'/4 \\ -(1+\kappa^2/4) & 0 & \kappa(3+\tau)/4 \\ -\kappa'/4 & \kappa(3-\tau)/4 & 0 \end{matrix} \right|$$

If $\boldsymbol{\alpha}_p(\phi)$ is the trajectory of $\mathbf{p}:(x,y,z)$ fixed in the moving frame then we have to the third order

$$\boldsymbol{\alpha}_p(\phi) = \left\{ \begin{matrix} x \\ y \\ z \end{matrix} \right\} + \left\{ \begin{matrix} -y \\ x \\ 0 \end{matrix} \right\} \phi + \left\{ \begin{matrix} -x + (\kappa/2)z \\ y \\ -(\kappa/2)x \end{matrix} \right\} \phi^2/2$$
$$+ \left\{ \begin{matrix} (1+\kappa^2/4)y + (\kappa'/4)z \\ -(1+\kappa^2/4)x + \frac{\kappa}{4}(3+\tau)z \\ -(\kappa'/4)x + \frac{\kappa}{4}(3-\tau)y \end{matrix} \right\} \frac{\phi^3}{6} \qquad (62)$$

We can now determine the geodesic curvature γ of $\boldsymbol{\alpha}_p$ which is defined by the relation

$$\gamma_p = \frac{\alpha \cdot \dot{\alpha} \times \ddot{\alpha}}{(\dot{\alpha} \cdot \dot{\alpha})^{3/2}} = \frac{z(x^2+y^2) - (\kappa/2)x(x^2+y^2+z^2)}{(x^2+y^2)^{3/2}} \qquad (63)$$

The set of points $\mathbf{p}:(x,y,z)$ in the moving frame with the property that $\gamma_p = 0$ satisfy the relation

$$2z(x^2+y^2) - \kappa x(x^2+y^2+z^2) = 0 \qquad (64)$$

which is the well-known inflection cone now written in terms of the curvature κ of the image curve $\mathbf{c}(\phi)$.

The rate of change of geodesic curvature γ_p' of the trajectory α_p is given by

$$\gamma_p' = \frac{(\dot{\alpha} \cdot \dot{\alpha})(\alpha \cdot \ddot{\alpha} \times \dddot{\alpha}) - 3(\dot{\alpha} \times \ddot{\alpha})(\alpha \cdot \dot{\alpha} \times \ddot{\alpha})}{(\dot{\alpha} \cdot \dot{\alpha})^{5/2}}$$
$$= \frac{(x^2+y^2+z^2)}{4(x^2+y^2)^{5/2}} \{ [\kappa' x + \kappa(3-\tau)y](x^2+y^2) - 3\kappa^2 xyz \} \qquad (65)$$

The set of points for which $\gamma_p' = 0$ is given by

$$[\kappa' x + \kappa(3-\tau)y](x^2+y^2) - 3\kappa^2 xyz = 0 \qquad (66)$$

which is the well known cubic cone of stationary geodesic curvature now written in terms of κ, κ' and τ of the image curve $\mathbf{c}(\phi)$.

Equations (64) and (65) form the link between the geometric properties of the image curve and those of the curves traced by points in the moving body. For example if $\kappa = 0$ the image curve locally follows a great circle of the hypersphere to the third order. Motions with this property have degenerate loci $\gamma_p = 0$ and $\gamma_p' = 0$: the first locus becomes $z = 0$, the second $x = 0$. If $\kappa' = 0$ the image curve has constant curvature to the third order and (66) divides into the two loci $y = 0$ and $x^2 + y^2 - (3\kappa/(3-\tau))xz = 0$, a plane and a circular cone respectively. The value $\tau = 3$ does not seem to have any special geometrical meaning for the image curve though it causes (66) to degenerate into the plane $x = 0$ and circular cone $x^2 + y^2 - (3\kappa^2/\kappa')yz = 0$. Finally, if $\kappa' = 0$ and $\tau = 3$, (66) becomes the three planes $x = 0$, $y = 0$ and $z = 0$.

The curves (64) and (66) are important in the instantaneous synthesis of spherical four bar linkages since they identify points which may be used for the moving pivots. All the results of instantaneous spherical kinematics can be cast into form which links the results directly to geometric properties of the image curve.

The Axodes of Spherical Motion. The axodes of spherical motion are cones with vertices at the origin. One π_F is fixed in the fixed reference and the other π_M is fixed in the moving frame. The spherical motion may be considered to be generated as π_M rolls without slipping over π_F. The intersections of the axodes with the unit sphere are called the polhodes. In this section, we derive the equations for the fixed and moving polhodes; we will also use π_F and π_M to denote the respective polhodes.

Points on the moving polhodes π_M are those which have zero velocity at some time during the motion. From the velocity equation

$$\dot{\alpha}_p(\phi) = \dot{A}\mathbf{p} \qquad (67)$$

we seek $\mathbf{p}$ in M such that $\dot{\alpha}_p = 0$. This is easily obtained by multiplying (67) by A^T and setting $\dot{\alpha}_p = 0$ to yield

$$\mathbf{0} = A^T \dot{A}\, \mathbf{p} = {}^M\Omega(\phi)\mathbf{p}. \qquad (68)$$

Since ${}^M\Omega$ is skew-symmetric the solution of (68) for $\mathbf{p}$ is simply the vector formed from ${}^M\Omega$. Carrying out this computation we obtain to the second order

$$\pi_M(\phi) = \left\{ \begin{matrix} 0 \\ 0 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 0 \\ \kappa/2 \\ 0 \end{matrix} \right\} \phi + \left\{ \begin{matrix} \kappa(1-\tau)/4 \\ \kappa'/4 \\ -\kappa^2/4 \end{matrix} \right\} \phi^2/2 \qquad (69)$$

The equation for the fixed polhode is obtained in a similar way. We substitute $A^T p$ for $\mathbf{p}$ in (67) and set $\dot{\alpha}_p = 0$ to yield

$$\mathbf{0} = \dot{A} A^T P = {}^F\Omega(\phi) P \qquad (70)$$

The solution for P is the fixed polhode:

$$\pi_F(\phi) = \left\{ \begin{matrix} 0 \\ 0 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 0 \\ \kappa/2 \\ 0 \end{matrix} \right\} \phi + \left\{ \begin{matrix} -\kappa(1+\tau)/4 \\ \kappa'/4 \\ -(1+\kappa^2/4) \end{matrix} \right\} \frac{\phi^2}{2} \qquad (71)$$

The geodesic curvatures γ_{π_M} and γ_{π_F} of these two curves are easily computed using (63) and we have

$$\gamma_{\pi_M} = \frac{-(1-\tau)}{\kappa} \text{ and } \gamma_{\pi_F} = \frac{(1+\tau)}{\kappa}. \tag{72}$$

Analytically Defined Spherical Motion. Bottema and Roth (1979) survey a wide class of motions which are defined by setting the independent parameters of a motion equal to various functions of a single parameter t. An example of this is the particular case of a symmetric spherical motion defined by setting the Euler parameters $c_i(t)$, $i = 0,1,2,3$ equal to the functions

$$\begin{aligned} c_0(t) &= 0 \\ c_1(t) &= a(1-t^2) \\ c_2(t) &= 2bt \\ c_3(t) &= c(1+t^2) \end{aligned}$$

where a, b and c are constants. This is the coupler motion of a special spherical four bar linkage.

We introduce this example to illustrate the generality of local analysis using the image curve $\mathbf{c}(t) = (c_1(t), c_2(t), c_3(t), c_0(t))$. Equation (73) defines $\mathbf{c}(t)$ as a curve in the projective space Σ'. In order to examine $\mathbf{c}(t)$ on the hypersphere we normalize it

$$\mathbf{X}(t) = \frac{\mathbf{c}(t)}{|\mathbf{c}(t)|} \tag{74}$$

and compute its curvature

$$\kappa(t) = 8abc\left[\frac{(c^2+a^2)t^4 + (2c^2+4b^2-2a^2)t^2 + (c^2+a^2)}{(4b^2c^2+4a^2b^2)t^4 + ((16a^2-8b^2)c^2+8a^2b^2)t^2 + 4b^2c^2 + 4a^2b^2}\right]^{3/2} \tag{75}$$

and torsion

$$\tau(t) = 0. \tag{76}$$

The torsion $\tau(t)$ of $\mathbf{X}(t)$ is zero as a result of the fact that $\dddot{\mathbf{c}} = 0$, (see equation 38). At the instant $t = 0$, we have

$$\kappa(0) = 8abc\left[\frac{c^2+a^2}{4b^2(c^2+a^2)}\right]^{3/2} = \frac{ac}{b^2} \tag{77}$$

and we conclude that the instantaneous invariants at this instant are:

$$\epsilon = \frac{2ac}{b^2}, \quad \gamma_x = 0, \quad \gamma_y = 0 \tag{78}$$

We can also determine, for example, the curvature properties of point trajectories at the instant $t = 0$, in view of equation (63) are given by

$$\gamma_p = [z(x^2+y^2) - \left(\frac{ac}{2b^2}\right)x(x^2+y^2+z^2)]/[x^2+y^2]^{3/2} \tag{79}$$

The rate of change of geodesic curvature of point paths of the motion at this instant then becomes (see equation 65):

$$\gamma'_p = \frac{x^2+y^2+z^2}{4(x^2+y^2)^{5/2}}\left[\frac{3ac}{b^2}y(x^2+y^2) - \frac{3a^2c^2}{b^4}xyz\right]$$

We can also determine the polhodes of the motion and the geodesic curvatures of the moving and fixed polhodes at any instant.

Conclusions

In this paper, we have shown that kinematic mapping can be used as an elegant geometric tool to study local kinematics of spherical and spatial motions. Similar to instantaneous invariants of a motion, the mapping curve also provides constants that uniquely characterize a motion. We have applied the results to the study of instantaneous kinematics of spherical motions and have cast important results of spherical curvature theory in terms of the intrinsic properties of image curves of spherial motions.

Acknowledgment

This work was supported by NSF grant MEA-8315796 to the University of Pennsylvania and by U. S. Army Research Office grant DAAG29-84-K-0182, to the University of Wisconsin—Madison.

References

Blaschke, W., and Müller, H. R., 1956, *Ebene Kinematik*, Munich.

Bottema, O. and Roth, B., 1979, *Theoretical Kinematics*, North-Holland, New York, 550 pp.

DeSa, S., and Roth, B., 1981a, "Kinematic Mappings, Part 1: Classification of Algebraic Motions in the Plane," ASME *J. of Mechanical Design*, Vol. 103, No. 3, July 1981, pp. 585–591.

DeSa, S., and Roth, B., 1981b, "Kinematic Mappings, Part 2: Rational Algebraic Motions," ASME *J. of Mechanical Design*, Vol. 103, No. 4, Oct. 1981, pp. 712–717.

Dimentberg, F. M., 1965, *The Screw Calculus and Its Application in Mechanics* (in Russian), Moscow 1965 (English translation: AD680993, Clearinghouse for Federal Technical and Scientific Information, Virginia).

Flanders, H., 1963, *Differential Forms With Applications to the Physical Sciences*, Academic Press, New York.

Herman, R., 1966, *Lie Groups for Physicists*, W. A. Benjamin, New York.

Kirson, Y., 1975, *Higher Order Curvature Theory for Space Kinematics*, Ph.D. Dissertation, University of California, Berkeley, Sept. 1975.

McCarthy, J. M., 1983, "Differential Kinematics on a 3-Sphere," *Proc. of the Sixth World Congress on the Theory of Machines and Mechanisms*, New Delhi, India, Dec. 1983.

O'Neill, B., 1966, *Elementary Differential Geometry*, Academic Press, (Harcourt Brace Jovanovitch), pp. 411.

Ravani, B., and Roth, B., 1983, "Motion Synthesis Using Kinematic Mappings," ASME *J. of Mechanisms, Transmissions and Automation in Design*, Vol. 105, No. 3, Sept. 1983, pp. 460–467.

Ravani, B., and Roth, B., 1984, "Mappings of Spatial Kinematics," ASME *J. of Mechanisms, Transmissions, and Automation in Design*, Vol. 196, No. 3, Sept. 1984, pp. 341–347.

Roth, B., and Yang, A. T., 1977, "Applications of Instantaneous Invariants to the Analysis and Synthesis of Mechanisms," ASME *J. of Engineering for Industry*, Vol. 99, 1977, pp. 97–103.

Study, E., 1891, "Von den Bewegungen und Umlegungen," *Math. Ann.*, Vol. 39.

Veldkamp, G. R., 1967, "Canonical Systems and Instantaneous Invariants in Spatial Kinematics," *J. of Mechanism*, Vol. 2, No. 3, 1967, pp. 329–333.

Veldkamp, G. R., 1975, "On the Use of Dual Numbers, Vectors and Matrices in Instantaneous Spatial Kinematics," *Mechanisms and Machine Theory*, Vol. 11, pp. 141–156.

S. B. Skaar
Assistant Professor.

D. Tucker
Graduate Assistant.

Department of Engineering Science
and Mechanics,
Iowa State University,
Ames, Iowa 50011

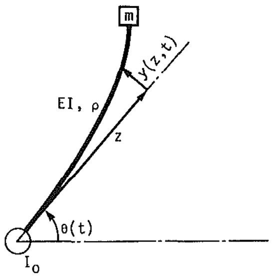
Point Control of a One-Link Flexible Manipulator

An alternative approach to the control of nonrigid, distributed parameter systems is presented. Transfer functions that relate the response of points on the system to a controlling force or torque are used in place of ordinary differential equations, which represent an approximation to the system dynamics. The implications of this "point control" approach are discussed with regard to plant modeling accuracy, uncontrolled regions, open-loop and closed-loop control strategies, system identification, and feedback estimation. Sample optimal control histories are illustrated for a single-link manipulator member with end load.

Introduction

Recently, efforts have been made to control maneuvers of mechanical systems which cannot be adequately modeled using a rigid body assumption for all or some of the system components, especially in the fields of satellite attitude control [1, 2] and robotics [3]. With few exceptions (e.g., [4]) these efforts deal with the distributed nature of such systems by attempting to develop an approximate model of the entire system using a finite number of time-dependent variables that are influenced by the control(s) according to a set of ordinary differential equations (ODEs). This general approach (which includes lumped parameter methods and controlled-modes methods) has two primary advantages and three significant disadvantages when compared with an alternative approach developed herein. The advantages of conventional discretization schemes are (1) that a large body of theory already exists to produce both optimal open-loop trajectories and closed-loop control laws for systems governed by ODEs, and (2) that, if control is successful, it is complete (in the sense that the entire system is modeled and controlled rather than a few regions or points within the system). The three disadvantages of the finite-dimensionalization approach are related to (1) the difficulty of adequately modeling a distributed parameter plant using a small number of variables, (2) the difficulty of acquiring from a few sensors accurate, real-time estimates of the controlled variables for state feedback, and (3) the difficulty of system identification.

An alternative method described below makes use of transfer functions which relate the response of one or more points (or rigid regions) within a distributed parameter system to the controlling input(s). The preferability of this system description from the standpoints of plant model accuracy, state estimation, and identification is discussed as the method is introduced immediately below. The issue of constructing optimal open-loop trajectories that control points of interest within the distributed parameter system is examined at length. (A similar approach is used to generate open loop trajectories for a second order system in [5].) Examples of optimal open-loop trajectories are provided. The availability of closed-loop control laws using the convolution integral descriptions of motion is discussed next. Finally, the problem of uncontrolled regions within the system is examined empirically from the open-loop test cases.



L = MANIPULATOR ARM LENGTH
EI = " " BENDING STIFFNESS
ρ = " " MASS/LENGTH
I_0 = RIGID BASE MASS MOMENT OF INERTIA
m = MASS OF END LOAD
θ = ANGULAR POSITION OF BASE
y = DEFLECTION FROM EQUILIBRIUM
z = POSITION LOCATION ALONG ARM
t = TIME (MEASURED FROM BEGINNING OF MANEUVER)

Fig. 1 Manipulator arm model

The Control Response Description

Rather than a set of linear ordinary differential equations, the present method makes use of a convolution integral description of the response of controlled points on a flexible system to one or more controlling inputs. Such convolution integrals are derivable for the single-member manipulator arm model of Fig. 1, the several system configurations discussed in [6, 7], and a great many other linear, nonrigid systems. The general forms of the integrals are

$$x(t) = \int_0^t u(\lambda) g(t-\lambda) d\lambda \tag{1}$$

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October 26, 1984; final revision, April 15, 1985.

$$\dot{x}(t) = \int_0^t u(\lambda)\dot{g}(t-\lambda)d\lambda \tag{2}$$

where

$$[\dot{\ }](t) = d/dt[\ \ ](t) \tag{3}$$

and where $g(t)$ is of the general form

$$g(t) = C_0 t + \sum_{n=1}^{\infty} C_n \exp[s_n t] + \bar{C}_n \exp[\bar{s}_n t] \tag{4}$$

The constants C_n and s_n are complex. If x is the deflection of some point on the system measured with respect to a reference point moving with the same system, C_0 is generally zero. If x is an absolute position, C_0 is nonzero and real. In the absence of damping, the s_n are purely imaginary.

The function g of equation (4) is the inverse Laplace transform of a transfer function $X(s)/U(s)$. This function is derived below for any point of interest on the system of Fig. 1. The derivation represents an exact solution to the governing partial differential equations. No order truncation of the system dynamics is required until the actual trajectory is determined. Furthermore, a large number of terms can be retained in the response descriptions (equation (4)) with little incremental computational cost. Hence, the "spillover" problem inherent in the controlled-modes methods [1] is avoided.

For the case of actual systems in which the first several parameters of equation (4) must be identified experimentally, sensors would, in principle, be required only for the purpose of measuring the input force and the corresponding acceleration of points to be controlled. This contrasts with the more demanding identification requirement of lumped parameter models in which sensors must be located at each of the modeled degrees of freedom [8]. Similarly demanding identification measures are needed for systems modeled with modal amplitude ODEs. Such requirements may be impractical, for instance, on an orbiting satellite.

The Manipulator Arm Integrals

Consider the system of Fig. 1. Along the length of the arm ($0 \leq z \leq L$), the modulus of elasticity (E), the transverse area moment of inertia (I), and the mass per unit length (ρ) are constant. Although the radius of the base axis of rotation is assumed for convenience to be zero, a motor armature and gear box are modeled by way of nonzero rigid mass moment of inertia, I_0, located at this base axis. The end mass, m, (located at the opposite end of the arm) is considered to occupy a point. The control torque, u, is continuously variable.

The variable $y(z, t)$ is the deflection of the arm at a point located a distance z from the torqued end, measured relative to the undeformed position of the arm. The angular displacement, $\theta(t)$, is the angular position of the base measured from its original or reference position.

To achieve point control of this system, four convolution integral types are required, given by

$$\theta(t) = \int_0^t u(\lambda)g(t-\lambda)d\lambda \tag{5}$$

$$\dot{\theta}(t) = \int_0^t u(\lambda)\dot{g}(t-\lambda)d\lambda \tag{6}$$

$$y(z,t) = \int_0^t u(\lambda)g_z(t-\lambda)d\lambda \tag{7}$$

$$\dot{y}(z,t) = \int_0^t u(\lambda)\dot{g}_z(t-\lambda)d\lambda \tag{8}$$

It should be emphasized that g_z and $\dot{g}_z$ in equations (7) and (8) will vary depending upon the choice of the point location z. As will be seen, these convolution integrals are useful not only for determining optimal control solutions, but also for the exact monitoring of the response to any given input of uncontrolled points in a system.

If the deflections are assumed to be small, the angular velocity low, and the member narrow, the governing partial differential equation is linear and given by

$$EI(\partial^4 y/\partial z^4) + \rho[\partial^2 y/\partial t^2 + z(d^2\theta/dt^2)] = 0 \tag{9}$$

Also,

$$I_t(d^2\theta/dt^2) + mL(d^2y(L,t)/dt^2) + \rho\int_0^L z(\partial^2 y/\partial t^2)dz = u \tag{10}$$

where

$$I_t = I_0 + \rho L^3/3 + mL^2 \tag{11}$$

The geometric boundary conditions at the torqued end are

$$y(0,t) = 0 \tag{12}$$

$$\partial y/\partial z\,|_{z=0} = 0 \tag{13}$$

The natural boundary conditions at the free end are

$$\partial^2 y/\partial z^2\,|_{z=L} = 0 \tag{14}$$

$$EI(\partial^3 y/\partial z^3) - m(Ld^2\theta/dt^2 + \partial^2 y/\partial t^2)\,|_{z=L} = 0 \tag{15}$$

The required transfer functions are found by taking a Laplace transformation of equations (9)-(15) in the time domain as follows:

$$EI(d^4 Y/dz^4) + \rho s^2(Y + z\Theta) = 0 \tag{16}$$

$$I_t s^2\Theta + mLs^2 Y(L) + \rho s^2\int_0^L zYdz = U \tag{17}$$

$$Y(0) = 0 \tag{18}$$

$$dY/dz\,|_{z=0} = 0 \tag{19}$$

$$d^2 Y/dz^2\,|_{z=L} = 0 \tag{20}$$

$$ms^2 Y(L) + mLs^2\Theta - EId^3 Y/dz^3\,|_{z=L} = 0 \tag{21}$$

where $Y(z)$, Θ, and U are the Laplace transforms of $y(z,t)$, $\theta(t)$, and $u(t)$, respectively.

A general solution to equation (16) is

$$Y(z) = \exp(\beta z)[A\cos\beta z + B\sin\beta z] + \exp(-\beta z)[C\cos\beta z + D\sin\beta z] - \Theta z \tag{22}$$

where

$$\beta^4 = \rho s^2/4EI \tag{23}$$

The constants A, B, C, and D are evaluated using equations (18)-(21). The resulting solution for Y is then substituted into the definite integral of equation (17), which is evaluated analytically. From this result and equation (22) the transfer functions $G(s)$ and $G_z(s)$ are found, where

$$G = \Theta/U \tag{24}$$

$$G_z = Y(z)/U \tag{25}$$

Both of these rather long expressions have the same purely imaginary poles. Summing the residues about these poles [9] yields g and g_z in the series form of equation (4). In the case of both series, the s_n are imaginary, and, in the case of g_z, C_0 = 0.

Optimal Open-Loop Trajectories

Although the following development is applicable to any system for which the response of points of interest are related to the control input according to equations (1, 2), the method will be illustrated in the context of the single member manipulator of Fig. 1 and equations (5)-(8).

Central to any optimal control problem is the performance index selection. The selection made here for controlling the flexible manipulator is based upon a trajectory planning strategy that has been offered for rigid manipulators [10]. In particular, the suggested trajectory for a rigid version of the single-degree-of-freedom model would interpolate in time between the initial angular position θ_0 at time t_0 and the final desired angular position θ_f at time t_f according to

Table 1 Physical properties of the system

Arm length (L)	– 3.0 m
Cross-sectional dimensions	– 1 cm × 6 cm
End mass (m)	– 1.0 kg
Base mass moment of inertia (I_0)	– 0.2 kg-m^2
Mass per unit length (ρ)	– 0.54 kg/m
Bending stiffness (EI)	– 27.8 N-m^2

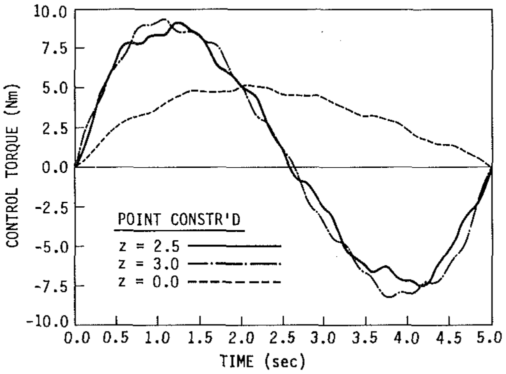


Fig. 2 Optimal control histories

$$\theta(t)=C_0+C_1t+C_2t^2+C_3t^3+C_4t^4+C_5t^5$$

The six constants in this polynomial are evaluated such that the desired angular positions are realized initially and finally, and such that initial and final velocities and accelerations are zero. A seldom observed fact is that the resulting smooth trajectory represents the minimum of

$$I=\int_{t_0}^{t_f}\dot{u}^2(\lambda)d\lambda \qquad (26)$$

where, for the rigid case, the control torque u is related to θ by

$$u=I_t\ddot{\theta}$$

Because of its association with a recommended rigid-manipulator trajectory plan, because it tends to yield smooth control histories, and because it allows for the initial and final specification $u(t_0) = u(t_f) = 0$ (such that initial and final control discontinuities are avoided), the performance index of equation (26) is selected for the flexible control problem.

To develop the necessary conditions for an optimal maneuver, we begin with an assumption that the system is initially at rest. A control function $u^*(t)$ is sought that, over the interval $t_0 = 0$ to t_f, will bring the base of the arm from an initial state of rest at the position $\theta = 0$ to a desired final angular position, θ_f, and velocity, $\dot{\theta}_f$, such that I is minimum.

In addition, N previously selected points ($z_1, z_2, \ldots z_N$) along the length of the arm may be brought simultaneously to a condition of zero deflection and zero rate of deflection. To this end, the performance index of equation (26) is augmented using $2N + 2$ constant Lagrange multipliers, K_i, in conjunction with equations (5)–(8).

Table 2 Trajectory comparisons

Case	Description[a]	Residual energy (J)	$u_{\max}$(N-m)
1	Deflec'n at $z = 2.5$ constrained @ t_f	1.77	8.8
2	Deflec'n at $z = 3.00$ constrained @t_f	1.17	9.0
3	No extra constraints	9.67	5.0

[a] Final conditions of $\theta(t_f)=\pi$, $\dot{\theta}(t_f) = 0$ specified for all three cases.

$$\begin{aligned}I=\int_0^{t_f}\{&\dot{u}^2(\lambda)+K_1[u(\lambda)g_{z_1}(t_f-\lambda)]+K_2[u(\lambda)\dot{g}_{z_1}(t_f-\lambda)]\\&+\ldots+K_{2N-1}[u(\lambda)g_{z_N}(t_f-\lambda)]+K_{2N}[u(\lambda)\dot{g}_{z_N}(t_f-\lambda)]\\&+K_{2N+1}[u(\lambda)g(t_f-\lambda)-\theta_f/t_f]\\&+K_{2N+2}[u(\lambda)\dot{g}(t_f-\lambda)-\dot{\theta}_f/t_f]\}d\lambda=\int_0^{t_f}F(\dot{u},u,\lambda)d\lambda\end{aligned} \qquad (27)$$

Applying the Euler-Lagrange necessary condition to equation (27)

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{u}}\right)-\frac{\partial F}{\partial u}=0 \qquad (28)$$

results in the requirement

$$\begin{aligned}\ddot{u}^*(t)=\frac{1}{2}\{&K_1g_{z_1}(t_f-t)+K_2\dot{g}_{z_1}(t_f-t)+\ldots\\&+K_{2N+1}g(t_f-t)+K_{2N+2}\dot{g}(t_f-t)\}\end{aligned} \qquad (29)$$

Integrating equation (29) yields

$$\begin{aligned}u^*(t)=\frac{1}{2}\int_0^t\int_0^\lambda\{&K_1g_{z_1}(t_f-\xi)+K_2\dot{g}_{z_1}(t_f-\xi)+\ldots\\K_{2N+1}g(t_f-\xi)&+K_{2N+2}\dot{g}(t_f-\xi)\}d\xi d\lambda+C_1t+C_2\end{aligned} \qquad (30)$$

Requiring that $u^*(0)=u^*(t_f) = 0$ leads to

$$C_2=0 \qquad (31)$$

$$C_1=-\frac{1}{2}\int_0^{t_f}\int_0^\lambda\{K_1g_{z_1}(t_f-\xi)+\ldots+K_{2N+2}\dot{g}(t_f-\xi)\}d\xi d\lambda \qquad (32)$$

Thus, the optimal control may be written in the form

$$u^*(t)=K_1f_1(t)+K_2f_2(t)+\ldots+K_{2N+2}f_{2N+2}(t) \qquad (33)$$

where, for instance,

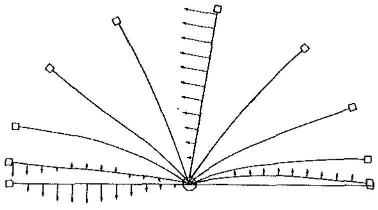
$$f_1(t)=\{\int_0^t\int_0^\lambda g_{z_1}(t_f-\xi)d\xi d\lambda-t\int_0^{t_f}\int_0^\lambda g_{z_1}(t_f-\xi)d\xi d\lambda\}/2 \qquad (34)$$

The multipliers K_i are found by substituting equation (33) into the appropriate forms of equations (5)–(8) evaluated at $t=t_f$. The result is a matrix equation of the form

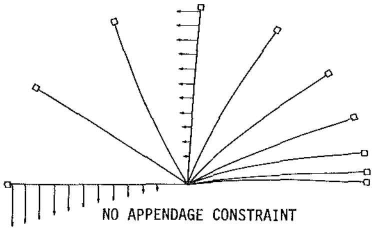
$$\begin{bmatrix}0\\0\\\cdot\\\cdot\\\cdot\\\theta_f\\\dot{\theta}_f\end{bmatrix}=\int_0^{t_f}\begin{bmatrix}f_1(\lambda)g_{z_1}(t_f-\lambda) & f_2(\lambda)g_{z_1}(t_f-\lambda) & \ldots f_{2N+2}(\lambda)g_{z_1}(t_f-\lambda)\\ \cdot & & \cdot\\ \cdot & & \cdot\\ \cdot & & \cdot\\ \\ \\ f_1(\lambda)\dot{g}(t_f-\lambda) & \cdots & f_{2N+2}(\lambda)\dot{g}(t_f-\lambda)\end{bmatrix}d\lambda\begin{bmatrix}K_1\\K_2\\\cdot\\\cdot\\\cdot\\ \\K_{2N+2}\end{bmatrix} \qquad (35)$$

from which K_1 through K_{2N+2} may be found.

To illustrate this method, consider three cases of a 180 deg rotation of the manipulator arm. In each case the total maneuver time, t_f, will be five seconds, with $\dot{\theta}_f = 0$ and $\theta_f = \pi$. The physical properties of the system provided in Table 1 are consistent with an aluminum arm. In case 1, a point 2.5



FINAL CONSTRAINT ON LOADED END



NO APPENDAGE CONSTRAINT

Fig. 3 Mid-trajectory arm positions

meters from the base of the 3-meter arm is controlled; in case 2, the end point is controlled; and in case 3, only the angle and angular velocity of the base are controlled.

The optimal control histories which result from these three cases are illustrated in Fig. 2. It is interesting to note that in the unconstrained case 3 no braking control is required. It is optimal, in fact, for the elastic member itself to supply the braking torque for the hub. Because only positive work is done on the arm for this case, the residual (elastic plus kinetic) energy in the system is high compared with the other two cases. This comparison is illustrated in Table 2 along with a comparison of the peak torque required for each case. The residual energy is computed by integrating the product $u\dot{\theta}$ over the maneuver time.

Figure 3 illustrates an additional comparison between cases 2 and 3. A sequence of member positions determined (via equations (5)–(8)) each 0.5 seconds through the maneuver is shown for these two cases. Certain of the frames indicate the velocity distribution along the member. As is clear from the final velocity profile in each case, most of the residual energy is kinetic rather than elastic. The relatively small deformation at $t = t_f$ probably occurs because $u^*(t_f) \equiv 0$.

The control of a second appendage point reduces the final velocity distribution to the extent that it cannot be detected using the velocity profile scale of Fig. 3. Fifteen terms were required in each infinite series expression in u^* to produce adequate solution convergence.

Closed-Loop Control

An interesting observation that can be made upon examination of the preceding open-loop control cases is that uncontrolled regions can be "brought into line" by controlling only a few points within a system. Many applications of the point control method, however, would require feedback to accommodate possible disturbances, modeling errors, or control errors. One closed-loop control approach which makes use of the convolution integrals, published first in [7], minimizes

$$I' = \int_0^{t_f} u^2(\lambda)d\lambda \tag{36}$$

The resulting trajectories tend to be unsmooth. The counterpart to this control law which minimizes

$$I = \int_0^{t_f} \dot{u}^2(\lambda)d\lambda \tag{37}$$

is developed below. The goal is as it was for the open-loop case; starting from a state of rest, bring the base and N appendage points to a prescribed position and velocity at $t = t_f$ while minimizing I. Now, however, the angular acceleration of the base, $\ddot{\theta}_m(t)$, is continuously available. (It should be noted that a similar development to that given below could be used for cases of angular position or angular velocity feedback.)

Consider the mid-maneuver time $t = T$, $0 < T < t_f$. Clearly, "I" will be minimum if $u(t)$, $T \le t \le t_f$, is selected such that

$$J(T) = \int_T^{t_f} \dot{u}^2(\lambda)d\lambda \tag{38}$$

is minimized subject to the final requirements on $\theta(t_f)$, $\dot{\theta}(t_f)$, and the like. In addition, u must be selected such that $u(t_f) = 0$ and $u(T)$ is continuous.

The necessary condition for minimizing equation (38) gives rise to equation (29) and equation (30) as before. However, C_1 and C_2 are no longer determined simply, as in equations (31) and (32). Furthermore, the value of the multipliers will change depending upon the prior acceleration history, $\ddot{\theta}_m(t)$, sensed at the base.

To illustrate this, consider the following form of the integral in equation (5) evaluated at $t = t_f$:

$$\theta_f = P_\theta(T) + \int_T^{t_f} u(\lambda)g(t_f - \lambda)d\lambda \tag{39}$$

where

$$P_\theta(T) = \int_0^T u_a(\lambda)g(t_f - \lambda)d\lambda \tag{40}$$

and where $u_a(\lambda)$ $(0 \le \lambda < T)$ is the "actual," sensed, control-plus-disturbance torque acting at the base. This "actual" torque can be determined from the measured acceleration $\ddot{\theta}_m(t)$ according to

$$u_a(t) = \int_0^t \ddot{\theta}_m(\lambda)\bar{g}(t - \lambda)d\lambda \tag{41}$$

where $\bar{g}$ may be found from equation (24) such that

$$\bar{g} = \mathcal{L}^{-1}\{[s^2G(s)]^{-1}\}$$

The sinusoidal form of $\bar{g}$ allows for the rearrangement of equation (41) into two integrals with time-varying coefficients of the form

$$u_a(t) = \sum_{n=1}^{\infty} \{h_n(t)\int_0^t \ddot{\theta}_m(\lambda)p_n(\lambda)d\lambda + r_n(t)\int_0^t \ddot{\theta}_m(\lambda)q_n(\lambda)d\lambda\} \tag{42}$$

Equations (40) and (42) may be integrated numerically throughout the maneuver to continually produce new updates for P_θ. Analogous quantities ($P_{\dot{\theta}}$, P_{y_1}, $P_{\dot{y}_1}$, P_{y_2}, etc.) can be computed for the other controlled quantities ($\dot{\theta}(t_f)$, $y(z_1, t_f)$, $\dot{y}(z_1, t_f)$, $y(z_2, t_f)$, etc.).

At time T, then, the multipliers K_i, along with the integration constants C_1 and C_2, can be reassessed using the matrix equation

$$\begin{bmatrix} -P_{y_1}(T) \\ -P_{\dot{y}_1}(T) \\ \cdot \\ \cdot \\ \cdot \\ \theta_f - P_\theta(T) \\ \dot{\theta}_f - P_{\dot{\theta}}(T) \\ u(T) \\ 0 \end{bmatrix} = [Q(T)] \begin{bmatrix} K_1 \\ K_2 \\ \cdot \\ \cdot \\ \cdot \\ K_{2N+1} \\ K_{2N+2} \\ C_1 \\ C_2 \end{bmatrix} \tag{43}$$

The elements of $[Q(T)]$ in equation (43) are given by

$$\begin{bmatrix} \int_T^{t_f} d_1(\lambda) g_{z_1}(t_f-\lambda)d\lambda & \int_T^{t_f} d_2(\lambda) g_{z_1}(t_f-\lambda)d\lambda & \cdots & \int_T^{t_f} \lambda g_{z_1}(t_f-\lambda)d\lambda & \int_T^{t_f} g_{z_1}(t_f-\lambda)d\lambda \\ \int_T^{t_f} d_1(\lambda)\dot{g}_z(t_f-\lambda)d\lambda & & \cdots & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ d_1(T) & d_2(T) & & T & 1 \\ d_1(t_f) & d_2(t_f) & & t_f & 1 \end{bmatrix} \tag{44}$$

where, from equation (30),

$$d_1(t) = \int_0^t \int_0^\lambda g_{z_1}(t_f-\xi)d\xi d\lambda \tag{45}$$

$$d_2(t) = \int_0^t \int_0^\lambda \dot{g}_{z_1}(t_f-\xi)d\xi d\lambda \tag{46}$$

and so on.

The frequency with which the multipliers are updated may vary depending upon the computational facilities which are available. The matrix $[Q(T)]$ is the same for any given maneuver regardless of the measured $\ddot{\theta}_m(t)$. Thus, $[Q(T)]^{-1}$ may be developed well in advance of time T, making plausible a real time implementation of the algorithm. The optimal closed-loop control takes the form of equation (30) in which the most current values for K_i and C_i are used.

This control law is, as stated, predicted upon the assumptions that all disturbances act on the rigid base, and that the rest of the system is modeled perfectly. Further investigation is required to determine the behavior of the algorithm when these assumptions do not hold.

Discussion

An alternative approach to the maneuver control of nonrigid systems is introduced. Rather than attempting to achieve a desired final distributed state of the system, the method makes use of transfer functions to control only a few points within the system. As demonstrated in an open-loop maneuver simulation, a judicious selection of the number and location of these points can result in acceptably good alignment of the uncontrolled regions at the end of the maneuver.

The approach offers certain inherent advantages over conventional finite-dimensionalization schemes in the areas of system modeling accuracy, system identification, and feedback estimation. The principal disadvantage of the approach lies in the absence of proven, stable, and robust feedback control algorithms that are based upon the convolution integral descriptions of motion.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. MEA-8318867.

References

1 Meirovitch, L., Baruh, H., and Öz, H., "A Comparison of Control Techniques for Large Flexible Systems," *Journal of Guidance, Control, and Dynamics*, Vol. 6, No. 4, July-Aug., 1983, pp. 302-310.

2 Turner, J. D., and Chun, H. M., "Optimal Distributed Control of a Flexible Spacecraft During a Large-Angle Maneuver," *Journal of Guidance, Control, and Dynamics*, Vol. 7, No. 3, May-June, 1983, pp. 257-264.

3 Schmitz, E., and Cannon, R., "Further Experiments of the End-Point Control of a One-Link Flexible Manipulator," paper presented at the 1984 ASME International Computers in Engineering Conference and Exhibit, Las Vegas, Nevada, Aug. 12-15, 1984.

4 Breakwell, J. A., "Optimal Control of Distributed Systems," *Journal of the Astronautical Sciences*, Vol. XXIX, No. 4, Oct.-Dec., 1981, pp. 472-479.

5 Skaar, S. B., "Closed Form Optimal Control Solutions for Continuous Linear Elastic Systems," *Journal of the Astronautical Sciences*, Vol. 32, No. 4, Oct.-Dec. 1984, pp. 447-461.

6 Skaar, S. B., and Tucker, D., "The Optimal Control of Flexible Systems Using a Convolution Integral Description of Motion," *Proceedings*, 22nd IEEE Conference on Decision and Control, San Antonio, Texas, Dec. 14-16, 1983, pp. 825-829.

7 Skaar, S. B., and Tang, L., "An Optimal Single-Sensor Feedback Control Law for Flexible Dynamic Systems," *Proceedings*, 25th Structures, Structural Dynamics, and Materials Conference, and AIAA Dynamics Specialists Conference, Palm Springs, Calif., May 14-18, 1984, pp. 825-829.

8 Hendricks, S. L., Rajaram, S., Kamat, M. P., and Junkins, J. L., "Identification of Large Flexible Structures Mass/Stiffness and Damping from On-Orbit Experiments," *Journal of Guidance, Control, and Dynamics*, Vol. 7, No. 2, Mar.-Apr., 1984, pp. 244-245.

9 Wylie, C. Ray, *Advanced Engineering Mathematics*, 4th edition, McGraw-Hill, New York, 1975, pp. 820-823.

10 Brady, M., ed., *Robot Motion: Planning and Control*, MIT Press, Cambridge, Mass., 1982, pp. 223-25.

Zhen-ni Wang
Doctor of Engineering,
Beijing University of
Aeronautics and Astronautics,
Beijing, China

Tong Fang[1]
Visiting Professor,
Duke University,
Durham, NC 27706

A Time-Domain Method for Identifying Modal Parameters

A time-domain method for identifying the modal parameters of a vibration system is presented. It is shown that system eigenvectors can be effectively estimated through the multivariate AR *model representation of the system response to white noise excitation. In contrast to the usual* ARMA *model approach, in this method only a linear least square algorithm is required, so that a great amount of calculation is saved. Results of digital simulations support the identification method.*

1 Introduction

Modal parameter identification is crucial for linear system identification (Berman, 1979). The key point in identification is to determine the relationship between the system parameters and the measured dynamical data. There are different ways to relate the modal parameters to the measured data, which leads to different identifying methods.

Modal parameters can be identified either by a frequency-domain method or by a time-domain method. These two methods are often complementary to each other. Each of them has its own features and merits. They not only provide different approaches to the problem, but also reveal deeper knowledge from different respects. The frequency-domain method for system identification is relatively mature and has been widely used in engineering research. The time-domain one is still developing.

Because in practice the system input data are often unavailable, in recent years attention has been paid to system identification when only output data are available. Two main methods have appeared in time-domain identification in this respect. One is the ITD method (Ibrahim et al., 1977a, 1977b), by which the system parameters are identified from the free vibration data. Another one is the ARMA model method (Gersch, 1975; Wu, 1977), which regards the random response as the time series of an ARMA process and identifies the system parameters by the ARMA model. Yet the latter has been applied mainly to identifying the eigenvalues, i.e., estimating the natural frequencies and damping of the system (Gersch and Liu, 1976; Pandit and Wu, 1983).

In this paper a new method is presented for identifying the system eigenvectors by the ARMA model. Since eigenvectors must be identified in a multidimensional space, usually one has to use a multivariate ARMA model. However, determining a multivariate ARMA model is laborious, since a nonlinear least-square estimation or two-stage least-square procedure is required. In our approach, a multivariate AR model is determined first from the random response data. Then it is used to estimate the Green function matrices of the corresponding ARMA model. Based on their relation to these Green functions, the modal parameters are then identified. In this approach it is not necessary to determine the ARMA model itself, so that the laborious computation for its modeling is saved. Results of digital simulations support this identification method.

2 The Multivariate ARMA Model and Its Main Properties

Consider a zero mean-valued k-dimensional (k may be any finite positive integer) stationary random sequence y_t, satisfying the following equation:

$$y_t - \sum_{i=1}^{p} a_i y_{t-i} = w_t - \sum_{j=1}^{q} c_j w_{t-j} \tag{1}$$

where a_i and c_j are $k \times k$ matrices, and w_t is a k-dimensional white noise sequence. It is assumed that w_t is uncorrelated to y_j when $j < t$, and has the properties:

$$E[w_t] = 0, \quad E[w_t w_s^T] = D\delta_{ts}$$

where D is a $k \times k$ real positive-definite matrix, and δ_{ts} is a Kronecker symbol.

By defining a backward shift operator B as

$$By_t = y_{t-1}$$

and letting

$$A(B) = I - \sum_{i=1}^{p} a_i B^i$$

$$C(B) = I - \sum_{j=1}^{q} c_j B^j$$

equation (1) may be written as

$$A(B) y_t = C(B) w_t \tag{2}$$

The random sequence y_t is said to be stationary (Priestly, 1981), if and only if "all the zeros of det$A(B)$ lie outside the unit circle" (3*a*) and y_t is said to be invertible, if and only if "all the zeros of det$C(B)$ lie outside the unit circle" (3*b*).

---

[1] On leave from Northwestern Polytechnical University, China.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 8, 1984; final revision, May 8, 1985.

Satisfying equation (1) and the stationary and invertible condition (3), y_t is called a k-variate ARMA sequence, or said to be governed by a k-variate ARMA model.

The roots of $\det A(B) = 0$ are defined as the eigenvalues of the k-variate ARMA model.

It can be proved (Wang, 1984) that, similar to the univariate one, the multivariate ARMA model also has its transfer form and inverted form:

$$y_t = A^{-1}(B)C(B)w_t = \sum_{j=0}^{\infty} G_j w_{t-j}$$

$$= \sum_{j=0}^{\infty} G_j B^j w_t = G(B)w_t, \quad G_0 = I \tag{4}$$

$$w_t = C^{-1}(B)A(B)y_t = -\sum_{j=0}^{\infty} L_j y_{t-j}$$

$$= -\sum_{j=0}^{\infty} L_j B^j y_t = L(B)y_t, \quad L_0 = -I \tag{5}$$

Some important properties of the Green functions and Inverse functions are as follows:

1. G_j in equation (4) is called the Green function matrix (or Impulse response matrix), which describes the "memory" effect of the system to the impulse, which acted at the instant $(t-j)$. The convergence of $\{G_j\}$ decribes the stability of the system.

Substituting equation (4) into equation (2), we have

$$A(B)G(B) = C(B)$$

Comparing the terms on the both sides with the same power in B, we have

$$A(B)G_j = 0, \quad j > q \tag{6}$$

The general solution of equation (6) has the form:

$$G_j = \sum_{i=1}^{k\times p} E_i R_i^j, \quad j > q \tag{7}$$

where R_i's are inverses of the eigenvalues of the ARMA model, and coefficient matrices E_i's are determined by the initial conditions, G_j's, $j \le q$. Since $G_j = 0$, for $j < 0$; when $q < k\times p$, the solution of equation (6), satisfying the initial conditions, may be written as

$$G_j = \begin{cases} \sum_{i=1}^{k\times p} E_i R_i^j, & \text{when } j \ge 0 \\ 0, & \text{when } j < 0 \end{cases} \tag{8}$$

2. L_j in equation (5) is called an Inverse function matrix. From equations (4) and (5), we obtain

$$G(B)L(B) = A^{-1}(B)C(B)C^{-1}(B)A(B) = I$$

or

$$(I + G_1 B + G_2 B^2 + \ldots)(I - L_1 B - L_2 B^2 - \ldots) = I$$

Comparing the terms on the both sides with the same power in B, we obtain

$$G_1 = L_1, \quad G_2 = L_1 G_1 + L_2, \ldots\ldots$$

$$G_n = \sum_{j=1}^{n-1} L_{n-j} G_j + L_n, \ldots\ldots \tag{9}$$

3 The Relation Between Damped Linear System and Its ARMA Model

Suppose that the differential equation of motion of a damped linear system may be written as

$$m\ddot{y}(t) + c\dot{y}(t) + ky(t) = w(t) \tag{10}$$

where matrices m, c, and k are all assumed to be $n\times n$ real symmetrical positive-definite matrices.

The transfer function matrix of the system (10) can be expressed as

$$H(s) = \sum_{i=1}^{2n} U_i U_i^T / s_i(s - s_i)$$

where s_i's are the system eigenvalues, and U_i's are corresponding eigenvectors, which are complex in general.

The impulse response matrix of system (10) can be expressed as

$$h(t) = \mathcal{L}^{-1}H(s) = \begin{cases} \sum_{i=1}^{2n} e^{s_i t} U_i U_i^T / s_i, & t \ge 0 \\ 0, & t < 0 \end{cases}$$

The stationary response of system (10) is

$$y(t) = \int_0^{\infty} h(u) w(t-u) du$$

Putting it in discrete form, we have

$$y_t = \sum_{j=0}^{\infty} h_j w_{t-j} \tag{11}$$

where

$$y_t = y(tv), \quad v\text{: sampling time interval}$$

$$w_{t-j} = w(tv - jv)$$

$$h_j = \int_{tv}^{(t+1)v} h(u) du$$

$$= \sum_{i=1}^{2n} e^{js_i v} U_i U_i^T (e^{s_i v} - 1)/s_i^2 \tag{12}$$

under the assumption that $w(t)$ is constant over the sampling time interval.

It is known that the correlation function of the response of a damped linear system under white noise excitation is asymptotically stationary. Hence, there always exists a stationary random response, for which the discretized version is a stationary random sequence.

Since the discretized random response of system (10) is a stationary time series, it satisfies the following n-variate ARMA model (Hanna, 1970):

$$A(B)y_t = C(B)w_t$$

Its transfer form may be written as

$$y_t = \sum_{j=0}^{\infty} G_j B^j w_t$$

The y_t's in the above equation and in equation (11) express the same response sequence of the system, so that the corresponding Green function matrices should be equal. We have

$$\sum_{i=1}^{2n} e^{js_i v} U_i U_i^T (e^{s_i v} - 1)/s_i^2 = \sum_{i=1}^{n\times p} R_i^j E_i$$

Comparing the corresponding parameters on both sides of the above equation, we obtain

$$\left.\begin{aligned} &p = 2 \\ &R_i = e^{s_i v} \\ &E_i = U_i U_i^T (e^{s_i v} - 1)/s_i^2 \end{aligned}\right\} \tag{13}$$

Equation (13) gives the relation between the parameters of a vibration system and the corresponding ARMA model.

4 Identifying the Modal Parameters of a Vibration System

It is easy to see from equation (13) that, having estimated all the R_i and the corresponding coefficient matrices E_i of the ARMA model, we can identify all modal parameters of the vibration system. The identifying process consists of the following three steps.

1 Estimating Eigenvalues of an ARMA Model Based on Correlation Function Data (Zhang and Qiu, 1983). From equation (2), we have

$$\det A(B)y_t = \{adjA(B)\}C(B)w_t \tag{14}$$

Noting that $\det A(B)$ is a polynomial of B, we choose a z_t, any one of the components of y_t, and write down the corresponding equation in equation (14):

$$\det A(B)z_t = Q(B)w_t \tag{15}$$

where $Q(B)$ is the corresponding row vector in
$$\{adj\,A(B)\}C(B).$$
Equation (15) may be regarded as a higher order univariate ARMA model. According to the definition of eigenvalues of an ARMA model, system (15) and (2) should have the same eigenvalues. Therefore, the estimation of eigenvalues of a multivariate ARMA model may be reduced to that of a higher order univariate one.

Assume that the order of $\det A(B)$ is m, and the order of $Q(B)$ is q. Let

$$\det A(B) = 1 - f_1B - f_2B^2 - \ldots - f_mB^m$$

The autocorrelation functions, u_k's, of a univariate ARMA model satisfy the following difference equation:

$$\det A(B)u_k = 0, \quad k > q$$

Since $m > q$, the above equations may be written as

$$\left.\begin{array}{c} u_{m-1}f_1 + u_{m-2}f_2 + \ldots + u_0f_m = u_m \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ u_{N-1}f_1 + u_{N-2}f_2 + \ldots + u_{N-m}f_m = u_N \end{array}\right\} N > m$$

or in the matrix form:

$$P_NF_m = Q_N \tag{16}$$

where

$$F_m = [f_1 \ldots\ldots f_m]^T$$
$$Q_N = [u_m \ldots\ldots u_N]^T$$
$$P_N = \begin{bmatrix} u_{m-1} \ldots\ldots u_0 \\ \ldots\ldots\ldots \\ u_{N-1} \ldots\ldots u_{N-m} \end{bmatrix}$$

The least square solution of equation (16) is

$$F_m = [P_N^TP_N]^{-1}P_N^TQ_N$$

Then we find the zeros of $\det A(B)$ for the ARMA model.

2 Determining the Multivariate AR Model. We know that a multivariate ARMA model has its inverted form:

$$w_t = -\sum_{j=0}^{\infty} L_jB^jy_t, \quad L_0 = -I$$

We can always find a finite order multivariate AR model to approximate this infinite order model with sufficient accuracy. We might as well let its order be J, and write it as

$$A(B)y_t = \Big(I - \sum_{i=1}^{J} A_iB^i\Big)y_t = w_t$$

Postmultiplying y_{t-s}^T to the above equation and taking ensemble averages, we have

$$r_s = \sum_{j=1}^{J} A_jr_{s-j}, \quad s = 1,2, \ldots, J \tag{17}$$

$$r_0 = \sum_{i,j=1}^{J} A_jr_{i-j}A_i^T + D$$

where

$$r_s = E[y_ty_{t-s}^T]$$

Equation (17) is known as Yule-Walker (Y-W) equation. Let

$$a_{JJ} = [A_1 \quad A_2 \quad \ldots \quad A_J]^T$$
$$v_{JJ} = \begin{bmatrix} r_0 & r_1 & \ldots & r_{J-1} \\ r_1^T & r_0 & \ldots & r_{J-2} \\ & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ r_{J-1}^T & r_{J-2}^T & \ldots & r_0 \end{bmatrix}$$
$$v_J = [r_1 \quad r_2 \quad \ldots \quad r_J]^T$$

Equation (17) may be written in the matrix form:

$$v_J = v_{JJ}a_{JJ} \tag{17a}$$

Therefore, we may obtain the Y-W estimation for the multivariate ARMA model. We shall take

$$\hat{r}_s = \frac{1}{N}\sum_{t=s+1}^{N} y_ty_{t-s}^T$$

for the estimation of r_s. One of the merits is that this estimation has the property of positive-definiteness, which is necessary for determining the AR model. Since $\hat{v}_{JJ}$ is positive-definite, its inverse, $\hat{v}_{JJ}^{-1}$, does exist. From equation (17a),we can obtain a unique $\hat{a}_{JJ}$:

$$\hat{a}_{JJ} = \hat{v}_{JJ}^{-1}\,\hat{v}_J$$

In practice, it is more convenient to use the well-known recursive algorithm for Y-W estimation:

$$\hat{a}_{11} = \hat{r}_1\hat{r}_0^{-1}$$

$$\hat{a}_{p+1,p+1} = (\hat{r}_{p+1} - \sum_{j=1}^{p} \hat{a}_{pj}\hat{r}_{p+1-j})(\hat{r}_0 - \sum_{j=1}^{p} \hat{c}_{pj}\hat{r}_j)^{-1}$$

$$\hat{a}_{p+1,j} = \hat{a}_{pj} - \hat{a}_{p+1,p+1}\hat{c}_{p,p+1-j}$$
$$\hat{c}_{11} = \hat{r}_1^T\hat{r}_0^{-1}$$

$$\hat{c}_{p+1,p+1} = (\hat{r}_{p+1} - \sum_{j=1}^{p} \hat{a}_{pj}\hat{r}_{p+1-j})^T(\hat{r}_0 - \sum_{j=1}^{p} \hat{a}_{pj}\hat{r}_j^T)^{-T}$$

$$\hat{c}_{p+1,j} = \hat{c}_{pj} - \hat{c}_{p+1,p+1}\hat{a}_{p,p+1-j}$$
$$j = 1,2, \ldots\ldots, p$$

3 Estimating Modal Parameters of a Vibration System. Having found the eigenvalues of the ARMA model, R_i^{-1}, written as

$$R_i^{-1} = a_i + jb_i$$

and letting the eigenvalues of the vibration system be

$$s_i = \mathrm{Re}(s_i) + j\mathrm{Im}(s_i)$$

we have

$$\mathrm{Re}(s_i) = -(1/v)\ln\sqrt{a_i^2 + b_i^2}$$
$$\mathrm{Im}(s_i) = -(1/v)\tan^{-1}(b_i/a_i)$$

Then the natural frequencies and modal damping ratios may be obtained as

$$\omega_i^2 = \mathrm{Re}^2(s_i) + \mathrm{Im}^2(s_i), \quad \zeta_i = -\mathrm{Re}(s_i)/\omega_i$$

Having obtained the approximate estimation of inverse functions of the stationary random sequence, we find by equation (9) the approximate estimation of the Green functions:

$$G_0, G_1, \ldots\ldots, G_m$$

Table 1 Results of digital simulations ($U_{11} = U_{12} = 1.00000$)

System	Parameter \ Method	Single-point excitation		Theoretical	Multi-point excitation	
		Accurate correlation	Pseudo-random		Accurate correlation	Pseudo-random
1	s_1	$-1.56010 + j21.5777$	$-1.53904 + j21.5069$	$-1.56010 + j21.5777$	$-1.56010 + j21.5777$	$-1.62435 + j21.5535$
	ζ_1	0.07211	0.07138	0.07211	0.07211	0.07515
	ω_1	21.63403	21.56190	21.63399	21.63403	21.62462
	U_{21}	1.87981	$1.92788 + j0.00639$	1.87980	1.87980	$1.83692 + j0.01058$
	s_2	$-4.27323 + j35.5487$	$-4.54904 + j35.2260$	$-4.27324 + j35.5487$	$-4.27323 + j35.5487$	$-4.64640 + j35.2484$
	ζ_2	0.11935	0.12808	0.11935	0.11935	0.13069
	ω_2	35.80462	35.51851	35.80462	35.80462	35.55332
	U_{22}	-3.54648	$-3.72934 - j0.02734$	-3.54647	-3.54648	$-3.55660 - j0.02359$
2	s_1	$-0.23374 + j21.9363$	$-0.25055 + j21.9321$	$-0.23374 + j21.9363$	$-0.23374 + j21.9363$	$-0.22154 + j21.9687$
	ζ_1	0.01066	0.01142	0.01066	0.01066	0.01008
	ω_1	21.93750	21.93353	21.93750	21.93750	21.96982
	U_{21}	$1.99008 - j0.10852$	$2.01963 - j0.17497$	$1.99008 - j0.10852$	$1.99008 - j0.10852$	$2.01550 - j0.10784$
	s_2	$-1.80793 + j35.2813$	$-2.24729 + j35.3364$	$-1.80793 + j35.2813$	$-1.80793 + j35.2813$	$-2.01896 + j35.2737$
	ζ_2	0.05117	0.06347	0.05117	0.05117	0.05714
	ω_2	35.32760	35.40779	35.32760	35.32760	35.33143
	U_{22}	$-3.34316 - j0.29415$	$-3.31824 - j0.37061$	$-3.34316 - j0.29415$	$-3.34316 - j0.29415$	$-3.33370 - j0.36933$

with $m > (2n-1)$. Letting the sth column vector of G_j be $G_j^{(s)}$, we define a matrix $g^{(s)}$ as follows:

$$g^{(s)} = [G_0^{(s)} \quad G_1^{(s)} \; . \; . \; . \; G_m^{(s)}]$$

Letting

$$e^{(s)} = [E_1^{(S)} \quad E_2^{(s)} \; . \; . \; . \; E_2^{(S)}]$$

$$r = \begin{bmatrix} 1 & R_1 & R_1^2 & \ldots & R_1^m \\ & \ldots & \ldots & \ldots & \\ 1 & R_{2n} & R_{2n}^2 & \ldots & R_{2n}^m \end{bmatrix}$$

from equation (8) we have

$$g^{(s)} = e^{(s)} r \tag{18}$$

Therefore, the least square estimation of $e^{(s)}$ may be obtained as

$$\hat{e}^{(s)} = \hat{g}^{(s)} \hat{r}^T [\hat{r}\hat{r}^T]^{-1}$$

Moreover, from equation (13) we have

$$\begin{aligned} e^{(s)} &= [U_1 U_{s1} (e^{s_1 v} - 1)/s_1^2 \; . \; . \; . \; . \; . \; U_{2n} U_{s,2n} (e^{s_{2n} v} - 1)/s_{2n}^2] \\ &= [\tilde{u}_1 \; . \; . \; . \; . \; . \; . \; . \; . \; \tilde{u}_{2n}] \end{aligned}$$

Up to an aribitrary multiplier, $\tilde{u}_i$ is the complex eigenvector associated with s_i.

5 Digital Simulations

The following digital simulations have been made to verify the effectiveness of the above method for identifying modal parameters. Two illustrative systems are used in the simulations. Their differential equations of motion may be written as

$$m\ddot{x} + c\dot{x} + kx = w(t) = [w_1(t) \quad w_2(t)]^T$$

where for system 1,

$$m = \begin{bmatrix} 100 & 0 \\ 0 & 15 \end{bmatrix}, c = \begin{bmatrix} 500 & -100 \\ -100 & 100 \end{bmatrix}, k = 150c$$

and for system 2,

$$m = \begin{bmatrix} 100 & 0 \\ 0 & 15 \end{bmatrix}, c = \begin{bmatrix} 75 & -50 \\ -50 & 50 \end{bmatrix}$$

$$k = \begin{bmatrix} 76915 & -14415 \\ -14415 & 14415 \end{bmatrix}$$

Although the two systems have similar natural frequencies, in system 1 damping is proportional, or classical, so that system 1 has classical normal modes, i.e., its eigenvectors are real. While in system 2 damping is non-proportional, so that system 2 has non-classical modeshapes, i.e., its eigenvectors are all complex. The excitation, $w(t)$, is treated either as zero-meaned ideal white noise in case 1, or as zero-meaned pseudo random white noise in case 2. In each case, both single-point excitation, i.e., $w_1 \neq 0$ and $w_2 = 0$, and multi-point excitation, i.e., $w_1 \neq 0$ and $w_2 \neq 0$, are used.

Case 1 Ideal White Noise Excitation. In this case, ideal white noise excitation is used in order to examine the identification program. Accurate correlation function matrices of stationary random responses are obtained by modal analysis method (Fang and Wang, 1985). Then, multivariate AR models are determined based on the accurate time series data. Finally, modal parameters are obtained. The results are satisfactory and they are listed in Table 1. This simulation is actually a deterministic one, because the correlation function matrix obtained for response to ideal white noise excitation is an auccurate one.

Case 2 Pseudo Random White Noise Excitation. To verify the effectiveness of our approach in practice, the pseudo random white noise is used to carry out the random simulations. The excitation we used is a pseudo random sequence with uniform spectrum, namely a time series, which is the IFFT of a uniform spectrum with random phases. The $w_1(t)$ and $w_2(t)$ are generated independently in the experiments. The responses are obtained by numerical integration. In order to ensure the numerical integration has sufficient accuracy, we resort to the conventional modal analysis method. However, the responses are random due to the random nature of the excitation. In every trial we use different random numbers to generate different time-histories of pseudo random noise. Furthermore, in every trial we discard the first few hundred of sampled response data to ensure a stationary state (Caughey, 1963), and then use the succeeding 1000–1500 sampled data to estimate the response correlation function matrix. The sampling interval is taken as 0.02 s for system 1, and 0.04 s for system 2. The results shown in Table 1 are based on the average outcome of 20 trials. The experiments reveal that both the eigenvalues and the eigenvectors can be estimated within reasonable accuracy, no matter whether the system is single-point excited or multi-point excited.

All the above experiments were conducted on the SIEMENS 7760 computer with a double precision algorithm. For brevity the results listed in Table 1 are only shown to the fifth decimal point.

6 Conclusions

A new approach for identifying modal parameters is developed. The method has the following features:

1. It is based only on response data for white noise excitation. Thus it is applicable even when precise excitation data are unavailable, but the excitation is approximately white in nature, e.g., in low-level ambient vibration tests of structures.
2. Since only a multivariate AR model for response data is required to determine the Green function matrix, it saves computational effort.
3. Not only eigenvalues, but also eigenvectors can be identified with reasonable accuracy.

Digital simulations support this time domain identification method.

Acknowledgment

The authors are grateful to Professor Ji Wen-mei (NPU) and Professor E. H. Dowell (Duke) for their valuable comments.

References

Berman, A., 1979, "Parameter Identification Techniques for Vibrating Structures," *The Shock and Vibration Digest*, Vol. 11, No. 1, pp. 13–16.

Caughey, T. K., 1963, "Nonstationary Random Inputs and Responses," *Random Vibrations*, S. H. Crandall, ed., Vol. 2, MIT Press.

Fang, T., and Wang, Z. N., 1985, "Complex Modal Analysis of Random Vibrations," to appear in *AIAA Journal*.

Gersch, W., 1975, "Parameter Identification: Stochastic Process Techniques," *The Shock and Vibration Digest*, Vol. 7, No. 11, pp. 71–86.

Gersch, W., and Liu, R. S-Z., 1976, "Time Series Methods for the Synthesis of Random Vibration Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, pp. 159–165.

Hanna, E. J., 1970, *Multiple Time Series*, New York.

Ibrahim, S. R., and Miluleik, E. C., 1977, "A Method for the Direct Identification of Vibration Parameters from Free Response," *Shock and Vibration Bulletin*, Bulletin 47, Part 4, pp. 183–198.

Ibrahim, S. R., 1977, "Random Decrement Technique for Modal Identification of Structures," *Journal of Spacecraft and Rockets*, Vol. 14, pp. 696–700.

Pandit, S. M., and Wu, S. M., 1983, *Time Series and System Analysis with Applications*, Wiley, New York.

Priestley, M. B., 1981, *Spectral Analysis and Time Series*, Vol. 2, Academic Press.

Wang, Z. N., 1984, "Modal Analysis of Random Vibrations and Time Domain Identification of Modal Parameters," Doctoral Thesis, Northwestern Polytechnical University, Xian, Shaanxi, China.

Wu, S. M., 1977, "Dynamic Data System: A New Modeling Approach," ASME *Journal of Engineering for Industry*, Vol. 99, pp. 708–714.

Zhang, J. H., and Qiu, Y., 1983, "Applications of Time Series Analysis in Vibrations," (in Chinese), *Journal of Vibration and Shock*, Vol. 2, No. 1.

ERRATA

Errata on "Beams in Receding/Advancing Contact: Dunders Problems" by T. P. Pawlak, N. J. Salamon, and F. F. Mahmoud, published in the December 1985 issue of ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 933–936.

On p. 934, line 2 of Table 1 should read as follows:

Moment of inertia $= 1.041 \times 10^{-7}$ m^4

On p. 935, line 21 in the Results section should read as follows:

" . . . where $\bar{D}$ is the center of deflection for $P = 2.224 \times 10^5$ N. In comparison . . . "

F. R. Vigneron
Research Scientist,
Communications Research Centre,
Ottawa, Canada, K2H 8S2

A Natural Modes Model and Modal Identities for Damped Linear Structures

A modal model is derived for a passive elastic structure with linear viscous damping, from a first-order state variable arrangement of the physical parameters model. The state variable form of the model is composed using the equations $\mathbf{K}\dot{\mathbf{q}} - \mathbf{K}\dot{\mathbf{q}} = 0$ *and* $\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}$*. An attribute of the particular formulation is that it facilitates a straightforward derivation of mass-properties-related modal identities for the associated damped natural modes. Transfer functions and normalizations used in experimental modal parameter estimation are also given special attention.*

1.0 Introduction

For currently developing modal test and analysis technology that is based on curve-fit type of parameter estimation techniques, such as the complex exponentials method, a modal model corresponding to a passive linear elastic structure with general linear viscous damping is of value and interest, and in use to some extent. This type of model is also referred to as a "damped natural modes model."

A derivation of the basic modal model for passive nongyroscopic systems from a first-order state variable arrangement of the physical parameters model, in the spirit of the classical normal modes analysis, was published by Foss (1958) and is well known and extensively used (Meirovitch, 1967; Ewins, 1984). An alternate modal derivation that involves Laplace Transformation of the physical parameters model and analysis of the resulting second-order algebraic matrix equation is also available (Richardson, 1974). Derivations using state variable formulations different than Foss's have also been given (e.g., Béliveau, 1977, Brandon, 1984). Analyses of modal properties applicable to active gyroscopic systems are given by Nelson and Glasgow (1979) and Fawzy (1977). These works are complementary and contribute a visibility into the structure of the system that is needed, particularly for efficient parameter estimation.

Foss's (1956) state variable arrangement of the physical parameters model, composed using $M\dot{\mathbf{q}} - M\dot{\mathbf{q}} = 0$ and $M\ddot{\mathbf{q}} + C\dot{\mathbf{q}} + K\mathbf{q} = \mathbf{f}$, has parameter matrices that are not positive definite. Consequently certain mass-related modal identities cannot be established conveniently from the formulation. The modal identities are the damped natural mode equivalents of those recently documented by Hughes (1980) for the undamped case. The identities have significance to future research in analytical and experimental techniques of assessing mode set completeness and modal truncation procedures. In a formulation put forth for a more general category of modal analysis by Meirovitch and Baruh (1981) and Vigneron (1981), $K\dot{\mathbf{q}} - K\dot{\mathbf{q}} = 0$ is used to compose the state variable form and this results in the presence of a positive definite parameter matrix. As a result it turns out to be evident that certain modal vectors constitute a basis in a Euclidean inner product space, and further that Bessel's equations can be conveniently obtained and used to establish the modal identities.

This paper first derives the modal model from the latter state variable arrangement, with the intent of adding further visibility into the structure of the passive linear damped case and clarifying certain normalization factors and other relations used in parameter estimation. Then the appropriate form of Bessel's and Parseval's equations are established, and the modal identities obtained.

2.0 Model in Terms of Physical Variables

To obtain modal identities in the form to follow in Chapter 4, a physically-based definition of the structure and reference coordinate system are needed. One appropriate for this purpose, and for experimental modal parameter estimation, is outlined in this Chapter.

The structure, depicted schematically in Fig. 1, is defined by N points, relative to a coordinate system $0xyz$. In parameter identification, where a finite number of points are instrumented, N is finite. A mass, m^i, is associated with each point. The coordinate system, $0xyz$, is considered to be chosen such that rigid body translations and rotations between it and the structure are not possible due to physical restraints, or have been mathematically eliminated. Define the deformation matrices, of order $N \times 1$, as

$$U = \begin{bmatrix} u^1 \\ u^2 \\ u^3 \\ \cdot \\ \cdot \\ \cdot \\ u^N \end{bmatrix} \quad V = \begin{bmatrix} v^1 \\ v^2 \\ v^3 \\ \cdot \\ \cdot \\ \cdot \\ v^N \end{bmatrix} \quad W = \begin{bmatrix} w^1 \\ w^2 \\ w^3 \\ \cdot \\ \cdot \\ \cdot \\ w^N \end{bmatrix}, \qquad (2\text{-}1)$$

and the corresponding position matrices of order $N \times 1$ 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 two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Mechanics Division, June 11, 1985; final revision, September 13, 1985.

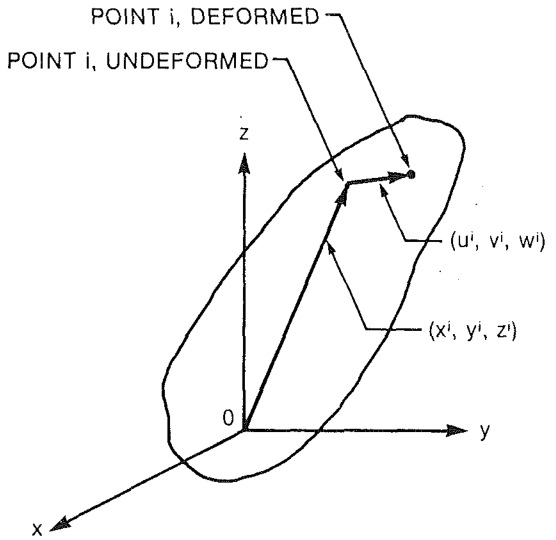


Fig. 1 Coordinates of the structure

$$X=\begin{bmatrix} x^1\\ x^2\\ x^3\\ \cdot\\ \cdot\\ \cdot\\ x^N \end{bmatrix}\quad Y=\begin{bmatrix} y^1\\ y^2\\ y^3\\ \cdot\\ \cdot\\ \cdot\\ y^N \end{bmatrix}\quad Z=\begin{bmatrix} z^1\\ z^2\\ z^3\\ \cdot\\ \cdot\\ \cdot\\ z^N \end{bmatrix}. \tag{2-2}$$

The model for this structure has the form

$$M\ddot{\mathbf{q}}+C\dot{\mathbf{q}}+K\mathbf{q}=\mathbf{f}, \tag{2-3}$$

The order of equation (2-3) is $3N$, and is further denoted herein by n; C and K and M are each of order $n \times n$. The $3N \times 1$ column matrices, $\mathbf{q}$ and $\mathbf{f}$, are

$$\mathbf{q}=\begin{bmatrix} U\\ V\\ W \end{bmatrix},\quad \mathbf{f}=\begin{bmatrix} f_x\\ f_y\\ f_z \end{bmatrix}. \tag{2-4}$$

The structure is considered to be passive, non-gryoscopic and with general linear viscous damping. Therefore, M, C, and K are symmetric and positive definite.

For the above choice of deformation variables and coordinates, M turns out to be diagonal and of the form

$$M=\begin{bmatrix} M^N & 0 & 0\\ 0 & M^N & 0\\ 0 & 0 & M^N \end{bmatrix},\quad M^N=\begin{bmatrix} m^1 & 0 & \cdot & \cdot\\ 0 & m^2 & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & m^N \end{bmatrix}, \tag{2-5}$$

where the dimension of M^N is $N \times N$.

When N is taken to infinity the model is exactly synonymous with the continuum representation of linear elasticity theory. The mathematical formulation to follow in Chapters 3 and part of 4 also applies for equations having the form of equation (2-3) with M non-diagonal. To obtain the form of the identities to follow in Chapter 4, the physical coordinates (and thus the diagonal M of this particular physical model) are employed.

3.0 Transformation to Damped Natural Modal Variables

Structural Model in State Variable Form. Equation (2-3) may be arranged in the first order state variable form,

$$A\dot{\mathbf{Q}}+B\mathbf{Q}=\mathrm{F} \tag{3-1}$$

$$\mathbf{Q}=\begin{bmatrix} \dot{\mathbf{q}}\\ \mathbf{q} \end{bmatrix},\quad \mathbf{F}=\begin{bmatrix} \mathbf{f}\\ \mathbf{0} \end{bmatrix},\quad A=\begin{bmatrix} M & 0\\ 0 & K \end{bmatrix},\quad B=\begin{bmatrix} C & K\\ -K & 0 \end{bmatrix}$$

where $\mathbf{Q}$ and $\mathbf{F}$ are of order $2n \times 1$; A is of order $2n \times 2n$ and rank $2n$, and is positive definite and symmetric; B is order $2n \times 2n$, and is the sum of a symmetric part involving C and a skew symmetric part involving K. As noted in the Introduction, the above state variable form is different than the usual one of Foss (1958).

Eigenproblem Analysis. The eigenproblem corresponding to equation (3-1) is

$$(\lambda_k A+B)\boldsymbol{\Upsilon}_k=0. \tag{3-2}$$

The eigenvalues λ_k, are solutions of

$$\det(\lambda_k A+B)=0. \tag{3-3}$$

Since A and B are of dimension $2n$, equation (3-3) is of degree $2n$. The equation yields $2n$ eigenvalues which are real or complex. Since the rank of B is also $2n$, the eigenvalues are nonzero. The complex eigenvalues occur in complex conjugate pairs

$$\lambda_k=-\sigma_k+i\nu_k \quad \lambda_k^*=-\sigma_k-i\nu_k. \tag{3-4}$$

The eigenvalues are assumed distinct. The λ's that are complex can be converted to the conventional natural modal frequencies and modal damping ratios, ω and ζ by the formula

$$\zeta_k=\sigma_k/(\nu_k^2+\sigma_k^2)^{1/2};\quad \omega_k^2=\nu_k^2+\sigma_k^2. \tag{3-5a}$$

The converse is

$$\sigma_k=\zeta_k\omega_k,\quad \nu_k^2=\omega_k^2(1-\zeta_k^2). \tag{3-5b}$$

A matrix column, $\boldsymbol{\Upsilon}_k$, is calculated (and determinable to within a complex scalar constant) by solving equation (3-3) for a particular λ_k. Also a matrix row, $\boldsymbol{\Gamma}_\tau^T$, can be calculated for a particular λ_τ from

$$\boldsymbol{\Gamma}_\tau^T(\lambda_\tau A+B)=0. \tag{3-6}$$

Taking the transpose of equation (3-6) and recognizing that $A = A^T$ yields

$$(\lambda_\tau A+B^T)\boldsymbol{\Gamma}_\tau=0. \tag{3-7}$$

$\boldsymbol{\Upsilon}_k$ and $\boldsymbol{\Gamma}_k$ are right and left eigenvectors of dimension $2n \times 1$, and are complex in general.

Premultiply equation (3-2) by $\boldsymbol{\Gamma}_\tau^T$:

$$\lambda_k\boldsymbol{\Gamma}_\tau^T A\boldsymbol{\Upsilon}_k+\boldsymbol{\Gamma}_\tau^T B\boldsymbol{\Upsilon}_k=0. \tag{3-8}$$

Post-multiply equation (3-6) by $\boldsymbol{\Upsilon}_k$:

$$\lambda_\tau\boldsymbol{\Gamma}_\tau^T A\boldsymbol{\Upsilon}_k+\boldsymbol{\Gamma}_\tau^T B\boldsymbol{\Upsilon}_k=0. \tag{3-9}$$

Subtract equation (3-9) from equation (3-8) to obtain,

$$(\lambda_k-\lambda_\tau)\boldsymbol{\Gamma}_\tau^T A\boldsymbol{\Upsilon}_k=0. \tag{3-10}$$

Therefore, for two non-equal eigenvalues λ_τ and λ_k,

$$\boldsymbol{\Gamma}_\tau^T A\boldsymbol{\Upsilon}_k=0;\quad \boldsymbol{\Gamma}_\tau^T B\boldsymbol{\Upsilon}_k=0 \quad \tau\neq k. \tag{3-11}$$

It follows from equation (3-8), and from equation (3-7) after multiplication $\boldsymbol{\Upsilon}_k^T$, that

$$\lambda_k=-\frac{\boldsymbol{\Gamma}_k^T B\boldsymbol{\Upsilon}_k}{\boldsymbol{\Gamma}_k^T A\boldsymbol{\Upsilon}_k}=-\frac{\boldsymbol{\Upsilon}_k^T B^T\boldsymbol{\Gamma}_k}{\boldsymbol{\Upsilon}_k^T A\boldsymbol{\Gamma}_k}. \tag{3-12}$$

λ_k and λ_k^* are two distinct eigenvalues, and consequently $\boldsymbol{\Gamma}_k^{*T}A\boldsymbol{\Upsilon}_k$ and $\boldsymbol{\Gamma}_k^{*T}B\boldsymbol{\Upsilon}_k$ are equal to zero.

To demonstrate the structure of $\boldsymbol{\Upsilon}_k$, consider the upper and lower internal columns, $\boldsymbol{\Upsilon}_k^u$ and $\boldsymbol{\Upsilon}_k^l$, each of order $n \times 1$:

$$\boldsymbol{\Upsilon}_k=\begin{bmatrix} \boldsymbol{\Upsilon}_k^u\\ \boldsymbol{\Upsilon}_k^l \end{bmatrix}. \tag{3-13}$$

Substitute equation (3-13) and the parameter values for A and B into equation (3-2):

$$\begin{pmatrix} \lambda_k M & 0\\ 0 & \lambda_k K \end{pmatrix}\begin{pmatrix} \boldsymbol{\Upsilon}_k^u\\ \boldsymbol{\Upsilon}_k^l \end{pmatrix}+\begin{pmatrix} C & K\\ -K & 0 \end{pmatrix}\begin{pmatrix} \boldsymbol{\Upsilon}_k^u\\ \boldsymbol{\Upsilon}_k^l \end{pmatrix}=0. \tag{3-14}$$

Equation (3-14) reduces to

$$\lambda_k M\boldsymbol{\Upsilon}_k^u+C\boldsymbol{\Upsilon}_k^u+K\boldsymbol{\Upsilon}_k^l=0 \tag{3-15a}$$

$$\lambda_k K\boldsymbol{\Upsilon}_k^l - K\boldsymbol{\Upsilon}_k^u = 0. \tag{3-15b}$$

From the latter equation, $\boldsymbol{\Upsilon}_k^u = \lambda_k \boldsymbol{\Upsilon}_k^l$. Then equation (3-13) becomes

$$\boldsymbol{\Upsilon}_k = \begin{bmatrix} \lambda_k \boldsymbol{\Phi}_k \\ \boldsymbol{\Phi}_k \end{bmatrix}. \tag{3-16}$$

$\boldsymbol{\Phi}_k$ is complex in general, and of dimension $n \times 1$. Equations (3-15) can be further arranged to obtain

$$(\lambda_k^2 M + \lambda_k C + K)\boldsymbol{\Phi}_k = 0. \tag{3-17}$$

A similar parallel development where $\boldsymbol{\Gamma}_k$ is partitioned into upper and lower parts and equation (3-7) is decomposed in a similar manner yields $\boldsymbol{\Gamma}_k^u = -\lambda_k \boldsymbol{\Gamma}_k^l$ and $(\lambda_k^2 M + \lambda_k C + K)\boldsymbol{\Gamma}_k^l = 0$. Thus $\boldsymbol{\Gamma}_k^l$ equals $\boldsymbol{\Phi}_k$, and

$$\boldsymbol{\Gamma}_k = \begin{bmatrix} -\lambda_k \boldsymbol{\Phi}_k \\ \boldsymbol{\Phi}_k \end{bmatrix}. \tag{3-18}$$

Vector Space Corresponding to Damped Natural Modes. The quantities, $\boldsymbol{\Upsilon}_k$ and $\boldsymbol{\Upsilon}_k^*$, $k = 1$ to n, are a set of $2n$ distinct column matrices (each of dimension $2n \times 1$). The set can be shown to be linearly independent, and each eigenvector can be shown to be unique in the sense that there is only one $\boldsymbol{\Upsilon}_k$ for each λ_k by proofs similar to one given in Wilkinson (1965). Thus the quantities $(\boldsymbol{\Upsilon}_k + \boldsymbol{\Upsilon}_k^*)$ and $i(\boldsymbol{\Upsilon}_k - \boldsymbol{\Upsilon}_k^*)$, $k = 1$ to n, are also $2n$ linearly independent column matrices and are real-valued, and define a vector basis of a real vector space of dimension $2n$. For any vector $\mathbf{G}$ (i.e., real-valued column matrix of dimension $2n \times 1$) contained in the space, $\mathbf{G}^T A\mathbf{G} \geq 0$ since A is positive definite. Thus $\mathbf{G}_1^T A\mathbf{G}_2$ defines a real inner product for two arbitrary vectors, $\mathbf{G}_1$ and $\mathbf{G}_2$, contained in the space. Thus the vector space is a real inner product space (a Euclidean Space). The corresponding dual basis consists of $(\boldsymbol{\Gamma}_k + \boldsymbol{\Gamma}_k^*)$ and $i(\boldsymbol{\Gamma}_k - \boldsymbol{\Gamma}_k^*)$, $k = 1$ to n. A real-valued vector, $\mathbf{G}$, may be represented in terms of the basis vectors by

$$\mathbf{G} = \sum_{k=1}^{n} a_k(\boldsymbol{\Upsilon}_k + \boldsymbol{\Upsilon}_k^*) + b_k i(\boldsymbol{\Upsilon}_k - \boldsymbol{\Upsilon}_k^*), \tag{3-19}$$

where a_k and b_k are real-valued scalars. The above expression rearranges to the form

$$\mathbf{G} = \sum_{k=1}^{n} (\alpha_k \boldsymbol{\Upsilon}_k + \alpha_k^* \boldsymbol{\Upsilon}_k^*) \tag{3-20}$$

where $\alpha_k = a_k + ib_k$. Thus one may regard the basis vectors of real vector space as $\boldsymbol{\Upsilon}_k$ and $\boldsymbol{\Upsilon}_k^*$, if at the same time appropriate pairing of complex and complex conjugate quantities is maintained to ensure that the total expression is real-valued. It proves convenient to work directly with the complex $\boldsymbol{\Upsilon}$'s and $\boldsymbol{\Gamma}$'s and the representation in the form of equation (3-20) as oppposed to the real-valued basis vectors and the form of equation (3-19).

System Model in Terms of Complex Modal Variables. The real-valued column matrix, $\mathbf{Q}(t)$, of equation (3-1), can thus be represented as

$$\mathbf{Q}(t) = \sum_{k=1}^{n} [\boldsymbol{\Upsilon}_k \rho_k(t) + \boldsymbol{\Upsilon}_k^* \rho_k^*(t)], \tag{3-21}$$

where $\rho_k(t)$ is a complex-valued scalar variable

$$\rho_k(t) = \xi_k(t) + i\eta_k(t). \tag{3-22}$$

To transform the state variable differential equation, equation (3-1), to be equivalent modal variable differential equation, substitute equation (3-21) into equation (3-1), premultiply by $\boldsymbol{\Gamma}_\tau^T$, and use the orthogonality properties of equations (3-11) and (3-12), to obtain

$$\dot{\rho}_k - \lambda_k \rho_k = \frac{\boldsymbol{\Gamma}_k^T \mathbf{F}(t)}{\boldsymbol{\Gamma}_k^T A \boldsymbol{\Upsilon}_k}. \tag{3-23}$$

The denominator, $\boldsymbol{\Gamma}_k^T A \boldsymbol{\Upsilon}_k$, can be reduced by substituting into it A of equation (3-1) and $\boldsymbol{\Upsilon}_k$ and $\boldsymbol{\Gamma}_k$ of equations (3-16) and (3-18), to obtain

$$-\lambda_k^2 \boldsymbol{\Phi}_k^T M \boldsymbol{\Phi}_k + \boldsymbol{\Phi}_k^T K \boldsymbol{\Phi}_k. \tag{3-24}$$

Multiplication of equation (3-17) by $\boldsymbol{\Phi}_k^T$ and combination with the above gives the final result,

$$\boldsymbol{\Gamma}_k^T A \boldsymbol{\Upsilon}_k = -\lambda_k \boldsymbol{\Phi}_k^T (2\lambda_k M + C)\boldsymbol{\Phi}_k. \tag{3-25}$$

Likewise, the numerator of equation (3-23) becomes $\boldsymbol{\Gamma}_k^T \mathbf{F} = -\lambda_k \boldsymbol{\Phi}_k^T \mathbf{f}$. Substitution of these results into equation (3-23) gives

$$\dot{\rho}_k(t) - \lambda_k \rho_k(t) = \frac{\boldsymbol{\Phi}_k^T \mathbf{f}(t)}{\boldsymbol{\Phi}_k^T (2\lambda_k M + C)\boldsymbol{\Phi}_k}. \tag{3-26a}$$

A similar procedure with $\boldsymbol{\Gamma}_\tau^{*T}$ as the premultiplier leads to

$$\dot{\rho}_k^*(t) - \lambda_k^* \rho_k^*(t) = \frac{\boldsymbol{\Phi}_k^{*T} \mathbf{f}(t)}{\boldsymbol{\Phi}_k^{*T} (2\lambda_k^* M + C)\boldsymbol{\Phi}_k^*}. \tag{3-26b}$$

Equation (3-26*a*) or (3-26*b*) is the differential equation of the system in terms of complex modal variables.

Transfer Matrix and Residues. The transformation between $\mathbf{q}$ and ρ may be deduced from equations (3-21), (3-1), and (3-16) to be

$$\mathbf{q}(t) = \sum_{k=1}^{n} \{\boldsymbol{\Phi}_k \rho_k(t) + \boldsymbol{\Phi}_k^* \rho_k^*(t)\}. \tag{3-27}$$

Equations (3-26) and (3-27) are next transformed by the Laplace Transform. In this context the two-sided transform of a complex variable is implied with $\mathbf{q}(t)$ taken to be zero at $t_0 = -\infty$. Then $\rho(-\infty)$ is zero. The transformed modal equation is

$$\bar{\rho}_k(s) = \frac{\boldsymbol{\Phi}_k^T \bar{\mathbf{f}}(s)}{\boldsymbol{\Phi}_k^T (2\lambda_k M + C)\boldsymbol{\Phi}_k} \cdot \frac{1}{(s - \lambda_k)}. \tag{3-28}$$

Substitution of equation (3-28) into a transformed version of equation (3-27) yields

$$\bar{\mathbf{q}}(s) = \sum_{k=1}^{n} \left\{ \frac{\boldsymbol{\Phi}_k \boldsymbol{\Phi}_k^T}{\boldsymbol{\Phi}_k^T (2\lambda_k M + C)\boldsymbol{\Phi}_k} \cdot \frac{1}{(s - \lambda_k)} \right.$$

$$\left. + \frac{\boldsymbol{\Phi}_k^* \boldsymbol{\Phi}_k^{*T}}{\boldsymbol{\Phi}_k^{*T} (2\lambda_k^* M + C)\boldsymbol{\Phi}_k^*} \cdot \frac{1}{(s - \lambda_k^*)} \right] \bar{\mathbf{f}}(s). \tag{3-29}$$

From the above equation follows the definition of the residue matrices, R_k and R_k^*, as

$$R_k = \frac{\boldsymbol{\Phi}_k \boldsymbol{\Phi}_k^T}{\boldsymbol{\Phi}_k^T (2\lambda_k M + C)\boldsymbol{\Phi}_k}\ ; \quad R_k^* = \frac{\boldsymbol{\Phi}_k^* \boldsymbol{\Phi}_k^{*T}}{\boldsymbol{\Phi}_k^{*T} (2\lambda_k^* M + C)\boldsymbol{\Phi}_k^*}\ , \tag{3-30}$$

and the 'transfer matrix', $\bar{H}(s)$, as

$$\bar{H}(s) = \sum_{k=1}^{n} \left[\frac{R_k}{s - \lambda_k} + \frac{R_k^*}{s - \lambda_k^*} \right]. \tag{3-31}$$

$\bar{H}(s)$ can be inverted to the time domain, to yield the unit impulse response function, $H(t)$:

$$H(t) = \sum_{k=1}^{n} (R_k e^{\lambda_k t} + R_k^* e^{\lambda_k^* t}). \tag{3-32}$$

R_k and $\bar{H}(s)$ are complex-valued and of order $n \times n$. $H(t)$ is real-valued and of order $n \times n$.

Normalization Constant. The normalization constant, Q_k, is consistent with experimental modal analysis conventions of Brown (1984) and with Fawzy (1977) when defined as

$$Q_k = \frac{1}{\boldsymbol{\Phi}_k^T (2\lambda_k M + C)\boldsymbol{\Phi}_k}. \tag{3-33}$$

Q_k is a complex scalar. Its numerical value and units depend on numerical value and units of the mode shape, or vice-versa.

With this definition of normalization constant, the residues assume the form,

$$R_k = Q_k \boldsymbol{\Phi}_k \boldsymbol{\Phi}_k^T; \quad R_k^* = Q_k^* \boldsymbol{\Phi}_k^* \boldsymbol{\Phi}_k^{*T}, \tag{3-34}$$

and $\bar{H}(s)$ takes the form

$$\bar{H}(s) = \sum_{k=1}^{n} \left\{ \frac{Q_k \boldsymbol{\Phi}_k \boldsymbol{\Phi}_k^T}{s-\lambda_k} + \frac{Q_k^* \boldsymbol{\Phi}_k^* \boldsymbol{\Phi}_k^{*T}}{s-\lambda_k^*} \right\}. \tag{3-35}$$

Three of several possible ways of assigning the arbitrary constant associated with each mode shape, and hence Q_k, follow.

(a) Choose the numerical scaling for each mode so that the scalar element of the mode shape at the main exciter's driving point in a modal survey test is $1 + i0$. This implies that Q_k equals the scalar value of the residue associated with the driving point. This choice of normalization is compatible with experimental modal analysis conventions (Brown (1984)).

(b) Choose the scaling for each mode so that $\boldsymbol{\Phi}_k^T(2\lambda_k M + C)\boldsymbol{\Phi}_k$ equals $2\lambda_k$. This degenerates to $\boldsymbol{\Phi}_k^T M \boldsymbol{\Phi}_k = 1$ when damping is zero, and is thus consistent with the unity modal mass convention of finite element and experimental modal engineering practice (Brown (1984)). For this choice, $Q_k = 1/2\lambda_k$.

(c) Choose the numerical scaling of each mode so that $\boldsymbol{\Gamma}_k^T A \boldsymbol{\Upsilon}_k$ equals $1 + i0$. This then implies that $Q_k = -\lambda_k$. This choice is the most convenient one for theoretical work because it simplifies algebra a great deal. Unfortunately, this normalization has no counterpart in the classical undamped or proportionally-damped modal theories, and consequently is not compatible with current engineering practice and developed software (e.g., SDRC (1985)).

Procedures for establishing the Q_k and mode shape scale factor from experiment-derived modal data (where M, C, K are not known) are described in Brown (1984) and in Vigneron (1985).

System Model in Terms of Real Modal Variables. The system model can be expressed in terms of real-valued modal variables, $\xi_k(t)$ and $\eta_k(t)$, as an alternate to $\rho(t)$ and $\rho^*(t)$. Substitution of equation (3-22) into (3-27) results in

$$\mathbf{q}(t) = 2 \sum_{k=1}^{n} \{ \mathrm{Re}\boldsymbol{\Phi}_k \bullet \xi_k(t) - \mathrm{Im}\boldsymbol{\Phi}_k \bullet \eta_k(t) \}. \tag{3-36}$$

Substitution of equations (3-22) and (3-4) into equations (3-26*a*) and (3-26*b*), and successive addition and subtraction of the two equations leads to

$$\dot{\xi}_k + \sigma_k \xi_k + \nu_k \eta_k = \mathbf{S}_k^T \mathbf{f} \tag{3-37a}$$

$$\dot{\eta}_k + \sigma_k \eta_k - \nu_k \xi_k = \mathbf{T}_k^T \mathbf{f}, \tag{3-37b}$$

where $\mathbf{S}_k^T$ and $\mathbf{T}_k^T$ are the real and imaginary parts, respectively, of $\boldsymbol{\Phi}_k^T / \{\boldsymbol{\Phi}_k^T(2\lambda_k M + C)\boldsymbol{\Phi}_k\}$. Equations (3-37) are two first-order modal differential equations that are the damped-natural-modes counterpart of the familiar single uncoupled second order modal equation

$$\ddot{\mathbf{p}}_k + 2\zeta_k \omega_k \dot{\mathbf{p}}_k + \omega_k^2 \mathbf{p}_k = \pi^T \mathbf{f}. \tag{3-38}$$

of the proportional-damping theory. Of significance, however, is the fact that equations (3-37) cannot be put into the form of equation (3-38), except for the special cases of proportional and zero damping. Because of this, the physical concepts of "modal mass," "modal damping," and "modal stiffness" are not rigorously-applicable in the damped natural modes theory. The transformation of equations (3-37) from variables (ξ_k, η_k) to (τ_k, p_k) by

$$\tau_k(t) = \frac{\sigma_k}{\nu_k} \xi_k(t) + \eta_k(t) \tag{3-39a}$$

$$p_k(t) = \frac{\sigma_k^2 + \nu_k^2}{\nu_k} \xi_k(t), \tag{3-39b}$$

together with appropriate substitutions and use of equations (3-5) lead to:

$$\ddot{p}_k + 2\zeta_k \omega_k \dot{p}_k + \omega_k^2 p_k = \omega^2{}_k \left(\frac{\zeta_k \mathbf{S}_k^T}{\sqrt{1-\zeta_k^2}} - \mathbf{T}_k^T \right) \mathbf{f} + \mathbf{S}_k^T \dot{\mathbf{f}}. \tag{3-40}$$

The above equation is different in structure from equation (3-38) as noted above, because of the presence of the term in $\dot{\mathbf{f}}$.

Modal mass, modal stiffness and modal damping for the damped natural modes model can be established from the properties $\boldsymbol{\Gamma}_k^{*T} A \boldsymbol{\Upsilon}_k = 0$ and $\boldsymbol{\Gamma}_k^{*T} B \boldsymbol{\Upsilon}_k = 0$. Additional information on the real modal variables formulation is available in Vigneron (1981) and Vigneron (1985).

4.0 Modal Identities

Bessel's and Parseval's Equations. Bessel's and Parseval's equations offer a convenient means of establishing certain modal identities that are useful for validating completeness of mode sets and modal truncation. The form of the equations that appear in standard references is not general enough for the situation at hand. The desired forms are briefly outlined below.

A real-valued arbitrary vector $\mathbf{G}$ of the Euclidean inner product space may be represented in the form

$$\mathbf{G} = \sum_{k=1}^{n} (\alpha_k \boldsymbol{\Upsilon}_k + \alpha_k^* \boldsymbol{\Upsilon}_k^*). \tag{4-1}$$

The Fourier coefficients, α_k, are obtained by multiplying equation (4-1) by $\boldsymbol{\Gamma}_r^T A$ and use of the orthogonality relationships, equations (3-11). Likewise an expression for α_k^* is obtained by multiplication by $\boldsymbol{\Gamma}_r^{*T} A$. The following are obtained.

$$\alpha_k = \frac{\boldsymbol{\Gamma}_k^T A \mathbf{G}}{\boldsymbol{\Gamma}_k^T A \boldsymbol{\Upsilon}_k}; \qquad \alpha_k^* = \frac{\boldsymbol{\Gamma}_k^{*T} A \mathbf{G}}{\boldsymbol{\Gamma}_k^{*T} A \boldsymbol{\Upsilon}_k^*} \tag{4-2}$$

Similarly, $\mathbf{G}$ may be represented in terms of $\boldsymbol{\Gamma}_k$ as

$$\mathbf{G}^T = \sum_{k=1}^{n} (\beta_k \boldsymbol{\Gamma}_k^T + \beta_k^* \boldsymbol{\Gamma}_k^{*T}). \tag{4-3}$$

The corresponding Fourier coefficients, derived by post-multiplication by $A\boldsymbol{\Upsilon}_k$ and $A\boldsymbol{\Upsilon}_k^*$ as above, are

$$\beta_k = \frac{\mathbf{G}^T A \boldsymbol{\Upsilon}_k}{\boldsymbol{\Gamma}_k^T A \boldsymbol{\Upsilon}_k}; \qquad \beta_k^* = \frac{\mathbf{G}^T A \boldsymbol{\Upsilon}_k^*}{\boldsymbol{\Gamma}_k^{*T} A \boldsymbol{\Upsilon}_k^*} \tag{4-4}$$

Consider the inner product,

$$\left\{ \mathbf{G}^T - \sum_{k=1}^{n} (\beta_k \boldsymbol{\Gamma}_k^T + \beta_k^* \boldsymbol{\Gamma}_k^{*T}) \right\} A \left\{ \mathbf{G} - \sum_{k=1}^{n} (\alpha_k \boldsymbol{\Upsilon}_k + \alpha_k^* \boldsymbol{\Upsilon}_k^*) \right\}. \tag{4-5}$$

The expression equals zero if the basis and dual, $\boldsymbol{\Upsilon}_k$ and $\boldsymbol{\Gamma}_k$, are complete. If they are incomplete, due to modal trucation for example, then the inner product is real and greater than or equal to zero, because A is positive definite and the right and left multiplying vectors are equal and greater than zero (if truncation is done, the dependence between the bases $\boldsymbol{\Upsilon}_k$ and $\boldsymbol{\Gamma}_k$ must be taken account of). The following result is achieved by multiplying equation (4-5) out, and simplifying with the orthogonality relations and the relationship $\boldsymbol{\Gamma}_k^T A \boldsymbol{\Upsilon}_k = \boldsymbol{\Upsilon}_k^T A \boldsymbol{\Gamma}_k$:

$$\mathbf{G}^T A \mathbf{G} \geq \sum_{k=1}^{n} \{ (\boldsymbol{\Gamma}_k^T A \boldsymbol{\Upsilon}_k) \alpha_k \beta_k + (\boldsymbol{\Gamma}_k^{*T} A \boldsymbol{\Upsilon}_k^*) \alpha_k^* \beta_k^* \} \tag{4-6}$$

The above is a form of Bessel's inequality. If the bases are complete, the equality holds and the relation is referred to as Parseval's Equation.

A second different arbitrary function $\bar{\mathbf{G}}$, may be further defined by

$$\bar{\mathbf{G}} = \sum_{k=1}^{n} (\bar{\alpha}_k \boldsymbol{\Upsilon}_k + \bar{\alpha}_k^* \boldsymbol{\Upsilon}_k^*), \tag{4-7}$$

$$\bar{\mathbf{G}}^T = \sum_{k=1}^{n} (\bar{\beta}_k \mathbf{\Gamma}_k^T + \bar{\beta}_k^* \mathbf{\Gamma}_k^{*T}), \tag{4-8}$$

where $\bar{\alpha}_k$, $\bar{\alpha}_k^*$, $\bar{\beta}_k$ and $\bar{\beta}_k^*$ are as per equation (4-4) with $\mathbf{G}$ replaced by $\bar{\mathbf{G}}$. It may further be shown that

$$\bar{\mathbf{G}}^T A \mathbf{G} \geq \frac{1}{2} \sum_{k=1}^{n} \{\mathbf{\Gamma}_k^T A \mathbf{\Upsilon}_k (\bar{\alpha}_k \beta_k + \alpha_k \bar{\beta}_k) + \mathbf{\Gamma}_k^{*T} A \mathbf{\Upsilon}_k^* (\bar{\alpha}_k^* \beta_k^* + \alpha_k^* \bar{\beta}_k^*)\}. \tag{4-9}$$

Equation (4-9) is referred to as the general form of Bessel's Inequality.

Identitites Involving Modal Linear Momentum. $\mathbf{\Phi}_k$ may be expressed in terms of components:

$$\mathbf{\Phi}_k^T = [\theta_k^T, \phi_k^T, \psi_k^T], \tag{4-10}$$

where θ_k, ϕ_k, and ψ_k are each $N \times 1$ column matrices and correspond to the U, V, and W coordinates of the displacement. Then

$$\mathbf{\Gamma}_k^T = [-\lambda_k \theta_k^T, -\lambda_k \phi_k^T, -\lambda_k \psi_k^T, \theta_k^T, \phi_k^T, \psi_k^T] \tag{4-11a}$$

$$\mathbf{\Upsilon}_k^T = [\lambda_k \theta_k^T, \lambda_k \phi_k^T, \lambda_k \psi_k^T, \theta_k^T, \phi_k^T, \psi_k^T] \tag{4-11b}$$

$$\mathbf{\Gamma}_k^T A \mathbf{\Upsilon}_k = -\lambda_k / Q_k. \tag{4-11c}$$

Let the arbitrary $\mathbf{G}$ of equation (4-1) assume the value, $\mathbf{J}_x$, where

$$\mathbf{J}_x^T = [\Sigma^T, 0^T, 0^T, 0^T, 0^T, 0^T], \tag{4-12}$$

each 0 is understood to be an $N \times 1$ column matrix, and Σ is an $N \times 1$ column matrix,

$$\Sigma^T = [1, 1, 1, \ldots\ldots, 1]. \tag{4-13}$$

The Fourier coefficients for the above $\mathbf{J}_x$, obtained from equations (4-2), (4-4) and (4-11), are

$$\alpha_k = Q_k \Sigma^T M^N \theta_k = Q_k \sum_{i=1}^{n} m^i \theta_k^i;$$

$$\alpha_k^* = Q_k^* \Sigma^T M^N \theta_k^* = Q_k^* \sum_{i=1}^{n} m^i \theta_k^{*i}$$

$$\beta_k = -Q_k \Sigma^T M^N \theta_k = -Q_k \sum_{i=1}^{n} m^i \theta_k^i;$$

$$\beta_k^* = -Q_k^* \Sigma^T M^N \theta_k^* = -Q_k^* \sum_{i=1}^{n} m^i \theta_k^{*i}. \tag{4-14}$$

Likewise

$$\mathbf{J}_x^T A \mathbf{J}_x = \Sigma^T M^N \Sigma = \sum_{i=1}^{N} m^i = m \tag{4-15}$$

where m equals the total mass of the structure. Substitution of equations (4-14) and (4-15) into Bessel's Inequality, equation (4-6) results in

$$m \geq \sum_{k=1}^{n} \{\lambda_k Q_k (\Sigma^T M^N \theta_k)^2 + \lambda_k^* Q_k^* (\Sigma^T M^N \theta_k^*)^2\}. \tag{4-16}$$

Define the quantities, P_{xk}, P_{yk}, and P_{zk} by

$$P_{xk} = \Sigma^T M^N \theta_k = \sum_{i=1}^{N} m^i \theta_k^i \quad P_{yk} = \Sigma^T M^N \phi_k = \sum_{i=1}^{N} m^i \phi_k^i$$

$$P_{zk} = \Sigma^T M^N \psi_k = \sum_{i=1}^{N} m^i \psi_k^i. \tag{4-17}$$

P_{xk}, P_{yk}, and P_{zk} are complex scalars and are the $0x$, $0y$ and $0z$ components of the modal linear momentum coefficient of the k^{th} vibrational mode of the structure. Then equation (4-16) may be written

$$\sum_{k=1}^{n} \{\lambda_k Q_k P_{xk}^2 + \lambda_k^* Q_k^* P_{xk}^{*2}\} \leq m. \tag{4-18a}$$

Equation (4-18a) thus states that the sum of the squares of the model linear momentum coefficients in the $0x$ direction (multipled by scale factors) is bounded, and is less than the total mass of the structure. In equations (4-14) to (4-17), M^N corresponds to the physical model of Chapter 2 and is diagonal.

For the normalization option (b) of the previous chapter $Q_k = 1/2\lambda_k$, $Q_k^* = 1/2\lambda_k^*$, and equation (4-18a) becomes

$$\sum_{k=1}^{n} \frac{1}{2} \{P_{xk}^2 + P_{xk}^{*2}\} \leq m. \tag{4-19}$$

For damping set equal to zero, $P_{xk} = P_{xk}^*$, in which case equation (4-19) compares to equation (25) of Hughes (1980).

A similar procedure with the general form of Bessel's inequality (equation (4-9)), and $\bar{\mathbf{G}} = \mathbf{J}_x$ and $\mathbf{G} = \mathbf{J}_y = [0^T, \Sigma^T, 0^T, 0^T, 0^T, 0^T]^T$ leads to

$$\sum_{k=1}^{n} \{\lambda_k Q_k P_{xk} P_{yk} + \lambda_k^* Q_k^* P_{xk}^* P_{yk}^*\} \leq 0. \tag{4-18b}$$

With other combinations of $\mathbf{J}_x$, $\mathbf{J}_y$, and $\mathbf{J}_z$, and the appropriate forms of Bessel's Equation, the following may likewise be obtained:

$$\sum_{k=1}^{n} \{\lambda_k Q_k P_{yk}^2 + \lambda_k^* Q_k^* P_{yk}^{*2}\} \leq m \tag{4-18c}$$

$$\sum_{k=1}^{n} \{\lambda_k Q_k P_{zk}^2 + \lambda_k^* Q_k^* P_{zk}^{*2}\} \leq m \tag{4-18d}$$

$$\sum_{k=1}^{n} \{\lambda_k Q_k P_{xk} P_{zk} + \lambda_k^* Q_k^* P_{xk}^* P_{zk}^*\} \leq 0, \tag{4-18e}$$

$$\sum_{k=1}^{n} \{\lambda_k Q_k P_{yk} P_{zk} + \lambda_k^* Q_k^* P_{yk}^* P_{zk}^*\} \leq 0, \tag{4-18f}$$

The above six relations can be expressed in a single matrix relation

$$\sum_{k=1}^{n} \{\lambda_k Q_k \mathbf{P}_k \mathbf{P}_k^T + \lambda_k^* Q_k^* \mathbf{P}_k^* \mathbf{P}_k^{*T}\} \leq \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix}, \tag{4-20}$$

where $\mathbf{P}_k^T = \{P_{xk}, P_{yk}, P_{zk}\}$. $\mathbf{P}_k^T$ has dimension 3×1, and is complex in general.

Identities Involving Modal Angular Momentum. Further appropriate assignments of the arbitrary $\mathbf{G}$ and $\bar{\mathbf{G}}$ of Bessel's Inequalities lead to the following identities amongst modal angular momentum coefficients and moments of inertia.

$$\sum_{k=1}^{n} \{\lambda_k Q_k \mathbf{H}_k \mathbf{H}_k^T + \lambda_k^* Q_k^* \mathbf{H}_k^* \mathbf{H}_k^{*T}\} \leq \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{bmatrix} \tag{4-22}$$

where the components of $\mathbf{H}_k$, the modal angular momentum coefficient vector, are

$$H_{xk} = Y^T M^N \psi_k - Z^T M^N \phi_k = \sum_{i=1}^{N} m^i (y^i \psi_k^i - z^i \phi_k^i) \tag{4-23a}$$

$$H_{yk} = Z^T M^N \theta_k - X^T M^N \psi_k = \sum_{i=1}^{N} m^i (z^i \theta_k^i - x^i \psi_k^i) \tag{4-23b}$$

$$H_{zk} = X^T M^N \phi_k - Y^T M^N \theta_k = \sum_{i=1}^{N} m^i (x^i \phi_k^i - y^i \theta_k^i) \tag{4-23c}$$

and I_{xx}, I_{xy}, etc, are the moments of inertia of structure, given by

$$I_{xx} = Z^T M^N Z + Y^T M^N Y = \sum_{i=1}^{N} m^i (y^{i2} + z^{i2}) \qquad (4\text{-}24a)$$

$$I_{xy} = -X^T M^N Y = -\sum_{i=1}^{N} m^i x^i y^i \qquad (4\text{-}24b)$$

and so forth. In equations (4-23) and (4-24), M^N corresponds to the physical model of Chapter 2 and is diagonal, and H_{xk}, H_{yk}, and H_{zk} are complex-valued scalars.

Identities Involving Modal Linear and Angular Momentum. The following identities may also be established.

$$\sum_{k=1}^{n} \{\lambda_k Q_k \mathbf{H}_k \mathbf{P}_k^T + \lambda_k^* Q_k^* \mathbf{H}_k^{*T} \mathbf{P}_k^{*T}\} \leq m \begin{bmatrix} 0 & c_z & -c_x \\ -c_z & 0 & c_y \\ c_x & -c_y & 0 \end{bmatrix}. \qquad (4\text{-}25)$$

In the above, the c's are coordinates of the center of mass,

$$c_x = \sum_{i=1}^{N} m^i x^i / m; \quad c_y = \sum_{i=1}^{N} m^i y^i / m; \quad c_z = \sum_{i=1}^{N} m^i z^i / m \qquad (4\text{-}26)$$

Discussion of Identities. The physical model of Chapter 2, upon which the foregoing derivation is based, becomes the familiar continuum mechanics model when N is taken to infinity. In the limiting case, $\mathbf{P}_k$ and $\mathbf{H}_k$ are defined by $\int \mathbf{\Phi}_k dm$ and $\int \mathbf{r} \times \mathbf{\Phi}_k dm$, respectively, where dm is an elemental mass and $\mathbf{r}$ is its position. Thus the identities given in equations (4-20), (4-22), and (4-25), when specialized to zero damping and $Q_k = 1/2\lambda_k$ ($\mathbf{\Phi}_k^T M \mathbf{\Phi}_k = 1$) and generalized to N equal to infinity, are exactly equal to equations D, E, and F of Hughes (1980).

The identities in the form presented herein possess the attractive features of having a physical interpretation, and of being a generalization of earlier published ones. This form of the identities is applicable to results obtained by finite element or substructure coupling models, if the results are put into the form of the physical variables model, or through mathematical transformation to diagonalize M.

The technique outlined in foregoing for derivation of identities (i.e., use of Bessels Inequalities, equations (4-6) and (4-9), which are applicable for M non-diagonal), can be used to derive equivalent identities directly for a general equation of the form of equation (2-3), without reference to a physical model and coordinates. The resulting algebraic forms turn out to be more complicated and lack the physical interpretation. The form presented herein seems to be more desirable.

6.0 Discussion and Conclusions

The foregoing has developed the natural modes and modal model for an elastic structure with linear viscous damping, via a formulation that is comparable to that of the classical normal modes formulation of the undamped case. Transfer functions and normalizations of use for experimental modal parameter estimation are given special attention. Mass-properties-related modal identities are obtained.

Complex numbers and variables are used herein in order to be compatible with earlier obtained equations. However, it should be noted that the formulation could be done in terms of real-valued modal vectors and real-valued modal variables (or stated another way, the appropriate vector space for this case is a real-valued inner product space). The use of complex quantities is a matter of convenience and not necessity.

The arrangement of the physical-variable equations in state vector form employed herein offers two advantages: it leads naturally to a relatively uncomplicated derivation of mass-properties-related modal identities, and it can be easily generalized further to include gyroscopic forces if desired.

References

Béliveau, J.-G., 1977, "Eigenrelations in Structural Dynamics," *AIAA Journal*, Vol. 15, No. 7, pp. 1039–1041.

Brandon, J. A., 1984, "A Discussion of Alternative Duncan Formulations of the Eigenproblem for the Solution of Non-classically, Viscously Damped Linear Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 904–906.

Brown, D. L., 1984, "Modal Analysis Short Course Notes," *Analytical and Experimental Modal Analysis*, Section 13, Structural Dynamics Research Laboratory, University of Cincinnati.

Caughey, T. K., 1960, "Classical Normal Modes in Damped Linear Dynamics Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 27, pp. 269–271.

Ewins, D. J., 1984, *Modal Testing: Theory and Practice*, Wiley, New York, Chapter 2.

Fawzy, I., 1977, "Orthogonality of Generally Normalized Eigenvectors and Eigenrows," *AIAA Journal*, Vol. 15, No. 2, pp. 276–278.

Foss, K. A., 1958, "Coordinates Which Uncouple the Equations of Motion of Damped Linear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, No. 25, pp. 361–364.

Hughes, P. C., 1980, "Modal Identities for Elastic Bodies, With Application to Vehicle Dynamics and Control," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 177–184.

Meirovitch, L., 1967, *Analytical Methods in Vibrations*, McMillan, New York, Chapter 9.

Meirovitch, L., and Baruh, H., 1981, "Optimal Control of Damped Flexible Gyroscopic Systems," *J. Guidance and Control*, Vol. 4, No. 2, pp. 157–163.

Nelson, H. D., and Glasgow, D. A., 1979, "Eigenrelations for Second Order Systems," *AIAA Journal*, Vol. 17, No. 7, pp. 795–797.

Richardson, M., and Potter, R., 1974, "Identification of the Modal Properties of an Elastic Structure From Measured Transfer Function Data," *Preprint, 20th International Instrumentation Symposium*, ISA ASI 74250, pp. 239–246.

SDRC, 1985, *User Manual For Modal Analysis 9.0*, General Electric CAE International Inc., Milford, Ohio.

Vigneron, F. R., 1981, "Natural Modes and Real Modal Variables for Flexible Spacecraft," CRC Report No. 1348, Department of Communications, Ottawa, Canada.

Vigneron, F. R., 1985, "A Natural Modes Formulation and Modal Identities For Structures With Linear Viscous Damping," CRC Report No. 1382, Department of Communications, Ottawa, Canada.

Wilkinson, J. H., 1965, *The Algebraic Eigenvalue Problem*, Oxford University Press, pp. 4–6.

J. R. Hutchinson
Professor,
Department of Civil Engineering,
University of California,
Davis, California 95616
Mem. ASME

S. D. Zillmer
Technical Staff,
Rockwell International,
North American Aircraft Operations,
El Segundo, Calif. 90245

On the Transverse Vibration of Beams of Rectangular Cross-Section

An exact solution for the natural frequencies of transverse vibration of free beams with rectangular cross-section is used as a basis of comparison for the Timoshenko beam theory and a plane stress approximation which is developed herein. The comparisons clearly show the range of applicability of the approximate solutions as well as their accuracy. The choice of a best shear coefficient for use in the Timoshenko beam theory is considered by evaluation of the shear coefficient that would make the Timoshenko beam theory match the exact solution and the plane stress solution. The plane stress solution is shown to provide excellent accuracy within its range of applicability.

Introduction

In a recent paper by these authors [1] an exact solution for the vibrations of a solid isotropic linearly elastic rectangular parallelepiped with traction-free boundaries was developed. The flexural vibrations of beams of rectangular cross-section represents an important subset of that exact solution. The exact solution is used in this paper to evaluate the accuracy and range of applicability of the Timoshenko beam theory as well as a plane stress approximation which is developed herein.

The Timoshenko beam theory, in which shear and rotary inertia effects are included, has been the subject of many papers, with particular emphasis placed on evaluation of a correct shear coefficient. A 1975 review paper by Kaneko [2] lists some twenty different values of shear coefficients which have been used by various authors for beams of rectangular cross section. Kaneko concludes that for rectangular beams the shear coefficient implied in Timoshenko's 1922 paper [3] gives the best match to the experimental results. That value of shear coefficient is $(5+5\nu)/(6+5\nu)$ where ν is Poisson's ratio. In this paper that particular shear coefficient will be referred to as Timoshenko's shear coefficient, even though Timoshenko himself used other values in his work. The general shear coefficient in the Timoshenko beam equation will simply be referred to as the shear coefficient. Timoshenko's shear coefficient was derived by matching with the plane stress solution for long wavelengths. Another shear coefficient which will be used for comparison is $\pi^2/12$ derived by Mindlin and Deresiewicz [4] in 1954. They derived their shear coefficient by matching the three-dimensional equations for simple thickness-shear motions of infinite beams. In this paper their shear coefficient will be referred to as Mindlin's shear coefficient.

In a 1981 paper [5] by the senior author an exact solution for the vibrations of a traction-free solid isotropic linearly elastic circular cylinder was used to evaluate both the Timoshenko beam theory, for free-free beams with circular cross section, and an approximate solution which was developed by Pickett [6]. It was found that the Pickett formulation led to highly accurate solutions in its range of applicability. In the Pickett solution the governing equations are satisfied identically in the domain and the boundary conditions on the lateral surfaces are also identially satisfied. The boundary conditions on the ends are then approximated by setting the resultant moment and shear to zero.

Because of the success of the Pickett solution for beams of circular cross section a similar type of solution was attempted for beams of rectangular cross section. For beams of rectangular cross section it is not possible, in general, to combine solution forms of the governing equations so as to satisfy identically the traction-free boundary conditions on the lateral surfaces of the beam. It was found to be possible, however, to find a plane strain solution for which the boundary conditions could be identically satisfied on the upper and lower beam surfaces and then the end boundary conditions could be satisfied as Pickett had done. The plane strain solution is then converted to a plane stress solution by a simple change in the elastic constants as suggested by Timoshenko in his 1922 paper. The range of applicability of this plane stress solution is shown to be the same as for the Timoshenko beam theory, but its accuracy is greater.

A plane strain solution was previously evaluated by Fromme and Leissa [7] in 1970. They compared the plane strain solution to the Timoshenko beam solution using a shear coefficient of 5/6. They found excellent correlation for $\nu=0$ but for $\nu=0.3$ there was a discrepancy in the frequencies. If they had converted their plane strain solution to plane stress by a change of elastic constants and had used Timoshenko's

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami Beach, Fla., November 17-22 1984 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. 10077, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 30, 1984; final revision, March 4, 1985. Paper No. 85-WA/APM-22.

Table 1 Plane strain solution forms. The dimensionless wave numbers β, δ, $\bar{\beta}$, and $\bar{\delta}$ and the frequency parameter ω are related by $\beta^2+\delta^2=\omega^2(1-2\nu)/(2-2\nu)$ and $\bar{\beta}^2+\bar{\delta}^2=\omega^2$.

	Form 1	Form 2
u	0	0
v	$\cos(\beta y)\begin{Bmatrix}\cos(\delta z)\\ \delta\sin(\delta z)\end{Bmatrix}$	$\cos(\bar{\beta} y)\begin{Bmatrix}\cos(\bar{\delta} z)\\ \bar{\delta}\sin(\bar{\delta} z)\end{Bmatrix}$
w	$\delta^2\dfrac{\sin(\beta y)}{\beta}\begin{Bmatrix}-\sin(\delta z)/\delta\\ \cos(\delta z)\end{Bmatrix}$	$-\bar{\beta}\sin(\bar{\beta} y)\begin{Bmatrix}-\sin(\bar{\delta} z)/\bar{\delta}\\ \cos(\bar{\delta} z)\end{Bmatrix}$
σ_x	$-\omega^2\left(\dfrac{\nu}{1-\nu}\right)\dfrac{\sin(\beta y)}{\beta}\begin{Bmatrix}\cos(\delta z)\\ \delta\sin(\delta z)\end{Bmatrix}$	0
σ_y	$-\left[\omega^2\left(\dfrac{\nu}{1-\nu}\right)+2\beta^2\right]\dfrac{\sin(\beta y)}{\beta}\begin{Bmatrix}\cos(\delta z)\\ \delta\sin(\delta z)\end{Bmatrix}$	$^{-2\bar{\beta}}\sin(\bar{\beta} y)\begin{Bmatrix}\cos(\bar{\delta} z)\\ \bar{\delta}\sin(\bar{\delta} z)\end{Bmatrix}$
σ_z	$-\left[\omega^2\left(\dfrac{\nu}{1-\nu}\right)+2\delta^2\right]\dfrac{\sin(\beta y)}{\beta}\begin{Bmatrix}\cos(\delta z)\\ \delta\sin(\delta z)\end{Bmatrix}$	$^{2\bar{\beta}}\sin(\bar{\beta} y)\begin{Bmatrix}\cos(\bar{\delta} z)\\ \bar{\delta}\sin(\bar{\delta} z)\end{Bmatrix}$
τ_{xy}	0	0
τ_{yz}	$2\delta^2\cos(\beta y)\begin{Bmatrix}-\sin(\delta z)/\delta\\ \cos(\delta z)\end{Bmatrix}$	$(\bar{\delta}^2-\bar{\beta}^2)\cos(\bar{\beta} y)\begin{Bmatrix}-\sin(\bar{\delta} z)/\bar{\delta}\\ \cos(\bar{\delta} z)\end{Bmatrix}$
τ_{zx}	0	0

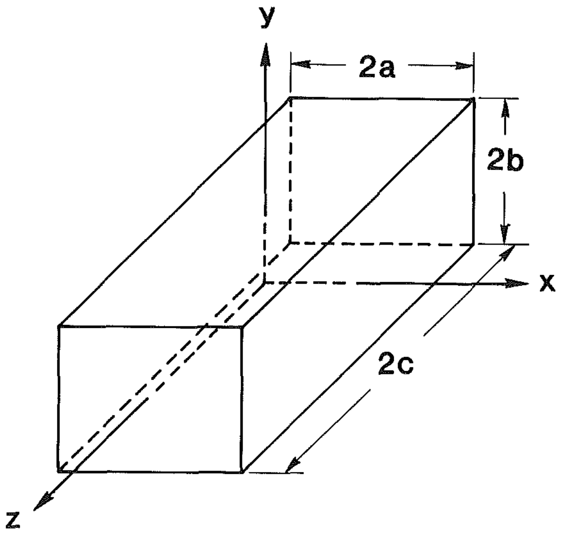


Fig. 1 Coordinates and dimensions of beam

shear coefficient they would have found the same excellent correlation for $\nu=0.3$ as they found for $\nu=0$. The solution of Fromme and Leissa can be viewed as an exact solution of the plane strain problem; whereas, the solution we have proposed herein only approximates the end boundary conditions. The new method is, however, computationally simpler, is easily modified to allow for arbitrary end boundary conditions, and produces excellent results.

Solutions

The solution forms for the various analytic solutions considered in this paper are discussed in this section. Both the elementary beam solution (i.e., the Euler-Bernoulli beam solution) and the Timoshenko beam solution appear in many references, for example, see [5]. For this reason their derivations will not be given here. As pointed out in the introduction a shear coefficient K, appears in the Timoshenko beam equation and has been the subject of many papers. In this paper the Timoshenko beam equation is solved to determine the value K would have to obtain in order to cause the Timoshenko beam solution to produce identical frequencies as the exact solution and the plane stress solution.

The coordinates and dimensions for the beam are shown in Fig. 1. The length direction is considered to be z. The direction in which the transverse beam motion is taking place is the y-direction. The length of the beam is thus $2c$, the depth is $2b$, and the width is $2a$. Choice of the centered origin allows a convenient splitting of the problem into even and odd functions of x, y, and z. Since we are considering only beam motion we limit the solution to forms where v, the displacement in the y direction, is an even function of both x and y.

The exact solution is derived in detail in [1] but will be briefly discussed.

Exact Solution. The exact solution is a series solution in which each term of the series identically satisfies the linear elasticity equations. There are four types of solution forms which can be derived from the Helmholtz Displacement Potential. These forms were tabulated in Table 1 of [1]. Linear combinations of the four solution types were chosen in three double series. Then by an appropriate choice of wave numbers and constants the boundary conditions on the shear stresses were identically satisifed. The boundary conditions on the normal stresses were then satisfied by orthogonalization on the boundary. This process leads to a matrix of coefficients whose determinant must be zero. The coefficients are transcendental functions of the natural frequency, and the order of the matrix is determined by the numer of terms retained in the infinite series.

The convergence of the solution to known experimental results, as more and more terms in the series were retained, was demonstrated in [1]. The convergence was monotonic from above. It was noted, however, that when the rectangular parallelepiped was a cube the convergence was excellent, but as one or two of the dimensions became large compared to the other(s) the rate of convergence diminishes. That is, it takes more terms, hence larger matrices, for equivalent accuracy, as the rectangular parallelepiped becomes less cube-like. Since all computations were performed on a minicomputer this precluded evaluating the exact solutions for long beams. The minicomputer solutions, however, were more than adequate for the comparisons performed in this paper.

It should perhaps be noted here that in [1] Figs. 4 and 5 appear above the captions for Figs. 6 and 7 and vice versa.

Plane Stress Solution. A plane strain solution is first formed from the two solution forms shown in Table 1. These solution forms are taken from the solution forms listed in columns 1 and 4 of Table 1 in [1] and can be seen to represent the plane strain case for the y-z-plane. That is, u the displacement in x-direction is zero and there is no functional dependence on x. The symmetric and antisymmetric forms in z are shown in braces with the symmetric (v as an even function of z) form on top. All values in Table 1 are dimensionless. The displacements u, v, and w were made dimensionless by dividing by the depth $2b$. All lengths were made dimensionless by dividing by the depth. All stress quantities were made dimensionless by dividing by the shear modulus. The wave numbers β, δ, $\bar{\beta}$, $\bar{\delta}$ were made dimensionless by multiplying by the depth. The frequency was made dimensionless by multiplying by the depth and dividing by the shear wave velocity. The solution is taken as the column 1 form multiplied by A plus the column 2 form multiplied by B.

The boundary conditions on the top and bottom surface of the beam are satisifed by letting $\delta=\bar{\delta}$, which allows both forms to have the same z dependence, and then setting

$$\sigma_y(\pm b,z)=0 \tag{1}$$

$$\tau_{yz}(\pm b,z)=0 \tag{2}$$

This yields a 2 × 2 set of homogeneous equations

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{Bmatrix} A \\ B \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{3}$$

where

$$a_{11} = -[\omega^2\nu/(1-\nu)+2\beta^2]\frac{\sin(\beta b)}{\beta} \tag{4}$$

$$a_{12} = -2\bar{\beta}\sin(\bar{\beta}b) \tag{5}$$

$$a_{21} = 2\delta^2\cos(\beta b) \tag{6}$$

$$a_{22} = -(\bar{\beta}^2-\delta^2)\cos(\bar{\beta}b) \tag{7}$$

For a solution to exist the determinant of the coefficients in equation (3) must be zero. As in Pickett's solution it is found that for values of ω less than about 8.5 there are only two real values of δ^2. A plot of the two values of δ^2 as a function of ω is shown in Fig. 2. This plot was for a specific Poisson's ratio. For other values of Poisson's ratio the curves will shift slightly but be similar to the plot shown. On finding the values of δ^2 the ratio of B to A can be expressed as

$$B/Z=-a_{11}/a_{12} \tag{8}$$

Using the two values of δ^2 allows satisfaction of the usual type of beam boundary conditions. For the free-free beam the boundary conditions are

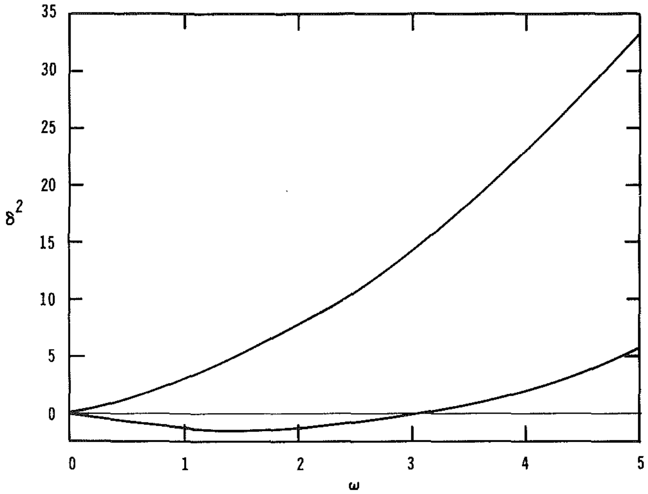


Fig. 2 The two values of δ^2 as a function of ω for $\nu=.3$

$$V = \int_{-b}^{b}\tau_{yz}(y,\pm c)dy=0 \tag{9}$$

$$M = \int_{-b}^{b}\sigma_z(y,\pm c)y\,dy=0 \tag{10}$$

Satisfying these boundary conditions leads to the set of equations

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{Bmatrix} A_1 \\ A_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{11}$$

where

$$b_{1j} = \left[2\delta^2\frac{\sin(\beta b)}{\beta}-\frac{B}{A}(\bar{\beta}^2-\delta^2)\frac{\sin(\bar{\beta}b)}{\bar{\beta}}\right]\begin{Bmatrix} -\sin(\delta c)/\delta \\ \cos(\delta c) \end{Bmatrix} \tag{12}$$

$$\begin{aligned} b_{2j} = {} & \left\{\left[\frac{\omega^2\nu}{1-\nu}+2\delta^2\right]\left[\frac{1}{2\beta^2}\cos(\beta b)-\frac{1}{\beta^3}\sin(\beta b)\right]\right. \\ & \left. -2\frac{B}{A}\left[\frac{\cos(\bar{\beta}b)}{2}-\frac{\sin(\bar{\beta}b)}{\bar{\beta}}\right]\right\}\begin{Bmatrix} \cos\delta c \\ \delta\sin\delta c \end{Bmatrix} \end{aligned} \tag{13}$$

where $j=1$, 2. Upper and lower forms in braces are the symmetric and antisymmetric forms respectively. Subscripts j on the wave numbers and on the ratio B/A are implied. The determinant of the coefficients of equation (11) must be zero for a solution to exist.

So far the solution is a plane strain solution. To convert to a plane stress solution a simple change is made in the elastic constants as done by Timoshenko [3] in 1922. In this formulation the shear modulus was used for the nondimensionalization. Since the shear modulus does not change from plane strain to plane stress, the only elastic constant left in our formulation is Poisson's ratio. If an equivalent Poisson's ratio of

$$\nu_{eq}=\nu/(1+\nu) \tag{14}$$

is used to replace ν in all previous formulas we have the plane stress solution.

The solution process is to choose ν, ω, and c, find the two values of δ^2 from equation (3), find the two values of B/A from equation (8), then check if the determinant of the coefficients in equation (11) is zero. If not a new value of either ω or c is assumed and the process repeated.

The solution process also can be applied to beams with arbitrary boundary conditions. The only difference is in the coefficients in equation (11). For the free-free beam, use was made of the fact that the solution separated into symmetric

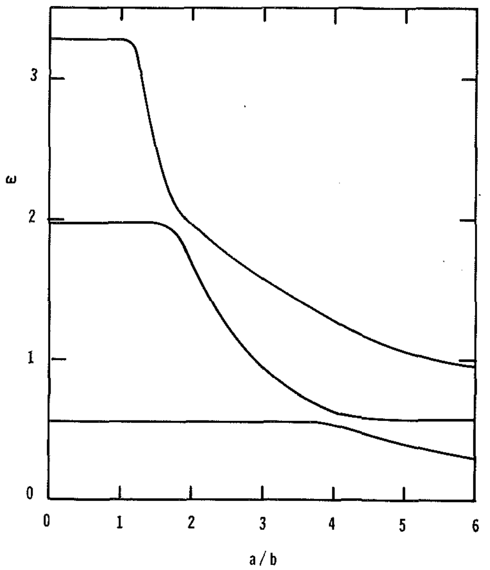


Fig. 3 Exact frequency versus the width-to-depth ratio a/b for the three lowest symmetric modes, $\nu = .3$ and $c/b = 4$

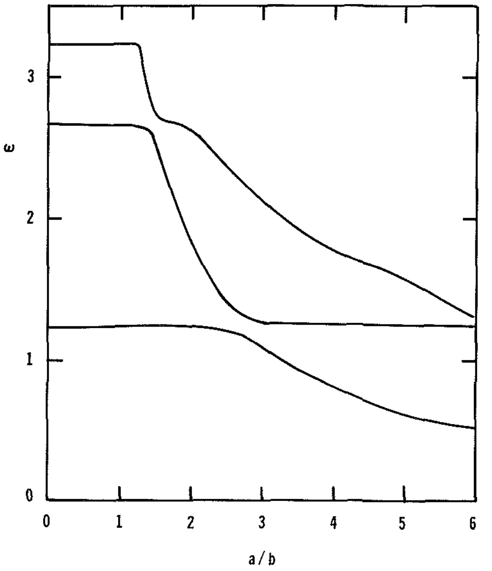


Fig. 4 Exact frequency versus the width-to-depth ratio a/b for the three lowest antisymmetric modes, $\nu = .3$ and $c/b = 4$

and antisymmetric parts which lead to equation (11) being a 2 × 2 system of equations. In more general cases (e.g., a cantilever beam) equation (11) will be replaced with a 4 × 4 system of equations. In setting up and solving these equations, however, the same procedure as outlined above is followed.

In the case of clamped boundary conditions a certain arbitrariness exists. This arbitrariness is the same as noted by Timoshenko and Goodier [8] for the polynomial solution of the static plane stress problem of an end loaded cantilever beam. In both cases it is not possible to identically satisfy the clamped boundary conditions. Timoshenko and Goodier restricted the rigid body displacement by fixing the displacement at the neutral axis at the clamped end. Then they considered several possibilities for restricting the rigid body rotation. One was to fix the slope of the neutral axis by setting,

$$v_{,z}(0,c) = 0 \tag{15}$$

where the comma subscript denotes differentiation, and, as before, v is the transverse displacement as a function of y and z, respectively. Another possibility considered was to fix the slope of the cross section at the neutral axis by setting

$$w_{,y}(0,c) = 0 \tag{16}$$

A third possibility was to set the axial displacement at the beam corner to zero

$$w(b,c) = 0 \tag{17}$$

By investigating Fig. 27 of [8] it is obvious that for the static problem the boundary condition expressed in equation (15) will produce displacements which are too small while those expressed in (16) will produce displacements which are too large. The boundary condition expressed in equation (17) represents a compromise between the two extremes but will still result in a displacement which is too small.

In the vibrating beam problem the result of applying either equations (15) or (17) is to over constrain the beam leading to high estimates of the natural frequencies. The result of applying equation (16) is to under constrain the beam leading to low estimates of the natural frequencies. Another better compromise is to choose some weighted average of the axial displacement as zero such as

$$\int_0^b y w(y,c) dy = 0 \tag{18}$$

The Timoshenko beam theory does not contain this ambiguity. The cross-sectional slope in the Timoshenko beam theory is actually an average slope. Thus, setting that average slope (along with the transverse displacement) to zero represents a correct clamped boundary condition.

In a recent paper by Levinson [9] a new beam formulation of the same order as the Timoshenko formulation was developed. Levinson's formulation, however, has several drawbacks. One of the drawbacks is that an equivalent shear coefficient of 5/6 automatically results. While 5/6 is not a bad choice for a shear coefficient there are slightly better values which can be used as shown in both [2] and [5] as well as in this paper. A more serious drawback to Levinson's theory, however, is that his cross-sectional slope term is the cross-sectional slope at the neutral axis. In applying clamped boundary conditions Levinson sets the cross-sectional slope at the neutral axis to zero. He shows that his solution is identical to the above mentioned Timoshenko and Goodier solution using the boundary condition expressed in equation (16). Therefore, for any beam problem involving clamped conditions the Levinson theory will result in deflections which are too large in static problems, and frequencies which are too low in vibration problems. The same drawbacks arise in Levinson's plate theory [10]. Thus, while fundamentally correct the Levinson beam and plate theories do not represent improvements on the Timoshenko beam and Mindlin plate theories.

Results

Figures 3 and 4 are plots of the three lowest natural frequencies for symmetric and antisymmetric beam type modes respectively. These plots show the exact solutions for the free rectangular parallelepiped. The variation of the natural frequencies with the ratio of width-to-depth of the beam is considered for a constant length-to-depth ratio of four. It can be seen that for small width-to-depth ratios the frequencies are virtually independent of the width-to-depth ratio. If the fundamental frequency in Fig. 3 is considered, for instance, it is seen to have a constant value near .5 until the width-to-depth ratio is almost four. The width-to-depth ratio of four in these plots corresponds to a square plate. For $a/b < 4$ the fundamental frequency corresponds to the fundamental beam frequency. At the knee of that curve where $a/b = 4$ the fundamental frequency corresponds to the fundamental frequency of a square plate. For $a/b > 4$ the fun-

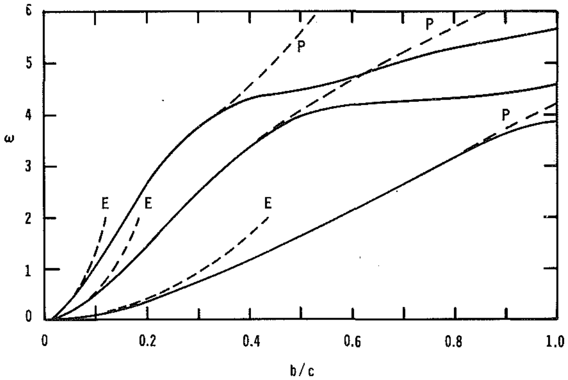


Fig. 5 Frequency versus depth-to-length ratio *b/c* for the three lowest symmetric modes for a square cross section $a/b = 1$ and $\nu = .3$

___ E: elementary beam solution
___ P: plane stress solution
___ : exact solution

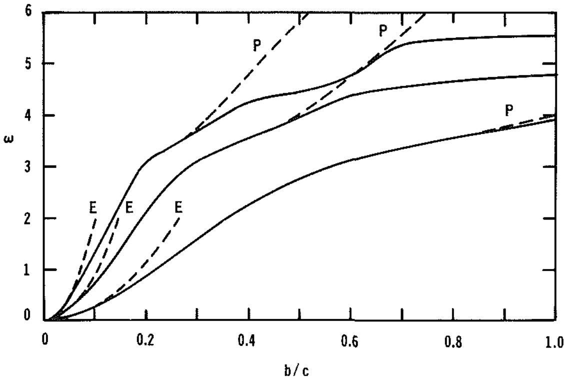


Fig. 6 Frequency versus depth-to-length ratio *b/c* for the three lowest antisymmetric modes for a square cross-section $a/b = 1$ and $\nu = .3$

___ E: elementary beam solution
___ P: plane stress solution
___ : exact solution

damental frequency again corresponds to a fundamental beam frequency, but not to a beam of width $2a$ and length $2c$ but rather to a beam of length $2a$ and width $2c$. For instance, the lowest frequency of the beam when a/b equals 6 corresponds to the fundamental frequency of a beam whose length is six times its depth. A similar observation can be made for all the curves in Figs. 3 and 4. That is, just to the right of the knee in each curve the negative sloping lines correspond to the fundamental beam frequency for a beam of length-to-depth ratio of a/b. Thus it may be concluded that the frequencies found in applying either the Timoshenko beam theory or the plane stress theory are limited to frequencies which are less than the fundamental frequency of a beam whose length is $2a$ (i.e., a beam whose length is the width dimension of the beam under consideration).

Figures 5 and 6 show the frequency variation with the depth-to-length ratio for a square beam and contrast the exact solution with the elementary beam solution and the plane stress solution. The elementary beam solution is shown to hold only over a very limited region; whereas, the plane stress solution is good over a much greater region. The Timoshenko beam solution is not shown on this plot because it is practically indistinguishable from the plane stress solution.

Figures 7 and 8 show the shear coefficient which would have to be used in the Timoshenko beam equation if identical matching with the exact solution or plane stress solution were required. These curves are for a beam of square cross section. It can be seen that as b/c approaches zero the plane stress solution approaches exactly Timoshenko's shear coefficient. This is also true for other values of Poisson's ratio. The exact solution was not carried out for small b/c for two reasons. First, as b/c gets smaller more terms are required in the solution for equivalent accuracy, and second, as b/c gets smaller greater accuracy is required in the frequencies in order to compute the shear coefficient. Computation was therefore limited to the values shown because of practical limitations of the minicomputer which mass used. The plane stress solution, however, does an excellent job of handling the solution in the region where the exact solution is not practical. For thinner beams, for example, a beam with $a/b = .5$, the exact curves, particularly for the first mode, are much closer to the plane stress curves than those shown in Figs. 7 and 8.

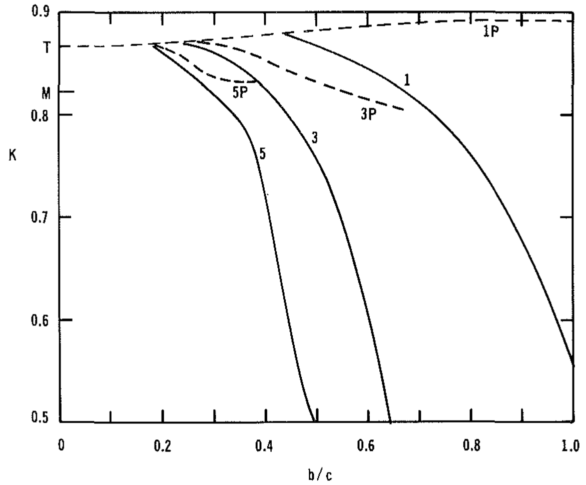


Fig. 7 Shear coefficient *K* versus the depth-to-length ratio *b/c* for the three lowest symmetric modes, $\nu = .3$, $a/b = 1$, ___ exact solution; ___ plane stress solution. Numbers refer to beam mode. T and M mark Timoshenko's and Mindlin's shear coefficient respectively.

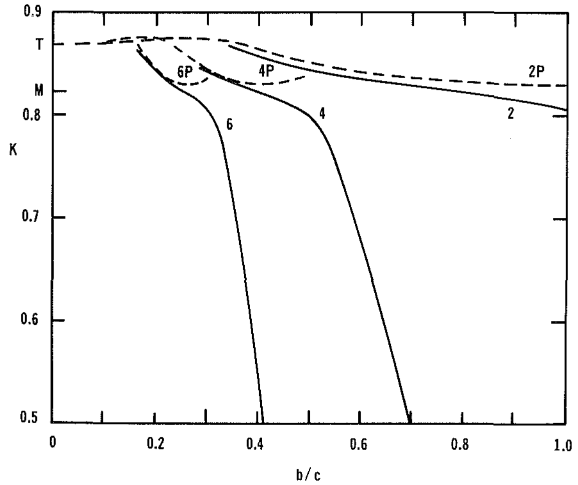


Fig. 8 Shear coefficient *K* versus the depth-to-length ratio b/c for the three lowest antisymmetric modes, $\nu = .3$, $a/b = 1$, ___ exact solution; ___ plane stress solution. Numbers refer to beam mode. T and M mark Timoshenko's and Mindlin's shear coefficient respectively.

Figures 7 and 8 also show that use of the Timoshenko shear coefficient in the Timoshenko beam theory will give excellent results for long wave lengths. A rule-of-thumb can be seen to be that if the wave length is greater than twice the beam depth the Timoshenko shear coefficient, $(5+5\nu)/(6+5\nu)$, will give good results. For shorter wave lengths, however, a smaller shear coefficient would have to be used to produce good matching with the exact solution.

Figures 9 and 10 show the frequency variation with depth-to-length ratio for a beam whose width is twice its depth. A comparison of Fig. 5 with Fig. 9 and Fig. 6 with Fig. 10 shows that they are identical for $\omega < 1.6$. A plateau is reached on the curves in Figs. 9 and 10 near $\omega = 1.7$ and the entire spectrum

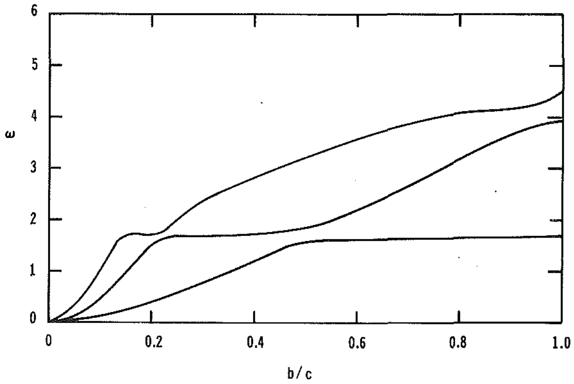


Fig. 9 Exact frequency versus the depth-to-length ratio for the lowest three symmetric modes for $a/b = 2$ and $\nu = .3$

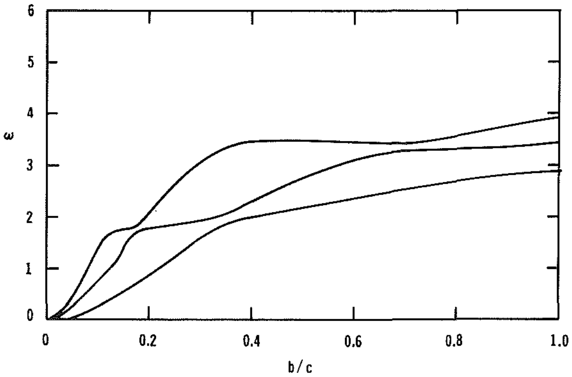


Fig. 10 Exact frequency versus the depth-to-length ratio for the lowest three antisymmetric modes for $a/b = 2$ and $\nu = .3$

shifts at that value. From Fig. 5 the fundamental frequency of a beam with $b/c = 0.5$ has a frequency of 1.7. The spectrum shift in Figs. 9 and 10 therefore occur at a frequency corresponding to the fundamental mode of vibration where the width direction is taken as the beam length. This supports the previous conclusion on the limit of high order beam theories. Similar curves were also generated for the case where $a/b = 1/2$. These curves were very close to the plane stress solution shown in Figs. 5 and 6 and so are not shown here.

It has been noted that the near coalescence and subsequent divergence of the curves in Figs. 3, 4, 5, 6, 9, and 10 is a function of Poisson's ratio. For a Poisson's ratio of zero the coupling is absent and an actual coalescence is found with the lines simply crossing instead of veering away. Since Poisson's ratio is seldom zero this observation is primarily of academic interest.

Figure 11 shows the variation of the shear coefficient with Poisson's ratio. The values corresponding to Timoshenko's and Mindlin's shear coefficients can be compared to those of two specific cases found using the exact solution and the plane stress solution. The specific cases were chosen by looking for a case near the Timoshenko coefficient value in Fig. 7 and a case near the Mindlin coefficient value in Fig. 8. It can be seen that for the first mode for $b/c = 0.5$ the Timoshenko coefficient best represents the variation; whereas, for the second mode for $b/c = 0.8$ the Mindlin coefficient best represents the variation. Kaneko [2] concluded that the Timoshenko coefficient was best on the basis of experimental results; however, all the experiments reported were for the fundamental frequency and a depth-to-length ratio less than 1/2, a region in which Timoshenko's coefficient does prove best.

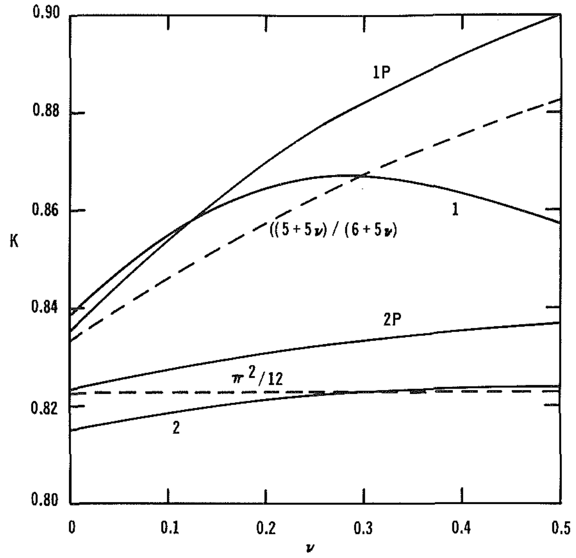


Fig. 11 Shear coefficient K versus Poisson's ratio for a square cross-section beam.

– – – Timoshenko's and Mindlin's shear coefficients
—— exact and plane stress solutions

1: first mode exact $b/c = .5$
1P: first mode plane stress $b/c = .5$
2: second mode exact $b/c = .8$
2P: second mode plane stress $b/c = .8$

Conclusions

- The plane stress and Timoshenko beam solution produce similar results with the same range of applicability.
- The plane stress solution gives greater accuracy than the Timoshenko beam theory but is slightly more difficult computationally.
- The plane stress and Timoshenko beam solutions are applicable for frequencies less than the fundamental frequency found by treating the width of the beam as the length dimension.
- For wavelengths longer than twice the beam depth a constant shear coefficient of $(5+5\nu)/(6+5\nu)$ gives good matching with experiment as well as the plane stress and exact solutions.

References

1 Hutchinson, J. R., and Zillmer, S. D., "Vibration of a Free Rectangular Parallelpiped," ASME Journal of Applied Mechanics, Vol. 50, 1983, pp. 123-130.

2 Kaneko, T., "On Timoshenko's Correction for Shear in Vibrating Beams," *Journal of Physics D: Applied Physics*, Vol. 8, 1975, pp. 1927-1936.

3 Timoshenko, S. P., "On the Transverse Vibrations of Bars of Uniform Cross-Section," *Philosophical Magazine*, Vol. 43, 1922, pp. 125-131.

4 Mindlin, R. D., and Deresiewicz, H., "Timoshenko Shear Coefficient for Flexural Vibrations of Beams," *Proceedings of 2nd U.S. National Congress of Applied Mechanics*, ASME, New York, 1954, pp. 175-178.

5 Hutchinson, J. R., "Transverse Vibrations of Beams, Exact Versus Approximate Solutions," ASME Journal of Applied Mechanics, Vol. 47, 1981, pp. 923-928.

6 Pickett, G., "Flexural Vibrations of Unrestrained Cylinders and Disks," *Journal of Applied Physics*, Vol. 16, 1945, pp. 820-831.

7 Fromme, J. A., and Leissa, A. W., "Free Vibration of the Rectangular Parallelepiped," *The Journal of the Acoustical Society of America*, Vol. 48, 1970, pp. 290-298.

8 Timoshenko, S. P., and Goodier, J. N., *Theory of Elasticity* (3rd ed.), McGraw-Hill, New York, 1970, pp. 42-46 and prob. 5 p. 63.

9 Levinson, M., "A New Rectangular Beam Theory," *Journal of Sound and Vibration*, Vol. 74, 1981, pp. 81-87.

10 Levinson, M., "An Accurate, Simple Theory of the Statics and Dynamics of Elastic Plates," *Mechanics Research Communications*, Vol. 7, 1980, pp. 343-350.

M. B. Rubin
Faculty of Mechanical Engineering,
Technion-Israel Institute of Technology,
Haifa 32000 Israel

Free Vibration of a Rectangular Parallelepiped Using the Theory of a Cosserat Point

Free vibration of a rectangular parallelepiped composed of a homogeneous linear elastic isotropic material is studied. The parallelepiped is modeled as an isotropic Cosserat point and simple formulas are developed to predict the lowest frequencies of vibration. Within the context of the theory, extensional and shear vibrations are uncoupled. The lowest extensional frequency predicted by the Cosserat theory is compared with available exact solutions and with predictions of thin rod theory. Finally, by introducing a simple modification of the director inertia coefficient it is shown that the Cosserat predictions of the extensional frequencies are correct.

Introduction

Recently the theory of a Cosserat point (Rubin, 1985a) has been developed to model the deformation of a body which is essentially a material point surrounded by some small volume. The development of this theory parallels the developments of the theory of a Cosserat surface (Naghdi, 1972), which models a shell-like body, and that of a Cosserat curve (Green et al., 1974 a, b), which models a rod-like body.

An important feature of the theory of a Cosserat point is that it can be used as a basis for developing numerical solution procedures for continuum problems. In particular, we mention that the theory has been successfully used to formulate the numerical solution of one-dimensional continuum problems (Rubin, 1985b).

In this paper, we consider the three-dimensional problem of linear vibration of a free rectangular parallelepiped composed of a homogeneous linear elastic isotropic material. In its reference configuration, the parallelepiped has dimensions L_1, L_2 and L_3 (see Fig. 1). As an example, it was shown in (Rubin, 1985a) that such a parallelepiped could be modeled as an isotropic Cosserat point. The equations describing motions of the parallelepiped were developed, but not solved. Here, we specialize the equations to describe free vibration of the parallelepiped and solve them for the natural frequencies.

This difficult problem has been considered by many researchers. Most recently, Hutchinson and Zillmer (1983) have developed a series solution which has been compared with elementary solutions and the experiments of Spinner et al. (1960). Here, we use the Cosserat theory to predict simple formulas for the lowest frequencies of vibration associated with particular modes. Specifically, it will be shown that the simple[1] theory of a Cosserat point predicts extensional and shear vibration that are uncoupled. Furthermore, we note that the deformation associated with the simple theory corresponds to a three-dimensional state of homogeneous strain so torsional and bending modes of the parallelepiped are not admitted.

In the following sections we develop the basic equations and solve for the frequencies of extensional and shear vibrations. The lowest extensional frequency is compared with available exact solutions for a square bar (Appendix A), a cube and with that predicted by thin rod theory. The comparison in each of these cases is good. We close by introducing a simple modification of the director inertia coefficient which significantly improves the prediction of each of these frequencies.

Basic Equations

In this section, we summarize the basic equations of the theory of a Cosserat point which was developed in (Rubin, 1985a). In the present configuration at time t, the Cosserat point occupies a region R of Euclidean 3-space bounded by the closed surface ∂R which consists of a finite set of sections ∂R_K $(K = 1, 2, \ldots, M)$. Furthermore, with respect to the present configuration, a Cosserat point is defined by its location $\mathbf{r}(t)$, relative to the origin of a fixed coordinate system, and by three directors $\mathbf{d}_i(t)$ $(i = 1, 2, 3)$. The position vector $\mathbf{r}$ and directors $\mathbf{d}_i$ are each a three-dimensional vector function of time only which in the reference configuration acquire the values $\mathbf{r} = \mathbf{R}$ and $\mathbf{d}_i = \mathbf{D}_i$. A motion of the Cosserat point is defined by

$$\mathbf{r}=\mathbf{r}(t),\ \ \mathbf{d}_i=\mathbf{d}_i(t),\ \ [\mathbf{d}_1,\mathbf{d}_2,\mathbf{d}_3]>0 \qquad (1a,b,c)$$

where the condition (1c) ensures that the directors form a right-handed set of linearly independent vectors.

The vector functions $\mathbf{r}$ and $\mathbf{d}_i$ in (1) are assumed to be suffi-

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February 20, 1984; final revision, June 6, 1985.

---

[1] We use the word simple to describe the theory which admits only three directors.

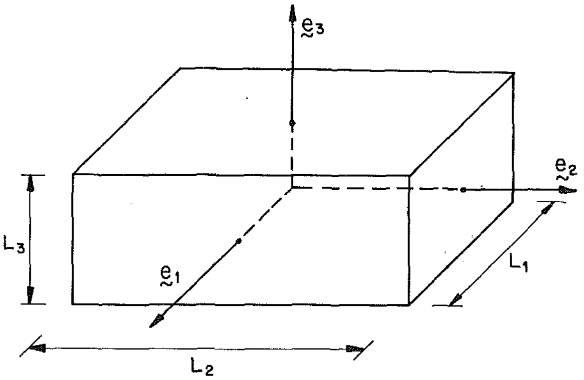


Fig. 1 A rectangular papallelepiped with dimensions L_1, L_2, L_3

ciently smooth in the sense that they are differentiable with respect to time t as many times as required. Let $\mathbf{v}$ and $\mathbf{w}_i$ denote, respectively, the velocity of the Cosserat point and the director velocities at time t. Then

$$\mathbf{v}=\dot{\mathbf{r}},\ \ \mathbf{w}_i=\dot{\mathbf{d}}_i \tag{2a,b}$$

where a superposed dot denotes time differentiation.

For later convenience, we introduce the reciprocal vectors $\mathbf{d}^i$ (i = 1, 2, 3) of the directors $\mathbf{d}_i$ and record the kinematical definitions

$$h_{ij}=\mathbf{d}_i\cdot\mathbf{d}_j,\ \ \mathbf{d}_i\cdot\mathbf{d}^j=\delta_i^j,\ \ h^{ij}=\mathbf{d}^i\cdot\mathbf{d}^j \tag{3a,b,c}$$

where δ_i^j is the Kronecker delta symbol. Furthermore, we introduce the strain γ_{ij} relative to the reference configuration through the expression

$$\gamma_{ij}=½(h_{ij}-H_{ij}) \tag{4}$$

where H_{ij} is the reference value of h_{ij}.

We now turn to a statement of the conservation and balance laws of the theory and, with reference to the present configuration, define the following quantities: the mass $m(t)$ of the point; the contact force $\mathbf{n}_K(t)$ and contact director couples $\mathbf{m}_K{}^i(t)$ applied to the section ∂R_K of the boundary ∂R; the external body force $\mathbf{f}(t)$ and external body director couples $\mathbf{l}^i(t)$; the intrinsic director couples $\mathbf{k}^i(t)$, which make no contribution to the supply of angular momentum, and the constant inertia coefficients y^i and y^{ij}, with y^{ij} being a symmetric tensor. With the above definitions, and with reference to the present configuration, the balance and conservation laws may be written in the forms[2]:

$$\dot{m}=0, \tag{5a}$$

$$m(\dot{\mathbf{v}}+y^i\dot{\mathbf{w}}_i)=\mathbf{f}+\sum_{K=1}^{M}\mathbf{n}_K, \tag{5b}$$

$$m(y^i\dot{\mathbf{v}}+y^{ij}\dot{\mathbf{w}}_j)=\mathbf{l}^i+\sum_{K=1}^{M}\mathbf{m}_K{}^i-\mathbf{k}^i, \tag{5c}$$

$$\mathbf{d}_i\times\mathbf{k}^i=0. \tag{5d}$$

Equation (5a) represents conservation of mass, (5b) represents the balance of linear momentum, (5c) represents the balance of director momentum, and (5d) represents the reduced form of the balance of angular momentum.

For the linearized theory we assume that in its reference configuration the Cosserat point is in its natural state and is free of body forces and director couples, contact forces and director couples, and intrinsic director couples. Let $\mathbf{u}(t)$ be the displacement of the Cosserat point and $\bar{\delta}_i(t)$ be the displacements of the directors relative to the reference configuration, so that

$$\mathbf{r}=\mathbf{R}+\mathbf{u},\ \ \mathbf{d}_i=\mathbf{D}_i+\bar{\delta}_i \tag{6a,b}$$

where $\mathbf{u}$ and $\bar{\delta}_i$ are considered to be small relative to $\mathbf{R}$ and $\mathbf{D}_i$, respectively. Neglecting quadratic terms in $\mathbf{u}$ and $\bar{\delta}_i$ the equations of motion (5b,c,d) may be written in the component forms

$$m(\ddot{u}^i+y^j\ddot{\bar{\delta}}_j{}^i)=f^i+\sum_{K=1}^{M}n_K{}^i, \tag{7a}$$

$$m(y^i\ddot{u}^j+y^{in}\ddot{\bar{\delta}}_n{}^j)=l^{ij}+\sum_{K=1}^{M}m_K{}^{ij}-k^{ij}, \tag{7b}$$

$$k^{ij}=k^{ji}, \tag{7c}$$

where

$$\mathbf{u}=u^i\mathbf{D}_i=u_i\mathbf{D}^i,\ \ \bar{\delta}_i=\bar{\delta}_i{}^j\mathbf{D}_j=\bar{\delta}_{ij}\mathbf{D}^j \tag{8a,b}$$

$$\mathbf{f}=f^i\mathbf{D}_i,\ \ \mathbf{n}_K=n_K{}^i\mathbf{D}_i \tag{8c,d}$$

$$\mathbf{m}_K{}^i=m_K{}^{ij}\mathbf{D}_j,\ \ \mathbf{k}^i=k^{ij}\mathbf{D}_j \tag{8e, f}$$

and where the components $\bar{\delta}_i^j$ and $\bar{\delta}_{ij}$ of the displacements $\bar{\delta}_i$ should not be confused with the values of δ_i^j and δ_{ij} of the Kronecker delta symbol. Also, $\mathbf{D}^i$ are the reference values of $\mathbf{d}^i$.

It was shown in (Rubin, 1985a) that a one-to-one correspondence exists between the theory of a Cosserat point and the three-dimensional theory when we assume that the position vector $\mathbf{p}$ of an arbitrary material point, within the Cosserat point, admits the representation

$$\mathbf{p}=\mathbf{p}(\theta^i,t)=\mathbf{r}(t)+\theta^i\mathbf{d}_i(t) \tag{9}$$

where θ^i (i = 1, 2, 3) are convected coordinates identifying the material point. For the linear theory the three-dimensional displacement $\mathbf{u}^*(\theta^i,t)$ may be represented by

$$\mathbf{u}^*(\theta^i,t)=\mathbf{u}(t)+\theta^i\bar{\delta}_i. \tag{10}$$

To determine the vibration of the parallelepiped shown in Fig. 1 it is convenient to specify the reference values $\mathbf{D}_i$ of the directors in the forms

$$\mathbf{D}_1=L_1\mathbf{e}_1,\ \ \mathbf{D}_2=L_2\mathbf{e}_2,\ \ \mathbf{D}_3=L_3\mathbf{e}_3 \tag{11a,b,c}$$

where $\mathbf{e}_i$ are base vectors of a fixed rectangular Cartesian coordinate system with coordinates x_i and where L_1, L_2 and L_3 are the lengths of the sides of the parallelepiped. Furthermore, we specify the convected coordinates θ^i by

$$\theta^1=\frac{x_1}{L_1},\ \ \theta^2=\frac{x_2}{L_2},\ \ \theta^3=\frac{x_3}{L_3}. \tag{12a,b,c}$$

In the above, we have identified the midpoint of the parallelepiped as the location of the Cosserat point so the region of space occupied by the Cosserat point is defined by

$$|\theta^1|\leqq ½,\ \ |\theta^2|\leqq ½,\ \ |\theta^3|\leqq ½ \tag{13a,b,c}$$

From (Rubin, 1985a), we recall that for a parallelepiped that is composed of a homogeneous, linear elastic material that has a uniform mass density $\rho_0{}^*$ in the reference configuration, and is three-dimensionally isotropic, the constitutive equations for the mass m, inertia quantities y^i and y^{ij}, and the intrinsic director couple k^{ij} become[3]

$$m=\rho_0{}^*L_1L_2L_3,\ \ y^i=0, \tag{14a,b}$$

$$y^{11}=y^{22}=y^{33}=\frac{1}{12},\ \text{all other }y^{ij}=0, \tag{14c,d}$$

$$k^{ij}=\bar{c}_1(H^{mn}\gamma_{mn})H^{ij}+2\bar{c}_2H^{im}H^{jn}\gamma_{mn}, \tag{14e}$$

$$H^{11}=\frac{1}{L_1^2},\ \ H^{22}=\frac{1}{L_2^2},\ \ H^{33}=\frac{1}{L_3^2}, \tag{14f,g,h}$$

$$\text{all other }H^{ij}=0, \tag{14i}$$

---

[2] *Throughout this paper we use the usual summation convention over repeated indicies.*

[3] *See additional remarks in the closing section regarding the specification of y^{ij}.*

$$\bar{c}_1 = \frac{E\nu L_1 L_2 L_3}{(1+\nu)(1-2\nu)}, \quad \bar{c}_2 = \frac{E L_1 L_2 L_3}{2(1+\nu)}, \tag{14j,k}$$

where H^{ij} are the reference values of h^{ij} and (14e) is a generalized form of Hooke's law. For later reference (see the conclusions) we note that the constitutive constants $\bar{c}_1$ and $\bar{c}_2$ were determined in terms of Young's modulus E and Poisson's ratio ν of the three-dimensional material by comparing Cosserat solutions with exact solutions of certain simple static problems. Also, the condition (14d) states that the directors $\mathbf{D}_i$ are parallel to the principal axes of inertia.

For free vibration of the parallelepiped the body force $\mathbf{f}$ and director couples $\mathbf{l}^i$, and contact forces $\mathbf{n}_K$ and director couples $\mathbf{m}_K{}^i$ all vanish. Then with the help of (14) the equations of motion (7a,b) reduce to

$$\ddot{u}^i = 0, \quad m y^{in} \ddot{\bar{\delta}}_n{}^j = -k^{ij}. \tag{15a,b}$$

Using (7c) and (14c,d) equation (15b) may be rewritten in the equivalent form

$$\frac{m}{12}\ddot{\bar{\delta}}_1{}^1 = -k^{11}, \quad \frac{m}{12}\ddot{\bar{\delta}}_2{}^2 = -k^{22}, \quad \frac{m}{12}\ddot{\bar{\delta}}_3{}^3 = -k^{33}, \tag{16a,b}$$

$$\frac{m}{12}\ddot{\bar{\delta}}_1{}^2 = -k^{12}, \quad \frac{m}{12}\ddot{\bar{\delta}}_1{}^3 = -k^{13}, \quad \frac{m}{12}\ddot{\bar{\delta}}_2{}^3 = -k^{23}, \tag{16c,d}$$

and

$$\ddot{\bar{\delta}}_1{}^2 = \ddot{\bar{\delta}}_2{}^1, \quad \ddot{\bar{\delta}}_1{}^3 = \ddot{\bar{\delta}}_3{}^1, \quad \ddot{\bar{\delta}}_2{}^3 = \ddot{\bar{\delta}}_3{}^2. \tag{17a,b,c}$$

Equation (15a) implies that the velocity of the center of mass of the Cosserat point is constant and equations (17) imply that the angular momentum about the center of mass is constant.

Now the strain components γ_{ij} can be determined in terms of the components of the displacements $\bar{\delta}_i$ by substituting (6b) into the definition (4) and neglecting quadratic terms in $\bar{\delta}_i$ to obtain

$$\gamma_{ij} = ½(\bar{\delta}_{ij} + \bar{\delta}_{ji}). \tag{18}$$

Using the transformation relations

$$\bar{\delta}_i{}^j = H^{jn}\bar{\delta}_{in}, \tag{19}$$

equations (17) may be written in the form

$$\frac{\ddot{\bar{\delta}}_{12}}{L_2{}^2} = \frac{\ddot{\bar{\delta}}_{21}}{L_1{}^2}, \quad \frac{\ddot{\bar{\delta}}_{13}}{L_3{}^2} = \frac{\ddot{\bar{\delta}}_{31}}{L_1{}^2}, \quad \frac{\ddot{\bar{\delta}}_{23}}{L_3{}^2} = \frac{\ddot{\bar{\delta}}_{32}}{L_2{}^2}. \tag{20a,b,c}$$

It follows from (18)–(20) that

$$\ddot{\gamma}_{11} = L_1{}^2\ddot{\bar{\delta}}_1{}^1, \quad \ddot{\gamma}_{22} = L_2{}^2\ddot{\bar{\delta}}_2{}^2, \quad \ddot{\gamma}_{33} = L_3{}^2\ddot{\bar{\delta}}_3{}^3 \tag{21a,b,c}$$

$$\ddot{\gamma}_{12} = \left(\frac{L_1{}^2 + L_2{}^2}{2}\right)\ddot{\bar{\delta}}_1{}^2, \quad \ddot{\gamma}_{13} = \left(\frac{L_1{}^2 + L_3{}^2}{2}\right)\ddot{\bar{\delta}}_1{}^3, \tag{21d,e}$$

$$\ddot{\gamma}_{23} = \left(\frac{L_2{}^2 + L_3{}^2}{2}\right)\ddot{\bar{\delta}}_2{}^3. \tag{21f}$$

Thus, with the help of the constitutive equations (14), the results (21) and the definition of the physical components of strain $\bar{\gamma}_{ij}$

$$\bar{\gamma}_{11} = \frac{\gamma_{11}}{L_1{}^2}, \quad \bar{\gamma}_{22} = \frac{\gamma_{22}}{L_2{}^2}, \quad \bar{\gamma}_{33} = \frac{\gamma_{33}}{L_3{}^2}, \tag{22a,b,c}$$

$$\bar{\gamma}_{12} = \frac{\gamma_{12}}{L_1 L_2}, \quad \bar{\gamma}_{13} = \frac{\gamma_{13}}{L_1 L_3}, \quad \bar{\gamma}_{23} = \frac{\gamma_{23}}{L_2 L_3}, \tag{22d,e,f}$$

the equations of motion (16) may be rewritten in the form

$$\left(\frac{1}{p_1{}^2}\right)\ddot{\bar{\gamma}}_{11} + c_1{}^2\bar{\gamma}_{11} + (c_1{}^2 - 2c_2{}^2)\bar{\gamma}_{22} + (c_1{}^2 - 2c_2{}^2)\bar{\gamma}_{33} = 0, \tag{23a}$$

$$\left(\frac{1}{p_2{}^2}\right)\ddot{\bar{\gamma}}_{22} + (c_1{}^2 - 2c_2{}^2)\bar{\gamma}_{11} + c_1{}^2\bar{\gamma}_{22} + (c_1{}^2 - 2c_2{}^2)\bar{\gamma}_{33} = 0, \tag{23b}$$

$$\left(\frac{1}{p_3{}^2}\right)\ddot{\bar{\gamma}}_{33} + (c_1{}^2 - 2c_2{}^2)\bar{\gamma}_{11} + (c_1{}^2 - 2c_2{}^2)\bar{\gamma}_{22} + c_1{}^2\bar{\gamma}_{33} = 0, \tag{23c}$$

and

$$\ddot{\bar{\gamma}}_{12} + c_2{}^2(p_1{}^2 + p_2{}^2)\bar{\gamma}_{12} = 0, \tag{24a}$$

$$\ddot{\bar{\gamma}}_{13} + c_2{}^2(p_1{}^2 + p_3{}^2)\bar{\gamma}_{13} = 0, \tag{24b}$$

$$\ddot{\bar{\gamma}}_{23} + c_2{}^2(p_2{}^2 + p_3{}^2)\bar{\gamma}_{23} = 0, \tag{24c}$$

where c_1 and c_2 are the dilitational and equi-voluminal wave speeds, respectively, defined by

$$c_1{}^2 = \frac{E(1-\nu)}{\rho_0{}^*(1+\nu)(1-2\nu)}, \quad c_2{}^2 = \frac{E}{2\rho_0{}^*(1+\nu)}, \tag{25a,b}$$

and p_1, p_2 and p_3 are constants which are related to wave numbers and are defined by

$$p_1{}^2 = \frac{12}{L_1{}^2}, \quad p_2{}^2 = \frac{12}{L_2{}^2}, \quad p_3{}^2 = \frac{12}{L_3{}^2}. \tag{26a,b,c}$$

It is interesting to note that, within the context of the theory of a Cosserat point, equations (23), which describe extensional vibration, are uncoupled from equations (24), which describe shear vibration. Each of these types of vibration are described in the following sections.

Extensional Vibration

To determine the natural frequencies of extensional vibration of the parallelepiped we may assume a solution of the form

$$\bar{\gamma}_{11} = a_1 \sin\omega t, \quad \bar{\gamma}_{22} = a_2 \sin\omega t, \quad \bar{\gamma}_{33} = a_3 \sin\omega t, \tag{27a,b,c}$$

where the constants a_i and ω are, respectively, the mode shape and frequency, each to be determined. Now, with the help of (27) the system of equations (23) may be reduced to

$$(K_{ij} - \omega^2 M_{ij})a_j = 0 \tag{28}$$

where the symmetric tensors K_{ij} and M_{ij} are defined by

$$K_{11} = K_{22} = K_{33} = c_1{}^2, \tag{29a}$$

$$K_{12} = K_{21} = K_{13} = K_{31} = K_{23} = K_{32} = (c_1{}^2 - 2c_2{}^2), \tag{29b}$$

$$M_{11} = \frac{1}{p_1{}^2}, \quad M_{22} = \frac{1}{p_2{}^2}, \quad M_{33} = \frac{1}{p_3{}^2}, \tag{29c,d,e}$$

$$\text{all other } M_{ij} = 0. \tag{29f}$$

Following usual procedures, the characteristic equation of the system (28) may be determined by setting the determinant of the tensor $(K_{ij} - \omega^2 M_{ij})$ equal to zero. After some algebraic manipulation, this characteristic equation may be written in the form

$$G(\omega, p_1, p_2, p_3) = 0, \tag{30a}$$

$$\begin{aligned} G = {} & \omega^6 - c_1{}^2(p_1{}^2 + p_2{}^2 + p_3{}^2)\omega^4 \\ & + 4(c_1{}^2 - c_2{}^2)c_2{}^2(p_1{}^2p_2{}^2 + p_1{}^2p_3{}^2 + p_2{}^2p_3{}^2)\omega^2 \\ & - 4(3c_1{}^2 - 4c_2{}^2)c_2{}^4 p_1{}^2p_2{}^2p_3{}^2. \end{aligned} \tag{30b}$$

Given the dimensions of the parallelepiped L_1, L_2, L_3 and the constitutive properties of the material in terms of the wave speeds c_1 and c_2, the characteristic equation (30a) may be solved for three natural frequencies of extensional vibration. The mode shape associated with each of these frequencies can then be determined by solving (28) for a_i using the appropriate value of ω. It is expected that the lowest of these frequencies should provide a good approximation to the lowest frequency of vibration of the parallelepiped associated with the particular mode shape considered.

A few special cases of interest may be analyzed simply if we consider a parallelepiped which has a square cross section. For such a parallelepiped

$$L_1 = L_2, \quad p_1 = p_2, \tag{31a,b}$$

and the solution of (30*a*) may be written in the form

$$\omega^2 = 2c_2{}^2 p_1{}^2, \quad \omega^2 = \frac{b - [b^2 - 4c]^{1/2}}{2}, \quad \omega^2 = \frac{b + [b^2 - 4c]^{1/2}}{2} \tag{32a,b,c}$$

$$b = 2(c_1{}^2 - c_2{}^2)p_1{}^2 + c_1{}^2 p_3{}^2, \tag{32d}$$

$$c = 2(3c_1{}^2 - 4c_2{}^2)c_2{}^2 p_1{}^2 p_3{}^2. \tag{32e}$$

In the limit that the parallelepiped becomes a cube, i.e.,

$$L_1 = L_2 = L_3, \quad p_1 = p_2 = p_3 \tag{33a,b}$$

the solution (32) takes the form

$$\omega^2 = 2c_2{}^2 p_1{}^2, \quad \omega^2 = 2c_2{}^2 p_1{}^2, \quad \omega^2 = (3c_1{}^2 - 4c_2{}^2)p_1{}^2. \tag{34a,b,c}$$

This result compares well with the exact solution presented in (Hutchinson and Zillmer, 1983).[4] In particular, the exact solution for a cube of length L_1 and Poisson's ratio $\nu = 0.3$ predicts that the lowest two frequencies are equal and take the value $\omega^* = 4.44\ c_2/L_1$. Comparing this value with the one predicted by (34*a,b*) we observe that $\omega/\omega^* = 1.103$. Thus, the Cosserat theory predicts a value about 10 percent higher than the exact value. However, the value of the highest frequency (34*c*) which is expected to be in error, is about twice the exact value.

To further examine the accuracy of the natural frequencies predicted by the theory of a Cosserat point, we consider the limiting cases when the parallelepiped becomes a thin rod or a thin plate, each with square cross section. Although one might expect the theory of a Cosserat point to be limited to the description of a body whose length dimensions are comparable, we will show that the Cosserat theory predicts good results for both of these limiting cases.

For the limiting case of a bar which is long in the $\mathbf{e}_3$ direction we have $p_3 << p_1$ and the solution (32) becomes

$$\omega^2 = \omega_b{}^2 = \frac{(3c_1{}^2 - 4c_2{}^2)c_2{}^2}{(c_1{}^2 - c_2{}^2)} p_3{}^2 = c_b{}^2 p_3{}^2, \tag{35a}$$

$$\omega^2 = 2c_2{}^2 p_1{}^2, \quad \omega^2 = 2(c_1{}^2 - c_2{}^2)p_1{}^2, \tag{35b,c}$$

where ω_b is the lowest frequency of the bar and where $c_b = (E/\rho_0{}^*)^{1/2}$ is the bar wave speed. The expression (35*a*) for the lowest frequency has the same form as that predicted by elementary thin rod theory except that the wave number p_3 is replaced by a different value $p_3{}^*$. More specifically, we recall from (Graff, 1975, p. 87)[5] that the frequency $\omega_b{}^*$ and wave number $p_3{}^*$ associated with the lowest extensional mode of vibration of the thin rod are given by

$$(\omega_b{}^*)^2 = c_b{}^2 (p_3{}^*)^2, \quad p_3{}^* = \frac{\pi}{L_3}. \tag{36a,b}$$

It follows from (26*c*), (35*a*) and (36) that

$$\frac{\omega_b}{\omega_b{}^*} = \frac{(12)^{1/2}}{\pi} = 1.103, \tag{37}$$

which states that the theory of a Cosserat point predicts a natural frequency that has the same functional dependence on the length L_3 of the bar as that predicted by thin rod theory, but has a value increased by about 10 percent.

For the limiting case of a plate which is thin in the $\mathbf{e}_3$ direction we have $p_3 >> p_1$ and the solution (32) becomes

$$\omega^2 = \omega_p{}^2 = 2c_2{}^2 p_1{}^2, \quad \omega^2 = \frac{2(3c_1{}^2 - 4c_2{}^2)c_2{}^2}{c_1{}^2} p_1{}^2, \tag{38a,b}$$

$$\omega^2 = c_1{}^2 p_3{}^2. \tag{38c}$$

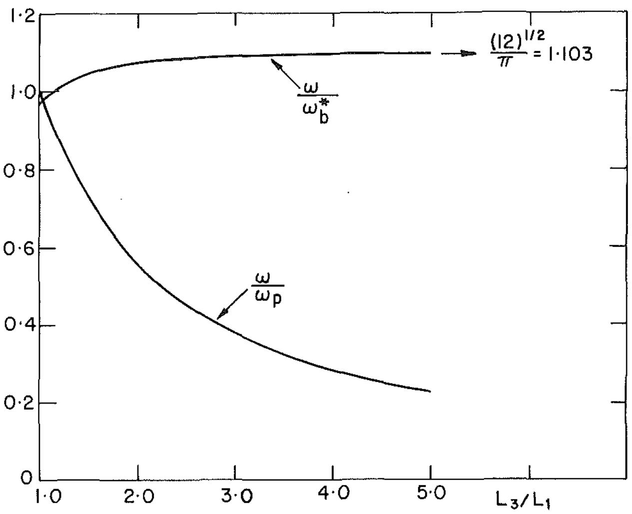


Fig. 2 Lowest extensional frequency ω for a square bar normalized by the variable bar frequency ω_b^* and the constant plate frequency ω_p

The expression (38*a*) has the same form as the exact solution (Appendix A) except that the wave number p_1 is replaced by a different value $p_1{}^*$. More specifically, we recall from Appendix A that the frequency $\omega_p{}^*$ and wave number $p_1{}^*$ associated with the lowest extensional mode of vibration of the plate[6] is given by

$$(\omega_p{}^*)^2 = 2c_2{}^2 (p_1{}^*)^2, \quad p_1{}^* = \frac{\pi}{L_1}. \tag{39a,b}$$

It follows from (26*a*), (38*a*) and (39) that

$$\frac{\omega_p}{\omega_p{}^*} = \frac{(12)^{1/2}}{\pi} = 1.103, \tag{40}$$

which states that the theory of a Cosserat point predicts a natural frequency that has the same functional dependence on the width dimension L_1 of the bar as the exact solution but has a value increased by about 10 percent. Here, it is worth emphasizing that the plate frequency ω_p (given by (38*a*)) is a *natural frequency of vibration of a bar of square cross section* for any length to width ratio L_3/L_1 (see (32*a*)). This result is consistent with the exact solution in Appendix A.

To investigate the detailed character of the lowest frequency of vibration we consider a parallelepiped with unit dimensions ($L_1 = L_2 = 1$) and with Poisson's ratio $\nu = 0.3$. Specifically, we are interested in the transitional character as the length to width ratio L_3/L_1 varies from that associated with a cube ($L_3/L_1 = 1$) to that associated with thin rod ($L_3/L_1 \to \infty$).

Figure 2 exhibits the transitional character of the lowest extensional frequency ω associated with a square bar (L_3 varies with L_1 held constant). The values of ω plotted in Fig. 2 are obtained using the expression (32) and the value $\nu = 0.3$ for Poisson's ratio. One curve in Fig. 2 shows the value of ω normalized by the variable frequency $\omega_b{}^*$ (given by (36*a*)) predicted by thin rod theory. The other curve shows the value of ω normalized by the constant frequency ω_p (given by (38*a*)) predicted by the Cosserat theory for a bar with constant square cross-section. The curve for ω/ω_p has the same character as that associated with the exact solution presented in the actual[7] Fig. 4 in (Hutchinson and Zillmer, 1983).

The curve $\omega/\omega_b{}^*$ in Fig. 2 shows that the frequency very rapidly approaches the value ω_b (given by (35*a*)) predicted by the Cosserat theory for a thin rod. This result suggests that the exact lowest frequency should be very closely approximated by the value $\omega_b{}^*$ predicted by thin rod theory even when the

---

[4] In a personal correspondence, Professor Hutchinson clarified that the printers placed the plots for Figs. 4 and 5 above the captions for Figs. 6 and 7, and vice versa.

[5] Our notation c_b, $\omega_b{}^*$, $p_3{}^*$ correspond to c_0, ω_1, and π/l, respectively, in (Graff, 1975).

[6] See the comment at the end of Appendix A.

[7] See footnote number 4.

length to width ratio L_3/L_1 is at low as about three. In this regard, it is not too surprising that the extensional (longitudinal) frequencies of vibration measured in the experiments of (Spinner et al., 1960) for the rectangular bar specimens are accurately predicted by elementary thin rod theory.

The transitional character of the lowest extensional frequency as the length to width ratio L_3/L_1 varies from that associated with a cube ($L_3/L_1 = 1$) to that associated with a thin plate ($L_3/L_1 \to 0$) was also investigated. Again the solution (32) was evaluated using the value $\nu = 0.3$ for Poisson's ratio. It was found that the plate frequency ω_p (given by (38*a*)) was the lowest frequency for the whole range of L_3/L_1 ($0 < L_3/L_1 \leqq 1$).

Shear Vibration

Equations (24) which describe shear vibration of the parallelepiped are uncoupled so that shear vibration can occur in the $\mathbf{e}_1 - \mathbf{e}_2$ plane, the $\mathbf{e}_1 - \mathbf{e}_3$ plane and the $\mathbf{e}_2 - \mathbf{e}_3$ plane, independently. The natural frequencies associated with shear vibration in each of these planes may be recorded in the form:

$$\omega^2 = c_2{}^2(p_1{}^2 + p_2{}^2) \text{ for } \mathbf{e}_1 - \mathbf{e}_2 \text{ plane,} \tag{41a}$$

$$\omega^2 = c_2{}^2(p_1{}^2 + p_3{}^2) \text{ for } \mathbf{e}_1 - \mathbf{e}_3 \text{ plane,} \tag{41b}$$

$$\omega^2 = c_2{}^2(p_2{}^2 + p_3{}^2) \text{ for } \mathbf{e}_2 - \mathbf{e}_3 \text{ plane,} \tag{41c}$$

At present, it does not appear that an exact solution exists with which these results can be compared.

Conclusions

A rectangular parallelepiped composed of a homogeneous linear elastic isotropic material has been modeled as an isotropic Cosserat point. For linear motions of a free parallelepiped the twelve equations of motion of the theory of a Cosserat point separate into four groups of three equations each. One group of equations states that the velocity of the center of mass vanishes while another group states that the angular momentum about the center of mass is constant. The remaining two groups describe, individually, extensional vibration and shear vibration.

The characteristic equation determining the three fundamental frequencies of extensional vibration was obtained for a parallelepiped with arbitrary major dimensions L_1, L_2, L_3 (see Fig. 1). The lowest natural frequency was analyzed in detail for the special case of a bar with square cross-section ($L_1 = L_2$). Both the limiting cases of a thin rod ($L_3/L_1 \to \infty$) and a thin plate ($L_3/L_1 \to 0$) were analyzed. It was shown that in the limit of a thin rod the Cosserat theory predicts a frequency with the same functional dependence on the length dimension L_3 as that predicted by thin rod theory but with a value increased by about 10 percent ($(12)^{1/2}/\pi$). Similarly, in the limit of a thin plate the Cosserat theory predicts a frequency with the same functional dependence on the width dimension L_1 as the exact solution but with a value increased by about 10 percent ($(12)^{1/2}/\pi$). Also, the value for the lowest frequency of a cube was shown to be about 10 percent (1.103) higher than the exact value.

The transitional character of the lowest extensional frequency was investigated as the length to width ratio L_3/L_1 varied (with L_1 constant). It was found that the lowest frequency rapidly approached the limiting value associated with a thin rod as the ratio L_3/L_1 varied from that associated with a cube ($L_3/L_1 = 1$) to that associated with a thin rod ($L_3/L_1 \to \infty$). This transitional character would suggest that the lowest extensional frequency is closely approximated by the value predicted by elementary thin rod theory even when the length to width ratio L_3/L_1 is as low as about three. Furthermore, the lowest extensional frequency was shown to be constant when the ratio L_3/L_1 varied (with L_1 constant) from that associated with a cube ($L_3/L_1 = 1$) to that associated with a thin plate ($L_3/L_1 \to 0$).

The lowest frequencies associated with shear vibration in the $\mathbf{e}_1 - \mathbf{e}_2$ plane, the $\mathbf{e}_1 - \mathbf{e}_3$ plane and the $\mathbf{e}_2 - \mathbf{e}_3$ plane were also determined using the Cosserat theory.

It is very interesting to note that the values predicted by the Cosserat theory for the lowest extensional frequency of vibration of a thin plate, a cube, and a thin rod are all increased by the same factor ($(12)^{1/2}/\pi = 1.103$). This suggests that it is possible to modify the Cosserat theory to obtain correct values for each of these limiting cases. In particular, we note that within the framework of the Cosserat theory the inertia quantities m, y^i, y^{ij} in (5) are not a priori determined but rather must be determined by constitutive equations. Such constitutive equations can be obtained by comparing with exact solutions or experiments. The values (14*a*-*d*) for these quantities were obtained by integrating expressions relating these quantites to three-dimensional quantities. Here, we retain the specifications (14*a*,*b*,*d*), and assume that the quantities y^{11}, y^{22}, y^{33} are equal, but we determine their value by comparing with a known solution. This is equivalent to specifying the wave numbers p_1, p_2, p_3 by

$$p_1{}^2 = \frac{1}{y^{11}L_1{}^2},\quad p_2{}^2 = \frac{1}{y^{11}L_2{}^2},\quad p_3{}^2 = \frac{1}{y^{11}L_3{}^2}, \tag{42a,b,c}$$

instead of by equations (26*a*,*b*,*c*). The value of y^{11} is determined by requiring the Cosserat prediction of the frequency (35*a*) for the limiting case of a thin rod to be accurate. Thus, from (35*a*), (36) and (42*c*) we obtain

$$y^{11} = y^{22} = y^{33} = \frac{1}{\pi^2}. \tag{43}$$

Now, using the specifications (42) and (43) it is easy to show that the Cosserat theory also predicts the correct extensional frequency for the limiting cases of a thin plate and a cube. Furthermore, with the help of (42), (43), the Cosserat prediction of the lowest shear frequency (41) is modified.

Finally, we note that the above procedure should be contrasted with the one suggested by Mindlin (1951). Specifically, Mindlin suggests that one of the material constants characterizing the strength of the material be modified instead of the inertia properties of the material. Within the context of the Cosserat theory this could be done by replacing $\bar{c}_1$ and $\bar{c}_2$ in (14*j*,*k*) by values which are scaled by the same factor. This scale factor could be determined by matching the frequency predicted in the limit of a thin rod. However, such a procedure would have the undesirable effect of causing errors in the predictions of the static problems of simple shear and simple tension. For this reason, we prefer to modify the specification of the director inertia instead of the material constants $\bar{c}_1$ and $\bar{c}_2$.

References

Graff, K. F., 1975, *Wave Motion in Elastic Solids*, Ohio State University Press.

Green, A. E., Naghdi, P. M., and Wenner, M. L., 1974a, "On the Theory of Rods—I. Derivations from the Three-Dimensional Equations," *Proc. R. Soc. Lond.* A337, pp. 451-483.

Green, A. E., Naghdi, P. M., and Wenner, M. L., 1974b, "On the Theory of Rods—II. Developments by Direct Approach," *Proc. R. Soc. Lond.* A337, pp. 485-507.

Hutchinson, J. R., and Zillmer, S. D., 1983, "Vibration of a Free Rectangular Parallelepiped," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 123-130.

Mindlin, R. D., 1951, "Influence of Rotary Inertia and Shear on Flexural Motions of Isotropic, Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 18, pp. 31-38.

Naghdi, P. M., 1972, "The Theory of Shells and Plates," Flügges *Handbuch der Physik*, Vol. VI a/2, Truesdell, C., ed., pp. 425-640.

Rubin, M. B., 1985a, "On the Theory of a Cosserat Point and Its Application to the Numerical Solution of Continuum Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 368-372.

Rubin, M. B., 1985b, "On the Numerical Solution of One-Dimensional Continuum Problems Using the Theory of a Cosserat Point," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 373-378.

Spinner, S., Reichard, T. W., and Tefft, W. E., 1960, "Comparison of Experimental and Theoretical Relations Between Young's Modulus and the Flexural and Longitudinal Resonance Frequencies of Uniform Bars," *Journal of Research*, NBS, Vol. 64A, pp. 147-155.

APPENDIX A

In this appendix we record a simple exact solution for extensional vibration of a parallelepiped with square cross-section ($L_1 = L_2$), which initially occupies the region of space

$$|x_1| \leqq \frac{L_1}{2}, \quad |x_2| \leqq \frac{L_1}{2}, \quad |x_3| \leqq \frac{L_3}{2}, \qquad (A1a,b,c)$$

The three-dimensional equations of motion of the linearized theory in the absence of body force are

$$\rho_0{}^* \ddot{u}_i{}^* = t_{ij,j}, \quad t_{ij} = t_{ji}, \qquad (A2a,b)$$

where $\rho_0{}^*$ is the mass density, $u_i{}^*$ is the displacement, t_{ij} is the stress, and where a comma denotes partial differentiation with respect to the coordinates x_i. The constitutive equations of a linear elastic isotropic material and the strain-displacement relations may be written in the forms

$$t_{ij} = \lambda\, e_{mm} \delta_{ij} + 2\mu\, e_{ij}, \qquad (A3a,b)$$

$$e_{ij} = \tfrac{1}{2}(u_{i,j}^* + u_{j,i}^*), \qquad (A3b)$$

where δ_{ij} is the Kronecker delta symbol. Furthermore, for free vibration of the parallelepiped, we impose the boundary conditions

$$t_{i1}\left(\pm\frac{L_1}{2},\ x_2,\ x_3,\ t\right) = 0, \qquad (A4a)$$

$$t_{i2}\left(x_1,\ \pm\frac{L_1}{2},\ x_3,\ t\right) = 0, \qquad (A4b)$$

$$t_{i3}\left(x_1,\ x_2,\ \pm\frac{L_3}{2},\ t\right) = 0. \qquad (A4c)$$

To describe extensional vibration in the $\mathbf{e}_1 - \mathbf{e}_2$ plane the displacements $u_i{}^*$ are specified by

$$u_1{}^* = a_1{}^* \sin \omega^* t \sin p_1{}^* x_1 \cos p_1{}^* x_2, \qquad (A5a)$$

$$u_2{}^* = -a_1{}^* \sin \omega^* t \cos p_1{}^* x_1 \sin p_1{}^* x_2 \qquad (A5b)$$

$$u_3{}^* = 0, \qquad (A5c)$$

where $a_1{}^*$, ω^*, $p_1{}^*$ are constants. These displacements characterize extensional vibration because $u_1{}^*$, $u_2{}^*$ are odd functions of x_1, x_2 and even functions of x_2, x_1, respectively. Substituting ($A5$) into ($A3b$) the expressions for the stress become

$$t_{11} = -t_{22} = 2\mu\, a_1{}^* p_1{}^* \sin \omega^* t \cos p_1{}^* x_1 \cos p_1{}^* x_2, \qquad (A6a)$$

$$\text{all other } t_{ij} = 0. \qquad (A6b)$$

Now, with the help of ($A5$) and ($A6$), the equations of motion ($A2a$) yield the dispersion relation ($39a$) and the boundary conditions ($A4$) determine the wave number $p_1{}^*$ through the equations

$$\cos\frac{p_1{}^* L_1}{2} = 0, \quad p_1{}^* = \frac{(2n-1)\pi}{L_1}\ n = 1,2,\ldots \qquad (A7a,b)$$

It follows that the lowest frequency $\omega^* = \omega_p{}^*$ (associated with $n = 1$) may be written in the form ($39a$) with $p_1{}^*$ given by ($39b$). Although $\omega_p{}^*$ in ($39a$) is the lowest extensional frequency of the type ($A5$) it is not necessarily the lowest extensional frequency of the square bar. For example, in the limit that the bar becomes a thin rod ($L_3/L_1 >> 1$) the lowest extensional frequency is the bar frequency $\omega_b{}^*$ given by ($36a$). On the other hand, in the limit that the bar becomes a thin plate ($L_3/L_1 << 1$) the lowest extensional frequency is the plate frequency $\omega_p{}^*$ given by (39).

ADDENDUM

The author has recently learned of some related work by other authors. Cohen (1981) developed the theory of pseudo-rigid bodies by a direct approach for bodies with a fixed point. Similar equations were developed by Muncaster (1984) as special solutions of the three-dimensional equations. Both of these approaches are different from that presented by Rubin (1985a). Also, Cohen and Muncaster (1984) have considered small strain-free vibrations of a symmetric body which are less general than those considered in this paper.

Additional References

Cohen, H., 1981, "Pseudo-Rigid Bodies," *Utilitas Mathematica*, Vol. 20, pp. 221-247.

Muncastor, R. G., 1984, "Invariant Manifolds in Mechanics II: Zero-Dimensional Elastic Bodies With Directors," *Arch. Rat. Mech. Anal.*, Vol. 84, pp. 375-392.

Cohen, H., and Muncaster, R. G., 1984, "The Dynamics of Pseudo-Rigid Bodies: General Structure and Exact Solutions," *J. of Elasticity*, pp. 127-154.

J. C. Simo[1]
Assistant Professor.

T. J. R. Hughes
Professor and Chairman of
Applied Mechanics.

Division of Applied Mechanics,
Stanford University,
Stanford, Calif. 94305

On the Variational Foundations of Assumed Strain Methods

So-called assumed strain methods *are based on the a-priori assumption of an interpolation for the discrete gradient operator, not necessarily derivable from the displacement interpolation. It is shown that this class of methods falls within the class of variational methods based on the Hu-Washizu principle. The essential point of this equivalence lies in the statement of the appropriate* stress recovery *procedure compatible with this variational structure. It is noted that most currently existing assumed strain methods fail to perform the stress recovery in a manner consistent with the variational structure discussed herein. Application is made to recently proposed methods such as mode decomposition techniques.*

1 Introduction

The denomination "assumed strain methods" is intended to encompass a variety of finite element procedures, often proposed on an ad-hoc basis, which are typically characterized by an interpolation of the discrete gradient operator assumed *a-priori*, independently of the interpolation adopted for the displacement field. The often referred to "B-bar procedure", proposed by Hughes [1], offers an example of an assumed strain method which has proven successful in a variety of situations, including widely used structural elements [3]. For the finite strain incompressible problem, this method has recently been precisely reformulated by Simo et al. [6] within the context of the Hu-Washizu principle. The so-called *mode decomposition technique*, proposed by Belytschko and coworkers (e.g., [1, 7]), furnishes another example of a B-bar type of method that leads to the formulation of successful structural elements.

The purpose of this paper is to show that *assumed strain methods* can be systematically formulated within the variational framework furnished by the Hu-Washizu principle. A crucial point in this development concerns the role played by the stress field, now entering the formulation as a Lagrange multiplier, and its recovery within the proposed variational structure.

It is first noted that the Lagrange multipliers drop out from the formulation leading to a *generalized displacement* model, provided a certain *orthogonality condition* on the *assumed strain field* is satisfied. In addition, as a result of the variational structure, the *admissible stress field* (Lagrange multipliers) is constrained by an *orthogonality condition* arising from the Hu-Washizu principle as an Euler-Lagrange equation. These orthogonality conditions result in a single *constraint* that, for a given form of assumed strain field, restricts the form of admissible stress fields. Therefore, once an assumed strain method has been formulated, the stress recovery becomes the central issue. It is interesting to observe that almost all of the stress recovery procedures employed in the assumed strain methods we are aware of fail to meet this condition, thus rendering the resulting method "nonvariational."

It should also be emphasized that the generality afforded by the Hu-Washizu principle is essential for the variational characterization of assumed strain methods discussed below. In this sense, the Hellinger-Reissner principle provides too narrow a variational framework in which this class of methods generally do not fit. In fact within this framework, in a recent paper [7], Stolarski and Belytschko conclude that the mode decomposition technique, a particular instance of a B-bar procedure, generally lacks a variational justification. However, as a particular example, we show below how to obtain this class of methods within the variational framework outlined here, by appropriately recovering the stress field.

For the sake of concreteness, we shall adopt the bending of a Mindlin plate as a model problem in the context of which our discussion is presented. The same ideas apply without modification to other problems such as three-dimensional elasticity.

2 Mindlin Plate Problem. Basic Notation.

Consider a Mindlin plate with mid-plane spanning $\Omega \subset IR^2$ and thickness h. Denote by $x \in \Omega$ a typical point of the mid-plane. Define the set of generalized displacements by

$$x \in \Omega \rightarrow u(x) := \left\{ \begin{array}{c} \vartheta(x) \\ w(x) \end{array} \right\} \in IR^3 \tag{1}$$

Here, $\vartheta(x)$ is the (infinitesimal) rotation of a line (director), perpendicular to Ω through x, and $w(x)$ the vertical deflection at $x \in \Omega$. Introduce the following differential operators

---

[1] Formerly post-doctoral research engineer at U.C. Berkeley.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 25, 1985; final revision, September 10, 1985.

$$L_B := \begin{bmatrix} \dfrac{\partial}{\partial x^1} & 0 & 0 \\ 0 & \dfrac{\partial}{\partial x^2} & 0 \\ \dfrac{\partial}{\partial x^2} & \dfrac{\partial}{\partial x^1} & 0 \end{bmatrix},$$

$$L_S := \begin{bmatrix} -1 & 0 & \dfrac{\partial}{\partial x^1} \\ 0 & -1 & \dfrac{\partial}{\partial x^2} \end{bmatrix}, \tag{2}$$

so that the *curvature*, $\kappa(x)$ and *shear strain*, $\gamma(x)$ at $x \in \Omega$ are given by

$$\kappa := L_B u, \quad \gamma := L_S u \tag{3}$$

We assume an *uncoupled* bending/shear stored energy function of the form

$$\Psi(\kappa,\gamma) = \Psi_B(\kappa) + \Psi_S(\gamma), \tag{4}$$

so that the bending moment m and shear force s are given by

$$m = \frac{\partial \Psi_B(\kappa)}{\partial \kappa}, \quad s = \frac{\partial \Psi_S(\gamma)}{\partial \gamma} \tag{5}$$

We note that for the developments that follow it is not necessary to assume that contitutive equations (5) are linear.

3 Variational Structure. Hu-Washizu Principle.

Consider the following partial Hu-Washizu principle[2]

$$\Pi(u,\gamma,s) := \int_\Omega [\Psi_B(L_B u) + \Psi_S(\gamma) + s \cdot (L_S u - \gamma)] d\Omega + \Pi_{\text{EXT}} \tag{6}$$

where Π_{EXT} is the total potential energy of the *external loading*. The space of *kinematically admissible variations* (*generalized displacements*) V may be defined as

$$V := \left\{ \eta \equiv (\vartheta, \eta) \in [H^1(\Omega)]^3 \mid \quad \eta \Big|_{\partial\Omega_u} = 0 \right\} \tag{7}$$

where $H^1(\Omega)$ denotes the usual Sobolev space of functions possessing finite energy, and $\partial\Omega_u$ is that part of the boundary where the generalized displacements are prescribed. In addition, *admissible shear strain* and *shear stress* fields are defined according to

$$\begin{aligned} \Gamma &:= \{\gamma : \Omega \to I\!R^2 \mid \quad \gamma \in [L^2(\Omega)]^2\} \\ S &:= \{s : \Omega \to I\!R^2 \mid \quad s \in [L^2(\Omega)]^2\} \end{aligned} \tag{8}$$

where $L^2(\Omega)$ denotes the usual Hilbert space of square integrable functions. By taking variations in the standard manner, we obtain the following *Euler-Lagrange* equations

$$G^1(u,s) := \int_\Omega \left(\frac{\partial \Psi_B}{\partial \kappa} : L_B \eta + s \cdot L_S \eta \right) d\Omega - G_{\text{EXT}} = 0 \qquad (\textit{equilibrium})$$

$$G^2(u,\gamma) := \int_\Omega \tau \cdot [L_S u - \gamma] d\Omega = 0 \quad (\textit{strain-displacement}) \tag{9}$$

$$G^3(\gamma,s) := \int_\Omega \epsilon \cdot \left[-s + \frac{\partial \Psi_S(\gamma)}{\partial \gamma} \right] d\Omega = 0 \qquad (\textit{constitutive equation})$$

The above equations hold for *arbitrary variations* $\eta \in V$, $\epsilon \in \Gamma$, and $\tau \in S$.

Approximation. Introduce a standard finite element discretization $\Omega = \cup_{e=1}^{E} \Omega_e$, and the finite dimensional approximating subspace

$$V^h := \left\{ u^h \in [C^0(\Omega)]^2 \mid \quad u^h \Big|_{\Omega_e} = \sum_{I=1}^{N_e} N^I(x) u_I \equiv N_e q_e \right\} \tag{10}$$

Here, $q_e \equiv \{u_I\}_{I=1,N_e}$ is the complete set of degrees of freedom of Ω_e. The discrete curvature and shear strain field corresponding to (10) are given by

$$L_B u^h \Big|_{\Omega_e} = B_B q_e, \text{ and } L_S u^h \Big|_{\Omega_e} = B_S q_e \tag{11}$$

Then, the discrete approximation to the weak forms (9) are

$$G_e^1 := p_e \cdot \left\{ \int_{\Omega_e} \left[B_B^T \frac{\partial \Psi_B}{\partial \kappa} \left(\kappa \Big|_{\Omega_e} \right) + B_S^T s \Big|_{\Omega_e} \right] d\Omega \right\} - G_{\text{EXT}} \Big|_{\Omega_e}$$

$$G_e^2 := \int_{\Omega_e} \tau \Big|_{\Omega_e} \cdot \left[B_S q_e - \gamma \Big|_{\Omega_e} \right] d\Omega \tag{12}$$

$$G_e^3 := \int_{\Omega_e} \epsilon \Big|_{\Omega_e} \cdot \left[-s \Big|_{\Omega_e} + \frac{\partial \Psi_S}{\partial \gamma} \left(\gamma \Big|_{\Omega_e} \right) \right] d\Omega$$

where $G^1 = \Sigma_{e=1}^{E} G_e^1$. Consider now an *assumed local strain field* generated by a *given B*-bar operator, $x \to \gamma^h |_{\Omega_e}(x) = \bar{B}_S(x) q_e$. That is, we assume (discontinuous) approximations for the strain field of the form

$$\begin{aligned} \Gamma^h := \{ \gamma^h \in [L^2(\Omega)]^2 \mid \gamma_e^h(x) &\equiv \gamma^h \Big|_{\Omega_e} (x) \\ &= \bar{B}_S(x) q_e, \text{ where } q_e \in I\!R^{N_e} \} \end{aligned} \tag{13}$$

Let us denote by $S^h \subset [L^2(\Omega)]^2$ the space of *discrete admissible stress fields* τ^h. Notice that in general $\Gamma^h \neq S^h$. In fact, as shown below and further elaborated upon in Section 4, this is the essential point in the variational stress recovery. By substitution of (13) into $(12)_3$, we obtain

$$G_e^3 := p_e \cdot \left(\int_{\Omega_e} \bar{B}_S^T \left[-s^h \Big|_{\Omega_e} + \frac{\partial \Psi_s}{\partial \gamma} (\gamma_e^h) \right] d\Omega \right) = 0 \tag{14}$$

for arbitrary variations $\epsilon^h |_{\Omega_e} = \bar{B}_S p_e \in \Gamma^h$. It will be shown next that the stress field s^h does not enter explicitly in the formulation provided certain orthogonality conditions are satisfied.

Orthogonality Constraints on the Stress Field. First, we observe that by making use of the identity $B_S = \bar{B}_S + [B_S - B_S]$ along with variational equation (14), the second term in the weak form of the momentum balance $(12.1)_1$ may be rewritten for arbitrary variations $\eta^h = N_e p_e \in V^h$, as

$$\begin{aligned} p_e \cdot \int_{\Omega_e} B_S^T s^h \Big|_{\Omega_e} d\Omega = p_e \cdot \int_{\Omega_e} \bar{B}_S^T \frac{\partial \Psi_S}{\partial \gamma} (\gamma_e^h) d\Omega \\ + p_e \cdot \int_{\Omega_e} [B_S - \bar{B}_S]^T s^h \Big|_{\Omega_e} d\Omega, \end{aligned} \tag{15}$$

Now assume that the second term in (15) vanishes. This condition along with the weak form $(12)_2$, yields the following two orthogonality relations

$$\left\{ \begin{array}{l} G_e^2 := \left\{ \int_{\Omega_e} \tau^h \Big|_{\Omega_e} \cdot [B_S - \bar{B}_S] d\Omega \right\} q_e = 0 \text{ for any } \tau^h \in S^h \\ \\ p_e \cdot \int_{\Omega_e} [B_S - \bar{B}_S]^T s^h \Big|_{\Omega_e} d\Omega = 0 \text{ for any } \eta^h = N_e p_e \in V^h \end{array} \right\} \tag{16}$$

where use has been made of (13). Note that for given $\bar{B}_S$, con-

---

[2] Independent stress and strain fields are only assumed for the transverse shear strain and stress fields. This presupposes no real loss of generality and is sufficient for the applications of interest [3].

ditions (16) place a restriction on the choice of admissible shear stress fields $\tau^h(x)$. In the next section, however, it is shown that *stress fields satisfying conditions* (16) *can always be constructed*. Substitution of (15) into $(12)_1$ yields, on account of $(16)_2$, the following *generalized* displacement model

$$\left\{\begin{aligned} G_e^1 &= p_e \bullet f_e - G_{\text{EXT}}\Big|_{\Omega_e} \\ &\equiv p_e \bullet \left(\int_{\Omega_e}\left[B_B^T \frac{\partial\Psi_B}{\partial\kappa}(\kappa_e^h) + \bar{B}_S^T \frac{\partial\Psi_S}{\partial\gamma}(\gamma_e^h)\right]d\Omega\right) - G_{\text{EXT}}\Big|_{\Omega_e}, \end{aligned}\right\} \tag{17}$$

where $\kappa_e^h = B_B q_e$ and $\gamma_e^h = \bar{B}_S q_e$. In addition, it readily follows from (17) that the *element tangent stiffness matrix*, K_e is given by

$$\left\{\begin{aligned} DG_e^1 \bullet u &= p_e \bullet K_e q_e \\ &\equiv p_e \bullet \left(\int_{\Omega_e}\left[B_B^T \frac{\partial^2\Psi_B(\kappa_e^h)}{\partial\kappa^2} B_B + \bar{B}_S^T \frac{\partial^2\Psi_S(\gamma_e^h)}{\partial\gamma^2} \bar{B}_S\right]d\Omega\right) q_e \end{aligned}\right\} \tag{18}$$

Therefore, the resulting discrete equations (17) and (18) are expressed exclusively in terms of B_B and $\bar{B}_S$.

Remark 3.1. The *assumed local strain field* defined by $\bar{B}_S$ is arbitrary. Formally, equations analogous to (17) and (18) would also arise in the context of the procedure proposed by Hughes [2] and are often referred to as the *B*-bar method. The present development furnishes the appropriate variational framework.

Remark 3.2. Note that the Lagrange multipliers *drop out* completely from the formulation, provided conditions (16) hold. This fact was also noted in Simo, Taylor and Pister [6] in the context of the *fully nonlinear* incompressible problem. For this problem, an identical formulation in the context of the Hu-Washizu principle can be carried out.

Now let us reverse our viewpoint. Assume we *start* with (17) and (18) (i.e., the *B*-bar method). The finite element discretization is considered *defined* by these equations independently of the way in which stresses are to be computed. Is this seemingly *ad-hoc* method variationally consistent? The exposition preceding (17) and (18) makes the answer manifest: It is variationally consistent if the stress field is constructed so that conditions (16) are satisfied. Thus, the method of stress calculation, or *recovery*, is the central issue. This is dealt with in the following section.

4 Variational Stress Recovery

An equivalent characterization of the admissible stress fields τ^h satisfying orthogonality conditions (16) is the following: For simplicity in the notation, the superscript h will be dropped. Let $\bar{E}$ be the set of *discrete strain fields* generated by $\bar{B}_S$; i.e.,

$$\bar{E} := \left\{\gamma(x) \in [L^2(\Omega)]^2 \mid \quad \gamma_e \equiv \gamma\Big|_{\Omega_e} = \bar{B}_S q_e\right\} \tag{19}$$

In addition, introduce the notation

$$E_b := \left\{\gamma_b(x) \in [L^2(\Omega)]^2 \mid \quad \gamma_{b_e} \equiv \gamma_b\Big|_{\Omega_e} = [B_S - \bar{B}_S]q_e\right\} \tag{20}$$

Conditions (16) then, may be re-stated simply as

$$(\tau, \gamma_b)_{L_2} := \int_{\Omega_e} \tau \bullet \gamma_b d\Omega = 0, \text{ for any } \tau \in S^h, \text{ and } \gamma_b \in E_b \tag{21}$$

Therefore, an assumed strain method is variational provided the space of admissible stress fields S^h is contained in the *orthogonal complement* to the space E_b relative to usual L_2 inner product.

An Explicit Construction of S^h. We now show that given $\bar{B}_S$, a space of admissible stress fields S^h satisfying (21) may always be explicitly constructed as follows. For simplicity, assume linear constitutive equations so that

$$D_s := \frac{\partial^2\Psi_S(\gamma)}{\partial\gamma^2} = \text{const.} \tag{22}$$

Now consider the *orthogonal complement* E_b^{orth} of the space E_b relative to the inner product induced by D_s, that is

$$E_b^{orth} := \{\gamma_b^{ort}(x) \in [L^2(\Omega)]^2 \mid \quad <\gamma_{b_e}, \gamma_{b_e}^{\text{ort}}> = 0\}, \tag{23}$$

where we have set

$$<\gamma_{b_e}, \gamma_{b_e}^{ort}> := \int_{\Omega_e} \gamma_b\Big|_{\Omega_e} \bullet D_s \gamma_b^{ort}\Big|_{\Omega_e} d\Omega \tag{24}$$

Then, we define *admissible* stress fields as the stresses associated with the projection of the strain fields $\bar{E}$ onto E_b^{orth}, that is

$$\tau \in S^h \text{ if and only if } \tau\Big|_{\Omega_e} = D_s \gamma_b^{ort}, \text{ where } \gamma_b^{ort} \in E_b^{orth} \tag{25}$$

Clearly, by virtue of (23) the local stress field τ obtained in this way satisfies constraint (21). The explicit form of the interpolation for $\tau \in S^h$ may be obtained by constructing a basis for E_b^{orth}. This can always be accomplished by orthogonalizing the columns of $\bar{B}_S$ against the columns of $B_S - \bar{B}_S$ in the $<.\,,.>$ inner product, according to the following generalized Gram-Schmidt procedure.

Stress Interpolation. Suppose that E_b^{orth} has dimension *dim* $E_b^{orth} = \bar{N}_e \leq N_e$. Then, the number of independent columns of $[B_S - \bar{B}_S]$ is $\bar{N}_e$. Let $\bar{b}_b^I$, $(I = 1, \ldots, N_e)$ denote the column vectors of the matrix $[B_S - \bar{B}_S]$. Without loss of generality, assume that the first $\bar{N}_e$ columns are the independent column vectors. We introduce the notation

$$\bar{B}_b := [\bar{b}_b^1 \bar{b}_b^2 \ldots \bar{b}_b^{\bar{N}^e}], \text{ where } B_S - \bar{B}_S = [\bar{b}_b^1 \bar{b}_b^2 \ldots \bar{b}_b^{N_e}] \tag{26}$$

Associated with the independent column vectors $\bar{B}_b$ we define the Gram-Schmidt matrix $\bar{H}_b$ according to

$$\bar{H}_b := \int_{\Omega_e} (\bar{B}_b)^T D_S \bar{B}_b d\Omega \tag{27}$$

Note that by construction $\bar{H}_b$ is regular. We now define a *modified* stress interpolation operator $\bar{\bar{B}}_S$ such that, for a given set of generalized displacements $q_e \in \mathbb{R}^{N_e}$, the admissible stress fields $\tau \in S^h$ are defined in terms of $\bar{\bar{B}}_S$ as

$$\left\{\begin{aligned} \tau &= D_s \bar{\bar{B}}_S q_e \\ \bar{\bar{B}}_S &:= \bar{B}_S - \bar{B}_b \bar{H}_b^{-1}\Big|_{\Omega_e} (\bar{B}_b)^T \bar{B}_S d\Omega \end{aligned}\right\} \tag{28}$$

To see that (28) furnishes the appropriate stress interpolation, simply observe that any $\gamma_b \in E_b$ may be expressed as $\gamma_b = [B_S - \bar{B}_S]p_e \equiv \bar{B}_b \bar{p}_e$. Here, $\bar{p}_e$ denote the nodal degrees of freedom corresponding to the $\bar{N}_e$ independent columns of $[B_S - \bar{B}_S]$. From (28) it follows at once that

$$(\tau,\gamma_b)_{L^2} = p_e \bullet \left\{ \int_{\Omega_e} (\bar{\bar{B}}_S)^T D_S \bar{B}_b d\Omega \right\} q_e \equiv 0 \tag{29}$$

Thus, the stress fields constructed according to (28) satisfies the key orthogonality condition (21).

Remark 4.1. It should be noted that stress fields of the form

$$s \bigg|_{\Omega_e} = D_s \bar{B}_S q_e \tag{30}$$

often used in applications, do not satisfy condition (21) in general, thus rendering the method nonvariational.

Remark 4.2. The crucial role played by the Hu-Washizu principle should be noted. In the more restrictive variational framework furnished by the Hellinger-Reissner principle, once the strains are assumed as $\gamma = \bar{B}_S q_e$, the only possible choice for the stress field is (30). By contrast, within the framework outlined above, the stress field may be chosen as the *orthogonal complement* of E_b in L_2.

Remark 4.3. The developments discussed so far yield no information regarding the structure of $\bar{B}_S$. The stress recovery (28) merely ensures that for a given $\bar{B}_S$ the method is variationally consistent. However, variational consistency is by no means a guarantee of success. Issues concerning stability and convergence of the method depend crucially on the form of $\bar{B}_S$ and ultimately, as any mixed method, on a discrete LBB condition (e.g., [5]).

An Example: Mode Decomposition Technique. As a particular example, consider the *mode decomposition* proposed by Belytschko and co-workers. The idea is to introduce a projection on the nodal displacements q_e so that

$$q_e = P_B q_e + P_S q_e \tag{31}$$

where P_B and P_S are projection operators which, accordingly, satisfy

$$P_B + P_S = I,\ P_B \mathrm{o} P_B = P_B,\ P_S \mathrm{o} P_S = P_S,\ P_B \mathrm{o} P_S = P_S \mathrm{o} P_B = O \tag{32}$$

Then, the procedure is based upon selecting a discrete gradient operator $\bar{B}_S$ of the form

$$\bar{B}_S := B_S \mathrm{o} P_S \tag{33}$$

The methodology for constructing P_S is based on the concept of a so-called *equivalent Kirchhoff* configuration, see, e.g., [1]. The essential purpose of the method is to eliminate the so-called *parasitic* shear strains. Clearly, the procedure falls within the class of B-bar methods. Hence, the structure outlined in Section 3 entirely applies. In particular, the shear strains are computed according to

$$\gamma \bigg|_{\Omega_e} = \bar{B}_S q_e = B_S \mathrm{o} P_S q_e, \qquad \gamma_b \bigg|_{\Omega_e} = B_S \mathrm{o} P_B q_e, \tag{34}$$

The shear stresses must be obtained so that condition (21) is satisfied. For this purpose, the procedure leading to (28) is applied.

Remark 4.4. It should be noted that the stress recovery in the mode decomposition approach is usually performed according to expression (31) which renders in general *nonadmissible stress fields* violating orthogonality condition (21), see, e.g., [7] and references therein. Other structural elements employing a B-bar type of approach, e.g., [3], have also employed this stress recovery.

Remark 4.5. Zienkiewicz and Nakazawa [8], generalizing results of Malkus and Hughes [4], have considered the equivalence between numerically integrated displacement formulations and mixed formulations based upon the Hu-Washizu principle. They also consider the issue of a consistent variational recovery of strain and stress fields, within the context of *smooth* (*continuous*) interpolations for these fields. It should be noted that numerically integrated displacement methods and B-bar type of methods represent *different* classes of generalized displacement methods, although in some instances they may overlap. Thus, the results in [8] and those reported herein may be considered complementary.

Acknowledgment

The authors are indebted to R. L. Taylor for helpful discussions. Partial support for the first author was provided by a grant from LLNL. The second author was supported by ONR Grant N00014-84-K-060. This support is gratefully acknowledged.

References

1 Belytschko, T., Stolarski, H., and Carpenter, N., [1984], "A Triangular Plate Element With One-Point Quadrature," *Int. J. Num. Meth. Engng.*, Vol. 20, pp. 787-802.

2 Hughes, T. J. R., [1980], "Generalization of Selective Integration Procedures to Anisotropic and Nonlinear Media," *Int. J. Num. Meth. Engng.*, Vol. 15, pp. 1413-1418.

3 Hughes, T. J. R., and Tezduyar, T. E., [1981], "Finite Elements Based Upon Mindlin Plate Theory With Particular Reference to the Four Node Bilinear Isoparametric Element," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 587-596.

4 Malkus, D. S., and Hughes, T. J. R., [1978], "Mixed Finite Element Methods - Reduced and Selective Integration Techniques: A Unification of Concepts," *Comp. Meth. Appl. Mech. Engng.*, Vol. 15, pp. 63-81.

5 Oden, J. T., and Carey, G., [1983], *Finite Elements: Mathematical Aspects*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

6 Simo, J. C., Taylor, R. L., and Pister, K. S., [1984], "Variational and Projection Methods for the Volume Constraint in Finite Deformation Elastoplasticity," *Comp. Meth. Appl. Mech. Engng.*, Vol. 51, pp. 177-208.

7 Stolarski, H., and Belytschko, T., "On the Equivalence of Mode Decomposition and Mixed Finite Elements," *Comp. Meth. Appl. Mech. Engng.* (to appear).

8 Zienkiewicz, O. C., and Nakazawa, S., [1984], "On Variational Formulation and Its Modifications for Numerical Solution," *Comp. Struct.*, Vol. 19, p. 303-313.

Invariance Considerations in Large Strain Elasto-Plasticity

P. A. Dashner
Department of Mechanical Engineering,
California State Polytechnic University,
Pomona, CA 91768

This paper addresses a number of issues in the ongoing debate over the relevance and/or appropriateness of the rotational invariance requirement which is generally associated with the intermediate unstressed configuration. In particular, it is argued that the principle of material frame invariance has been "misapplied" by the proponents of full rotational invariance. Insistence on a strictly "kinematic" interpretation of the deformation gradient ($\mathbf{F}=\mathbf{F}_e\mathbf{F}_p$) *constituents* $\mathbf{F}_e$ *and* $\mathbf{F}_p$ *justifies elimination of the plastic rotational component* $\mathbf{R}_p(\mathbf{F}_p=\mathbf{R}_p\mathbf{U}_p)$ *based on the principle of determinism for stress—not invariance of frame. However, simple physical considerations, including a physical example involving a "structurally anisotropic" crystal, suggest that a more intricate definition of the gradient constituents is required in order to adequately account for microstructural characteristics. These considerations suggest alternative definitions for the gradient constituents* $\mathbf{F}_e$ *and* $\mathbf{F}_p$ *with associated constitutive invariance requirements.*

Introduction

In recent years a number of articles concerning the admissibility of certain invariance requirements relating to the deformation gradient decomposition

$$\mathbf{F}=\mathbf{F}_e\mathbf{F}_p \qquad (1)$$

have appeared. This debate centers on the question of whether a specific requirement of invariance under the transformation

$$(\mathbf{F},\mathbf{F}_e,\mathbf{F}_p)\rightarrow(\mathbf{F},\mathbf{F}_e\mathbf{Q},\mathbf{Q}^T\mathbf{F}_p) \qquad (2)^1$$

for the full orthogonal group is

(*i*) required in order to obtain "admissible" constitutive laws, or

(*ii*) over-restrictive and leads to a general theory which is something less than "general."

Those who argue for (2) assert that constitutive equations for the symmetric Piola-Kirchhoff stress and (Helmholtz) free energy of the form

$$\mathbf{S}=\mathbf{S}(\mathbf{E},\mathbf{F}_p,\kappa);\quad \psi=\psi(\mathbf{E},\mathbf{F}_p,\kappa), \qquad (3)$$

must be independent of the rotational component of $\mathbf{F}_p$ and therefore equivalent to equations having the Green-Naghdi [1] form

$$\mathbf{S}=\mathbf{S}(\mathbf{E},\mathbf{E}_p,\kappa);\quad \psi=\psi(\mathbf{E},\mathbf{E}_p,\kappa), \qquad (4)$$

expressed in terms of the total and plastic strain

$$2\mathbf{E}\equiv\mathbf{C}-\mathbf{I};\quad \mathbf{C}\equiv\mathbf{F}^T\mathbf{F}$$

$$2\mathbf{E}_p\equiv\mathbf{C}_p-\mathbf{I};\quad \mathbf{C}_p\equiv{\mathbf{F}_p}^T\mathbf{F}_p. \qquad (5)$$

The case for this position has been stated most recently by Casey and Naghdi [2, 3] while proposition (*ii*) has had recent support from a number of authors including Mandel [4], Lubliner [5], Lubarda and Lee [6], and Lee [7]. Here, arguments in support of this latter group are presented. Through simple microstructural considerations, I intend to demonstrate that the imposition of this additional invariance requirement severely compromises our ability to model materials which possess a "persistent" anisotropic lattice structure (e.g., single crystals). This conclusion follows from the fact that it effectively allows for one and only one lattice orientation in each plastically deformed rest configuration, regardless of the complexity of the plastic deformation path. This restrictive property ascribes a path-independent characteristic to the plastic deformation mechanism which is not supported by available evidence. Relating to this latter point, and in lieu of evidence which suggests that the accumulated plastic strain has any relevance to the determination of a deformed element's mechanical state, I will also suggest a move to alternative constitutive formalisms which do not rely on the identification of plastic strain as a "primitive" variable.

The basic assumptions which underpin this line of argument are not new. Those familiar with the body of literature authored by Havner, Hill, Kocks, Rice, and others[2] dealing with the plastic deformation of crystalline metals will recognize, and no doubt feel comfortable with, the underlying phenomenology. The intent here is to bring this line of reasoning to bear on the present and ongoing debate over invariance criteria and to cast this sort of theory in a form which may be more familiar to those accustomed to the formalisms of Lee and/or Green and Naghdi.

---

[1] This invariance requirement is in addition to the frame invariance (isotropy of space) requirement associated with the transformation $(\mathbf{F},\mathbf{F}_e,\mathbf{F}_p)\rightarrow(\mathbf{QF},\mathbf{QF}_e,\mathbf{F}_p)$.

[2] The recent review article by Havner [8] seems to give a thorough accounting of this line of research.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 28, 1983; final revision, May 23, 1985.

The present considerations nevertheless lead to a general constitutive formalism [9] which, although fully consistent with that elucidated by Havner [8], has a number of noteworthy features. Apart from the elimination of explicit dependence on the accumulated plastic strain the most striking feature of this alternative format is that it is set in an Eulerian, or spatial, context with the present deformed configuration adopted as the instantaneous reference. This approach is based on the assumption that an inelastic element's present state is wholly determined by

(*i*) its present geometric configuration,
(*ii*) the instantaneous geometry and orientation of the characteristic material lattice structure,
(*iii*) an "adequate" characterization of the instantaneous dislocation structure.

In this formulation the inherent path dependent nature of the inelastic deformation mechanisms is recognized through the explicit non-reliance on a fixed-material reference configuration (as distinct from a fixed-lattice reference) and the elimination of all measures of the accumulated plastic strain. These considerations, which draw upon all previously cited formalisms, offer a resolution to the present debate over invariance criteria and establish a suitable theoretical structure for the development of realistic, large-strain, inelastic material models.

The Difficulty

In order for the Cauchy (true) stress response equation

$$\sigma = \sigma(\mathbf{F},\mathbf{F}_p,\kappa) \tag{6}$$

to be concordant with the principle of determinism for stress, and thereby establish a theoretically valid point of departure for inelastic modeling, it is necessary that $\mathbf{F}_p$, like $\mathbf{F}$, be determined from the history of the local material motion relative to some specified reference. This is explicitly acknowledged by Lee in [10] when he states that ". . . the plastic deformation $\mathbf{F}_p$ is a functional which represents the past history of plastic flow that has already occurred, . . . " The fallacy of the current debate is that an argument has been joined over the invariance requirements which apply to a quantity which has not been properly defined in this causal sense. There is no debate over the meaning of the stretch component

$$\mathbf{U}_p = \sqrt{\mathbf{F}_p{}^T\mathbf{F}_p} \tag{7}$$

of $\mathbf{F}_p$ as it is understood to represent the unique positive definite square root of the Green plastic deformation tensor

$$\mathbf{C}_p = \mathbf{F}_p{}^T F_p. \tag{8}$$

The existence of $\mathbf{U}_p$ as a functional of past deformation history follows from the assumption that each material element, at each instant, possesses a unique relaxed or unstressed geometry. A similar causal definition of the rotational component $\mathbf{R}_p = \mathbf{F}_p\mathbf{U}_p{}^{-1}$ has never been fixed. In this regard, the demonstrated "non-uniqueness" of the gradient decomposition

$$\begin{gathered} \mathbf{F} = \mathbf{F}_e\mathbf{F}_p = \mathbf{F}_e{}^*\mathbf{F}_p{}^* \\ \mathbf{F}_e{}^* = \mathbf{F}_e\mathbf{Q} \\ \mathbf{F}_p{}^* = \mathbf{Q}^T\mathbf{F}_p \rightarrow \mathbf{R}_p{}^* = \mathbf{Q}^T\mathbf{R}_p, \end{gathered} \tag{9}$$

in terms of an arbitrary proper orthogonal $\mathbf{Q}$, can be taken as proof of the fact that purely kinematic considerations alone do not, and cannot, establish the causal functional nature of $\mathbf{R}_p$.

Proponents of the contested invariance requirement seize on this non-uniqueness of the gradient decomposition and state, in effect, that the instantaneous response cannot possibly depend on how one orients the unstressed element. Thus, they invoke invariance under the transformation

$$(\mathbf{F},\mathbf{F}_e,\mathbf{F}_p) \rightarrow (\mathbf{F},\mathbf{F}_e\mathbf{Q},\mathbf{Q}^T\mathbf{F}_p) \tag{10}$$

for the full proper orthogonal group, while suggesting that this represents a logical extension of the principle of invariance of frame. I submit that while their conclusion is formally correct, given their characterization of the primitive variables as having only kinematic significance, it is not appropriate to appeal to the principle of invariance of frame. Moreover, by choosing to ignore the essential microstructural aspects of solids, they arrive at a general theory which is antithetical to the nature of crystalline or polycrystalline solids which exhibit material anisotropy.

First of all, the constitutive hypotheses of *determinism* and *invariance of frame* seem to have a natural ordering in which the former precedes the latter. Indeed, they are based on completely independent philosophical principles. The principle of determinism for stress, i.e.,

> The instantaneous stress response of a material element at time $t > 0$ is a functional of the history of the element deformation proceeding from a "well-characterized" reference state at $t = 0$, i.e.,

$$\sigma(t) = \mathcal{R}_o[\mathbf{F}(s):0 < s \le t], \tag{11}$$

is based on an intuitive understanding of the nature of time and the supposition that causes necessarily precede effects. This principle underpins and, in some sense, justifies all scientific investigation. The hypotheses of invariance of frame, as it applies here, i.e.,

> The instantaneous stress response of a material element at time $t > 0$ depends on the history of element orientation only as a function of its current value, and only insofar as it "orients" the resultant stress, i.e.,

$$\begin{gathered} \mathbf{Q}\sigma(t)\mathbf{Q}^T = \mathcal{R}_o[\mathbf{QF}(t);\mathbf{C}(s):0 < s \le t]; \\ \mathbf{C} = \mathbf{F}^T\mathbf{F}; \quad \text{for all proper orthogonal } \mathbf{Q} \end{gathered} \tag{12}$$

is based on the notion of isotropy of space. Since this hypothesis freely incorporates the idealization of Galilean relativity, among other things, it is clearly subordinate to the more fundamental assertion of determinism. In fact, the specific wording given above presupposes stress determinism through (11).

In view of these considerations it is difficult to justify the imposition of frame invariance on a constitutive form which already stands in violation of the principle of determinism. Within the context of the present debate, an explicit definition of $\mathbf{R}_p$, and thus of $\mathbf{F}_p$, in terms of prior deformation history must precede the consideration of additional invariance requirements. Put differently, the existence of a constitutive relation

$$\mathbf{R}_o(t) = \mathcal{N}_o[\mathbf{F}(s):0 < s \le t] \tag{13}$$

must be established before additional restrictions can be meaningfully imposed. If no such physically relevant definition of $\mathbf{R}_p$ exits, then it must be excluded from (6), leaving the form

$$\sigma = \sigma(\mathbf{F},\mathbf{U}_p,\kappa), \tag{14}$$

as a consequence of the principle of determinism for stress, not invariance of frame. On the other hand, if a physically relevant, causal definition of $\mathbf{R}_p$ can be formulated, then the validity of any additional invariance requirement can be assessed only in light of this definition. It should also be clear that no such definition of $\mathbf{R}_p$ is possible without the introduction of microstructural, in addition to purely kinematic, considerations. To exclude consideration of microstructure in modeling the behavior of solids is to ignore the inherent physical nature of this class of materials.

Additional Considerations

Within the context of a general theory, it is clearly necessary to allow for material anisotropy. While this is recognized by all parties there is not yet agreement on how this should be done. In fact, this issue seems to provide the focus for the ongoing debate. In view of the considerations of the previous Section, I believe that it is crucial to develop a "meaningful" causal definition of $\mathbf{R}_p$ which bears some relationship to anisotropic characteristics through material microstructure. For this purpose I have adopted the phenomenology elucidated by Rice [11; Sec. 3.2], in incremental form by Asaro and Rice [12; Sec. 2], and most recently in the review article by Havner [8; Sec. 3.2 & 4.1] in connection with the slip-shear deformation of single crystals.

In an effort to cast this discussion in a somewhat more general context I shall begin by dividing all anisotropic characteristics into two mutually exclusive categories. Specifically, any directional characteristic which is present and recognizable in all unstressed elements of a given material will be referred to as a *structurally anisotropic* characteristic. All others, i.e., those which may or may not be present depending on prior deformation history, will be referred to as *induced anisotropic* characteristics. Structural anisotropy will most likely be associated with chemical composition and the corresponding *crystallographic*, or *cell*, structure of the material while induced anisotropy is most likely to result from the local density and distribution of dislocations within the crystallites.[3]

Once the existence of structurally anisotropic materials is admitted (the paradigm example being the single crystal), then a causal definition of $\mathbf{R}_p$ of obvious physical relevance becomes apparent. In order to assign a particular value to $\mathbf{F}_p$, and therefore to its rotational component $\mathbf{R}_p$, it is first necessary to select and set aside a particular "virgin" element of this material which exhibits only the characteristic structural anisotropy. This element shall henceforth serve as a reference map for the underlying material bond structure and is to be regarded as an integral part of any constitutive relation. Having established such an element, the so called *reference cell*, as a standard for comparison, the value of $\mathbf{F}_p$ assigned to any other element of this material is taken as that which "places" the corresponding relaxed or unstressed element with structural orientation identical to that of the reference cell. Based on this definition of $\mathbf{F}_p$ and the gradient decomposition (1), it is evident that $\mathbf{F}_e$ can then be regarded as the linear map which fixes the current "lattice configuration" by "placing" the characteristic structure of the reference cell in the deformed material element. This makes it possible to visualize the total element deformation as a two step process consisting of an initial "plastic" deformation characterized by a spatially fixed cell structure, followed by an "elastic" deformation in which the characteristic cell structure is "dragged along" with the material. For reasons of convenience I choose to regard $\mathbf{F}$ and $\mathbf{F}_e$ as primative, with $\mathbf{F}$ placing the material relative to a chosen material reference and $\mathbf{F}_e$ providing the essential outline for the description of lattice or material bond structure (instantaneous elastic distortion and "structural" orientation) by placing the preselected reference cell in the deformed material element.

In the event of full structural anisotropy (orthotropy), this definition assigns unique values to the gradient constituents of each material element based entirely on local deformation history. On the other hand, if the material exhibits no structural anisotropy then, based on the above definition of $\mathbf{F}_p$, proper orthogonal $\mathbf{R}_p$ would clearly be indeterminant. Thus, adoption of the above definitions establishes the invariance requirement (2) as a consequence of structural isotropy. Based on these considerations it can be argued that the contested invariance requirement should rightfully be replaced by a requirement of invariance under the transformation

$$(\mathbf{F},\mathbf{F}_e,\mathbf{F}_p)\rightarrow(\mathbf{F},\mathbf{F}_e\mathbf{Q},\mathbf{Q}^T\mathbf{F}_p); \quad \text{for each } \mathbf{Q}\in\mathcal{G} \qquad (15)$$

where $\mathcal{G}$ represents the orthogonal symmetry group for the virgin reference cell.

The above discussion would seem to suggest that the theory of Green and Naghdi is incapable of modeling structurally anisotropic materials, as the imposition of invariance under (2) would appear to deny the influence of structural orientation. This, however, is not necessarily the case since elimination of $\mathbf{R}_p$ from (3), and subsequent reduction to (4), would also result from the explicit assumption that

$$\mathbf{R}_p = \mathbf{R}_p(\mathbf{F},\mathbf{U}_p,\kappa). \qquad (16)$$

Since elastic unloading to $\mathbf{F}=\mathbf{U}_p$ would presumably not further alter the structural orientation (relative to the material) in the relaxed element, (16) is seen to imply a simpler relationship of the form

$$\mathbf{R}_p = \mathbf{R}_p(\mathbf{U}_p,\kappa). \qquad (17)$$

The obvious implication of (17) is that there is one and only one cell orientation corresponding to each state of plastic strain at a given hardness. In other words, it implies that the evolution of lattice orientation, relative to the material, is independent of the plastic deformation path and thereby fixed by the accumulated plastic strain. There is, however, no evidence to suggest that this is characteristic of inelastic deformation in metals. In fact, it is a simple excercise to construct a physical counterexample (see Appendix) by demonstrating the existence of a closed plastic deformation path in a non-hardening crystal, consisting of a sequence of four single shear deformations along a pair of perpendicular slip systems, which results in a net rotation of the underlying crystal lattice relative to the material. Such examples, although highly idealized, serve to suggest that the above "path independent" characteristic is not a property which is generic to metal inelasticity.

This apparent deficiency in the Green-Naghdi format suggests one which is even more fundamental in nature. This relates to the selection of the accumulated plastic strain as a "primitive" state variable. As demonstrated above, functional dependence on accumulated plastic strain (as opposed to the entire history of plastic strain) appears to assign inherent path independent characteristics to the plastic deformation mechanism which are not borne out by available evidence. In effect, the plastic deformation mechanism disassociates "material distribution" from bond structure, with $\mathbf{F}_p$ giving the current accounting of this relationship. Havner [8; Sec. 3.1] expresses this by noting that, "The single most important phenomenological aspect of gross inelastic behavior in metal crystals (indeed, distinguishing them by kinematics alone from polymers and soils) is *the movement of the macroscopic material relative to the underlying crystalline structure.*" The key assumption which underlies all theoretical development based on the above phenomenology is that the mechanical state of a deformed material element is determined by the instantaneous spatial bond distribution (lattice placement plus dislocation distribution), independent of the material distribution. This assumption shall be formalized in the concluding Section through the introduction of a new invariance requirement which may be referred to as *replacement invariance*. Further, due to the fundamentally path dependent nature of the plastic deformation mechanism, I consider it improbable that accumulated plastic deformation has any relevance to the characterization of bond structure. To identify plastic strain as a primitive state variable is to suggest otherwise. As shall be demonstrated in the concluding Section, the requirement that all constitutive forms be invariant with respect to *material element replacement* provides a justification for the elimination of plastic strain as a primitive variable

---

[3] Based on statistical averaging, most annealed polycrystalline metals should probably be regarded as structurally isotropic.

and establishes a needed connection between the different formalisms employed by the contending parties.

An Alternative Formalism

The fundamental premise which guides the following development is that the instantaneous lattice placement and dislocation structure fix the mechanical state of deformed material elements. That is, any two deformation paths, regardless of final material placement, which result in identical lattice placements and spatial dislocation distributions should produce mechanically indistinguishable elements. Although these assertions would support a direct statement of the attendant constitutive forms, a more formal approach is adopted wherein the appropriate response forms are deduced through a systematic application of invariance requirements to a more general (and less controversial) form. For convenience, the response equations are first cast in the form

$$[\sigma,\psi]=\mathcal{R}(\mathbf{F},\mathbf{F}_e,\eta), \tag{18}$$

in terms of the Cauchy (true) stress σ, the (Helmholtz) free energy density ψ, the familiar total and elastic deformation gradients $\mathbf{F}$ and $\mathbf{F}_e$, and a set of frame invariant, Eulerian[4] tensor state variables $\eta=\{\mathbf{q}_i\}$. These additional variables will likely describe certain geometric aspects of the instantaneous dislocation distribution and are introduced in order to model the dislocation induced variation in mechanical properties. The conjecture of the previous Section is that these variables are inherently history dependent and that there is little likelihood that they will exhibit explicit dependence on accumulated plastic strain.

The first invariance requirement that must be imposed on (18) is that of material frame invariance. Since $\mathbf{F}$ and $\mathbf{F}_e$, respectively, "place" the material element and the characteristic cell lattice in the current configuration, it is clearly necessary and sufficient to require that

$$[\mathbf{Q}\sigma\mathbf{Q}^T,\psi]=\mathcal{R}(\mathbf{QF},\mathbf{QF}_e,\mathbf{T}_Q\eta), \tag{19}$$

for all proper orthogonal $\mathbf{Q}$ corresponding to post-deformation rotation of the material element. Here, $\mathbf{T}_Q$ represents the rotational group transformation operator associated with the set of tensor state variables η. Structural symmetries in the characteristic lattice are accounted for by demanding invariance under (15). This insensitivity to certain prerotations of the reference cell gives rise to the formal requirement that

$$[\sigma,\psi]=\mathcal{R}(\mathbf{F},\mathbf{F}_e\mathbf{Q},\eta);\quad \text{for all orthogonal } \mathbf{Q}\in\ \underset{\sim}{\mathcal{G}}. \tag{20}$$

Replacement invariance, as described above, takes the form of invariance of response under the transformation

$$(\mathbf{F},\mathbf{F}_e,\mathbf{F}_p,\eta)\rightarrow[\bar{\mathbf{F}}\mathbf{F},\mathbf{F}_e,(\mathbf{F}_e{}^{-1}\bar{\mathbf{F}}\mathbf{F}_e)\mathbf{F}_p,\eta] \tag{21}$$

for all possible non-singular replacements $\bar{\mathbf{F}}$ (see Fig. 1). "Replacement" of a deformed material element (placed by $\mathbf{F}$) can be visualized as one or a series of perfectly plastic shear-slip deformations which do not effect the geometry or orientation of the underlying lattice structure.[5] With the substitution

$$\mathbf{F}^*=\mathbf{F}^{-1}\bar{\mathbf{F}}\mathbf{F};\ \ \bar{\mathbf{F}}=\mathbf{F}\mathbf{F}^*\mathbf{F}^{-1}, \tag{22}$$

this transformation is seen to be equivalent to

$$(\mathbf{F},\mathbf{F}_e,\mathbf{F}_p,\eta)\rightarrow(\mathbf{F}\mathbf{F}^*,\mathbf{F}_e,\mathbf{F}_p\mathbf{F}^*,\eta). \tag{23}$$

Thus, replacement invariance is assured if and only if

$$[\sigma,\psi]=\mathcal{R}(\mathbf{F}\mathbf{F}^*,\mathbf{F}_e,\eta), \tag{24}$$

for all possible nonsingular $\mathbf{F}^*$. If dilitation is regarded as a purely elastic phenomenon (as it usually is), then the set of allowable replacements would be restricted by the determinant constraint

$$\det(\bar{\mathbf{F}})=\det(\mathbf{F}^*)=1. \tag{25}$$

Due to (25), (24) would render arbitrary only the unimodular component of $\mathbf{F}$ and hence it is more convenient to impose these invariance requirements on the equivalent response form

$$[\sigma,\psi]=\mathcal{R}(\mathbf{F}_p,\mathbf{F}_e,\eta) \tag{26}$$

expressed in terms of $\mathbf{F}_p=\mathbf{F}_e{}^{-1}\mathbf{F}$. With the adoption of (26) it is clear that the invariance requirements (19, 20, 24) are replaced by

Frame invariance

$$[\mathbf{Q}\sigma\mathbf{Q}^T,\psi]=\mathcal{R}(\mathbf{F}_p,\mathbf{QF}_e,T_Q\eta);$$
$$\text{for each proper orthogonal } \mathbf{Q}, \tag{27}$$

Lattice orthotropy

$$[\sigma,\psi]=\mathcal{R}(\mathbf{Q}^T\mathbf{F}_p,\mathbf{F}_e\mathbf{Q},\eta);\quad \text{for each orthogonal } \mathbf{Q}\in\mathcal{G}, \tag{28}$$

Replacement invariant

$$[\sigma,\psi]=\mathcal{R}(\mathbf{F}_p\mathbf{F}^*,\mathbf{F}_e,\eta];\quad \text{for each unimodular } \mathbf{F}^*. \tag{29}$$

Now, since $\mathbf{F}_p$ would also be subject to the unimodular constraint (25), invariance under (29) would render its value arbitrary and thereby force its removal from the list of arguments in the response equation (26).[6]

Thus, imposition of replacement invariance leads to a new formalism for the modeling of inelastic solids based on response functions of the form

$$[\sigma,\psi]=\mathcal{R}(\mathbf{F}_e,\eta). \tag{30}$$

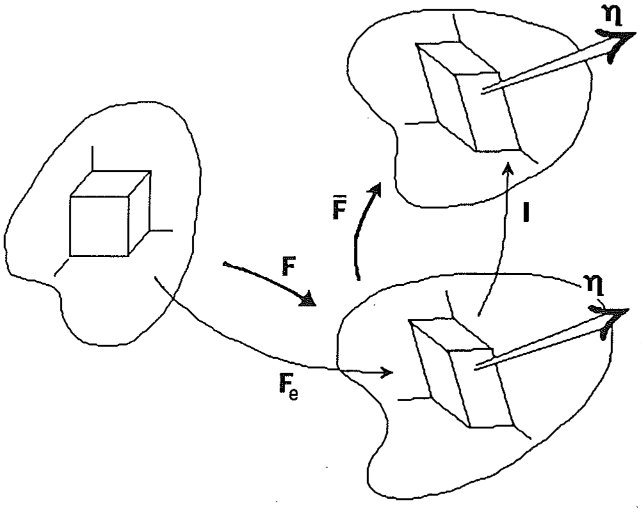


Fig. 1 Element "replacement" resulting in unaltered mechanical characteristics

---

[4] Eulerian in the sense that they are defined and interpreted relative to the current element configuration.

[5] If the material being modeled supports such deformations, then element replacement could be achieved through this mechanism. More generally, however, replacement invariance should be regarded as nothing more nor less than a mathematical statement of the assertion made in the second sentence of this Section.

[6] It is a simple matter to verify that the form

$$\mathbf{S}=\mathbf{S}(\mathbf{C},\mathbf{C}_p,\kappa)$$

is constrained by the requirement that

$$\mathbf{F}^{*-1}\mathbf{S}(\mathbf{F}^{*-1})^T=\mathbf{S}(\mathbf{F}^{*T}\mathbf{C}\mathbf{F}^*,\mathbf{F}^{*T}\mathbf{C}_p\mathbf{F}^*,\kappa);\quad \text{for all unimodular } \mathbf{F}^*,$$

under the imposition of replacement invariance. Since this guarantees invariance under all element prerotations, it must be concluded that this particular theoretical form cannot incorporate replacement invariance without being restricted to materials which have isotropic structure. An invariance requirement of this type was employed in [13] in the modeling of a special class of isotropic elastic-plastic solids having invariant elastic properties.

This form is further restricted by the usual frame invariance requirement

$$[\mathbf{Q}\sigma\mathbf{Q}^T,\psi]=\mathcal{R}(\mathbf{QF}_e,T_Q\eta);\quad \text{for each proper orthogonal } \mathbf{Q}, \tag{31}$$

and, the lattice orthotropy constraint

$$[\sigma,\psi]=\mathcal{R}(\mathbf{F}_e\mathbf{Q},\eta);\qquad \text{for each orthogonal } \mathbf{Q}\epsilon\mathcal{G}. \tag{32}$$

In view of (32), full structural isotropy would make it possible to rewrite (30) in either of the alternative forms

$$[\sigma,\psi]=f(\mathbf{b},\eta)=g(\mathbf{c},\eta), \tag{33}$$

in terms of the positive definite, symmetric, elastic deformation measures

$$\mathbf{b}\equiv\mathbf{F}_e\mathbf{F}_e{}^T \tag{34}$$

$$\mathbf{c}\equiv\mathbf{b}^{-1}=(\mathbf{F}_e{}^{-1})^T\mathbf{F}_e{}^{-1}.$$

These reduced forms are subject to frame invariance through

$$[\mathbf{Q}\sigma\mathbf{Q}^T,\psi]=f(\mathbf{QbQ}^T,T_Q\eta)=g(\mathbf{QcQ}^T,T_Q\eta), \tag{35}$$

but satisfy (32) identically.

The considerations of this Section underlie, and establish the point of departure for, the alternative constitutive formalism detailed in [9]. This format, which is fully compatible with the previously cited theoretical formulations of Havner, Hill, Rice, et. al., is not burdened with the physically overrestrictive invariance requirement (2) and deemphasizes the mechanical significance of accumulated plastic strain. This is accomplished by assigning full responsibility for the modeling of plastically induced mechanical effects on a collection of history dependent internal variables consisting of the "cell placement" tensor $\mathbf{F}_e$ and a set of (as yet unspecified) state variables $\eta=\{\mathbf{q}_i\}$. The history dependence of these variables is explicitly incorporated into this general theory by postulating the existence of frame invariant, causal, evolution equations which take the form

$$[\mathbf{QF}_e(t),\ T_Q\eta(t)]=\mathcal{G}_o[\mathbf{QF}(t);\ \mathbf{C}(s):\ 0<s\leq t];$$
$$\text{for all proper orthogonal } \mathbf{Q}, \tag{36}$$

when referenced to some "well-characterized" state at $t=0$. Havner [8] traces the development of incremental evolution laws for $\mathbf{F}_e$ (using the symbol $\mathbf{J}$ in place of $\mathbf{F}_e$) under the circumstance wherein slip-shear is the dominant plastic deformation mechanism.

References

1 Green, A. E., and Naghdi, P. M., "A General Theory of an Elastic-Plastic Continuum," *Arch. Rati. Mech. Anal.*, Vol. 18, 1965, pp. 251–281.

2 Casey, J., and Naghdi, P. M., "A Remark on the Use of the Decomposition $\mathbf{F}=\mathbf{F}_e\mathbf{F}_p$ in Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 672–675.

3 Casey, J., and Naghdi, P. M., discussion of "A Correct Definition of Elastic and Plastic Deformation and Its Computational Significance," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 983–984.

4 Mandel, J., "Equations Constitutives et Directeurs dans les Milieux Plastiques et Viscoplastiques," *Int. J. Solids Structures*, Vol. 9, 1973, pp. 725–740.

5 Lubliner, J., "A Thermodynamic Yield Criterion in Viscoplasticity," *Acta Mechanica*, Vol. 30, 1978, pp. 165–174.

6 Lubarda, V. A., and Lee, E. H., "A Correct Definition of Elastic and Plastic Deformation and Its Computational Significance," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 35–40.

7 Lee, E. H., "Some Comments on Elastic-Plastic Analysis," *Int. J. Solids Structures*, Vol. 17, 1981, pp. 859–872.

8 Havner, K. S., "The Theory of Finite Plastic Deformation of Crystalline Solids," *Mechanics of Solids: The Rodney Hill 60th Anniversary Volume*, H. G. Hopkins and M. J. Sewell, eds., Pergamon Press, Oxford, 1982, pp. 265–302.

9 Dashner, P. A., "Large Strain Inelastic State Variable Theory," *Int. J. Solids Structures*, in press.

10 Lee, E. H., "Elastic-Plastic Deformation at Finite Strains," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, 1969, pp. 1–6.

11 Rice, J. R., "Inelastic Constitutive Relations for Solids: An Internal Variable Theory and Its Application to Metal Plasticity," *J. Mech. Phys. Solids*, Vol. 19, 1971, pp. 433–455.

12 Asaro, R. J., and Rice, J. R., "Strain Localization in Ductile Single Crystals," *J. Mech. Phys. Solids*, Vol. 25, 1977, pp. 309–338.

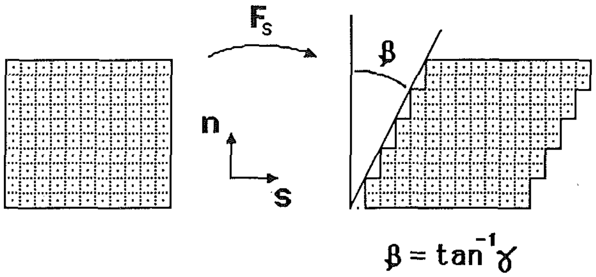


Fig. 2 Slip-shear deformation

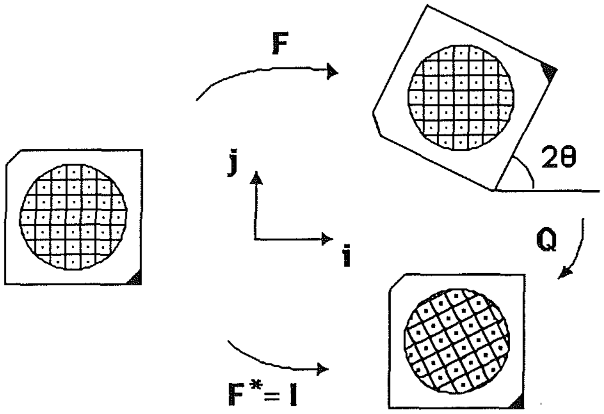


Fig. 3 Crystal reorientation via slip-shear deformation

13 Dashner, P. A., "A Finite Strain Work-Hardening Theory for Rate Independent Elasto-Plasticity," *Int. J. Solids Structures*, Vol. 15, 1979, pp. 519–528.

APPENDIX

An Interesting Example

For this example consider a unit cell of a perfect single crystal having horizontal and vertical families of slip planes. This element shall be subjected to a sequence of perfect single slip-shear deformations which alternately activate the horizontal and vertical slip systems. As is well-known, a perfect slip-shear deformation (see Fig. 2) is characterized by a displacement field of the form

$$\mathbf{u}(\mathbf{r})=\mathbf{u}_o+\gamma(\mathbf{s}\otimes\mathbf{n})\cdot\mathbf{r} \tag{37}$$

and deformation gradient

$$\mathbf{F}_s=\mathbf{I}+\gamma(\mathbf{s}\otimes\mathbf{n}) \tag{38}$$

in terms of the normal $\mathbf{n}$ to the active slip planes, the shearing or glide direction $\mathbf{s}$, and a positive shear strain parameter γ. Moreover, it is clear that such a deformation changes the material geometry without altering the cell size [$\det(\mathbf{F}_s)=1$], or the local geometry or spatial orientation of the underlying crystallographic structure.

Now, let $\mathbf{i}$ and $\mathbf{j}$ represent the horizontal and vertical directions, respectively, and consider the four step slip-shear deformation

$$\mathbf{F}=\mathbf{F}_4\mathbf{F}_3\mathbf{F}_2\mathbf{F}_1$$
$$=[\mathbf{I}+\gamma_2(\mathbf{j}\otimes\mathbf{i})][\mathbf{I}-\gamma_1(\mathbf{i}\otimes\mathbf{j})][\mathbf{I}+\gamma_1(\mathbf{j}\otimes\mathbf{i})][\mathbf{I}-\gamma_2(\mathbf{i}\otimes\mathbf{j})] \tag{39}$$

$$\{\mathbf{F}\}_{xyz}=\begin{pmatrix} 1-\gamma_1{}^2 & -[\gamma_2(1-\gamma_1{}^2)+\gamma_1] \\ \gamma_2(1-\gamma_1{}^2)+\gamma_1 & (1-\gamma_1\gamma_2)^2-\gamma_2{}^2 \end{pmatrix}$$

expressed in terms of the shearing parameters γ_1 and γ_2. The immediate objective is to show that γ_1 and γ_2 can be chosen so that the accumulated effect, as measured by **F**, amounts to a counterclockwise rigid material rotation (see Fig. 3) by an angle $\beta = 2\theta$, for any angle θ satisfying $-\pi/2 < \theta < \pi/2$, that is

$$\begin{bmatrix} 1-\gamma_1{}^2 & -[\gamma_2(1-\gamma_1{}^2)+\gamma_1] \\ \gamma_2(1-\gamma_1{}^2)+\gamma_1 & (1-\gamma_1\gamma_2)^2-\gamma_2{}^2 \end{bmatrix} = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}. \tag{40}$$

Note that this is accomplished by satisfying any two of the three equations

$$\begin{aligned} F_{xx} &= 1-\gamma_1{}^2 = \cos 2\theta \\ F_{yy} &= (1-\gamma_1\gamma_2)^2 - \gamma_2{}^2 = \cos 2\theta \\ F_{yx} &= -F_{xy} = \gamma_2(1-\gamma_1{}^2)+\gamma_1 = \sin 2\theta \end{aligned} \tag{41}$$

by virtue of the fact that **F** identically satisfies the determinant constraint $\det(\mathbf{F}) = 1$. Simultaneous solution of the first and third equations yields the result

$$\begin{aligned} \gamma_1 &= \sqrt{2}\sin\theta \\ \gamma_2 &= \sqrt{2}\sin\theta/(1+\sqrt{2}\cos\theta). \end{aligned} \tag{42}$$

With reference to Fig. 3, it therefore follows that for any choice of $\theta \epsilon (-\pi/2, \pi/2)$, the five step deformation process

$$\mathbf{F}^* = \mathbf{QF} = \mathbf{QF}_4\mathbf{F}_3\mathbf{F}_2\mathbf{F}_1, \tag{43}$$

which terminates with a clockwise rigid rotation **Q** through an angle 2θ, describes a closed deformation cycle ($\mathbf{F}^* = \mathbf{I}$) which results in a net rotation of the underlying lattice structure. Having assumed an initial absence of dislocations (perfect crystal), no hardening can occur during the perfect slip-shear deformations. Consequently, it is clear that the initial and final states of this element are characterized by

$$\{\mathbf{E} = \mathbf{E}_p, \kappa\}_f = \{\mathbf{E} = \mathbf{E}_p, \kappa\}_i. \tag{44}$$

Thus, the particular version of the Green and Naghdi theory in which the primitive variable $\mathbf{E}'$ is associated with the plastic strain $\mathbf{E}_p$ is not able to distinguish between these obviously distinct (provided of course that proper orthogonal **Q** does not belong to the lattice symmetry group) mechanical states.

The lesson of this example is inescapable insofar as it clearly demonstrates that the evolution of material characteristics associated with "structural" anisotropy depends in a complicated manner on the entire plastic deformation history. Inasmuch as the Green and Naghdi formulation depends on a "realistic" description of the anisotropic characteristics of subsequent unstressed elements in terms of the accumulated plastic strain and hardness (and perhaps a number of fixed material tensors associated with the initial anisotropic reference configuration) it must be regarded as overrestrictive. It is noteworthy that these conclusions follow from the consideration of a highly specialized example in which only structural anisotropic characteristics (i.e., those associated with lattice structure) are relevant. One can only speculate on how much stronger a plastic path dependence is exhibited by the induced anisotropic characteristics which arise from the generation and interaction of lattice dislocations.

J. E. Michaels[1]

Yih-Hsing Pao
Fellow ASME

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

Determination of Dynamic Forces From Wave Motion Measurements

An experimental method has been developed for generating oblique forces with known orientations and time histories. Recorded signals from several forces were analyzed by an iterative deconvolution method to determine their orientations and time histories. The recovered values agree closely with the exact ones for these controlled sources. These experiments are a valuable test of source characterization methods that may be applied to seismic data from earthquake sources or to signals recorded from the acoustic emission of cracks.

1 Introduction

A time-dependent concentrated force with fixed orientation that is applied to a structure generates wave motion which may be recorded on the surface. If the location, time history, and orientation of the force are known, the displacement response can be calculated for simple structures for which the Green's functions are known. To record the response to either a known or unknown source, a transducer can be mounted on the surface of the structure. In this paper a method is described and demonstrated to experimentally determine the orientation and time history of an oblique force applied to the surface of an elastic plate.

A closely related problem is determining the time history of a concentrated force of known orientation. Goodier et al. [1959] solved an integral equation to calculate the time history of a vertical force applied to a half-space from the far-field response. Hsu et al. [1977] and Michaels et al. [1981] discretized and inverted a time convolution integral to determine the time history of a vertical force applied to a plate from the near-field response. An important factor in the success of their work was the availability of an artificial source, which is generated by fracturing a glass capillary against the structure surface. As was first noted by Breckenridge et al. [1967], this source is a concentrated vertical force that has the time-dependence of a step-like unloading function.

The response of a structure to a force of known orientation is given by a convolution in time of a source function with a single Green's function. Thus, the problem of determining the source function from the response may be solved by deconvolution. However, for a force of unknown orientation, the response is given by a convolution of the source time function with a linear combination of Green's functions, where the unknown coefficient of each Green's function is proportional to a direction cosine of the force. Therefore, methods for deconvolution with a single Green's function cannot be directly applied to determine both the source time function and the direction cosines. In this paper, a deconvolution method recently developed by Michaels and Pao [1985] for multiple Green's functions is applied to experimental data.

2 Theory

Before the orientation and time history of an oblique force can be determined from the measured wave motion, it must be understood how the response depends upon the source, the medium, and the receiver. It is assumed here that the medium is an infinite elastic plate and that the receivers are piezoelectric transducers sensitive to normal motion.

Displacement Response in a Plate. Consider the plate geometry shown in Fig. 1. In cylindrical coordinates (r, θ, z), the source is located at $\mathbf{x}^o = (0, 0, 0)$ and a typical receiver is at $\mathbf{x} = (r, \theta, h)$, where h is the plate thickness. It is assumed that the horizontal dimensions of the plate are large enough so that it can be modeled as infinite in extent.

The Green's displacement tensor, $G_{ij}(\mathbf{x}, t; \mathbf{x}^o)$, is defined to be the displacement response in the ith direction at $\mathbf{x}$ and t due to an impulsive concentrated force of unit magnitude in the jth direction at $\mathbf{x}^o$ and $t=0$. Thus, for a point force $F_j(\mathbf{x}^o, t)$ acting at $\mathbf{x}^o$ that is zero for $t<0$, the resulting displacement is,

$$\begin{aligned} u_i(\mathbf{x}, t) &= \sum_{j=1}^{3} \int_0^t d\tau G_{ij}(\mathbf{x}, t-\tau; \mathbf{x}^o) F_j(\mathbf{x}^o, \tau) \\ &= \sum_{j=1}^{3} G_{ij}(\mathbf{x}, t; \mathbf{x}^o) * F_j(\mathbf{x}^o, t) \end{aligned} \tag{1}$$

In this and subsequent equations, an asterisk denotes a convolution integral in the time variable.

It is assumed here that the orientation of the oblique force F_j does not change with time. It then may be expressed as,

$$F_j(\mathbf{x}^o, t) = f_j(\mathbf{x}^o) s(t) \tag{2}$$

The function $s(t)$ is the time history of the force, which is the same for all three components of F_j. The vector $f_j(\mathbf{x})$ is the time-independent orientation. This decomposition in equation (2) is not unique because it is defined only to within a scale factor.

---

[1] Present address: JTM Systems and Consulting, Inc., Terrace Hill, Ithaca, NY 14850.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 20, 1984; final revision, August 15, 1985.

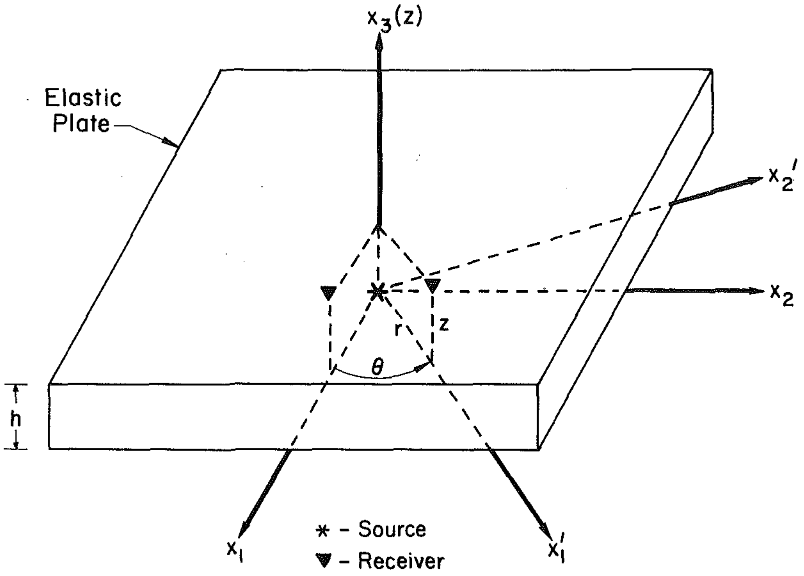


Fig. 1 Source and receiver geometry for an infinite elastic plate

Since the plate is axially symmetric and the source is located on the x_3 (or z) axis, the normal displacement $u_3(r, \theta, h, t)$ may be evaluated from the Green's functions at $\mathbf{x} = (r, 0, h)$ by rotating the components of the force f_j,

$$u_3(r, \theta, h, t) = \left\{ \sum_{j=1}^{3} G_{3j}(r, 0, h, t; \mathbf{0}) f_j' \right\} * s(t) \tag{3}$$

The components f_j' are,

$$\begin{aligned} f_1' &= f_1 \cos\theta + f_2 \sin\theta \\ f_2' &= -f_1 \sin\theta + f_2 \cos\theta \\ f_3' &= f_3 \end{aligned} \tag{4}$$

Equation (3) may be simplified by noting that $G_{32}(r, 0, h, t; \mathbf{0})$ vanishes identically because of the axial symmetry of the plate (Ceranoglu and Pao, 1981).

$$\begin{aligned} u_3(r, \theta, h, t) = \{ &(f_1 \cos\theta + f_2 \sin\theta) G_{31}(r, 0, h, t; \mathbf{0}) \\ &+ f_3 G_{33}(r, 0, h, t; \mathbf{0}) \} * s(t) \end{aligned} \tag{5}$$

Thus, if the Green's functions are known, the normal displacement response to a known oblique force may be calculated.

Deconvolution with Multiple Green's Functions. To determine f_1, f_2, f_3 and $s(t)$ from the measured normal displacement u_3, we must consider the problem of deconvolution where the kernel is a sum of Green's functions with unknown coefficients. The method used is described in detail by Michaels [1984] and Michaels and Pao [1985] and is only briefly reviewed here.

Equation (5) for displacement is of the form,

$$u(t) = \left\{ \sum_{m=1}^{M} c_m G_m(t) \right\} * s(t) \tag{6}$$

where we identify,

$$\begin{aligned} u(t) &= u_3(r, \theta, h, t) \\ M &= 2 \\ c_1 &= f_1 \cos\theta + f_2 \sin\theta \\ G_1(t) &= G_{31}(r, 0, h, t; \mathbf{0}) \\ c_2 &= f_3 \\ G_2(t) &= G_{33}(r, 0, h, t; \mathbf{0}) \end{aligned} \tag{7}$$

In general, there are several receivers, and equation (6) is valid for each receiver but with different c_m and $G_m(t)$.

The first step is to calculate c_1, c_2 and $s(t)$ at each receiver location by an iterative deconvolution procedure. The coefficients c_1 and c_2 are first set to non-zero initial values. Using these values, the source time function $s(t)$ is estimated by least squares deconvolution. Then, improved estimates of c_1 and c_2 are calculated from the estimate of $s(t)$, again by least squares. This procedure of alternately calculating c_m and $s(t)$ is continued until they converge to stable values.

The final estimate of $s(t)$ is obtained by averaging the signals obtained by deconvolution at all of the receivers. Similarly, f_3 is obtained by averaging the coefficient c_2. However, the iterative deconvolution procedure does not recover f_1 and f_2 directly. They are imbedded in the coefficient c_1 as shown in equation (7). Thus, to calculate f_1 and f_2, there must be at least two receivers located at different angular positions. For more receivers, a least squares minimization is performed to obtain f_1 and f_2 from the coefficients c_1 at all of the receivers.

Thus, the iterative deconvolution method yields $\hat{\mathbf{f}}$ and $\hat{s}(t)$, estiamtes of $\mathbf{f}$ and $s(t)$, the parameters of the oblique force. As discussed previously, $\hat{\mathbf{f}}$ and $\hat{s}(t)$ may be mulitiplied and divided, respectively, by an arbitrary scale factor. Here we use the convention that $\hat{\mathbf{f}}$ is a unit vector, and thus scale $\hat{s}(t)$ such that

$$\sum_{i=1}^{3} \hat{f}_i^2 = 1 \tag{8}$$

Since $\mathbf{f}$ and $\hat{\mathbf{f}}$ are both unit vectors, the angle between them is given by,

$$\Delta\phi = \cos^{-1}(\mathbf{f} \cdot \hat{\mathbf{f}})$$

This angle is a measure of the error in determining the orientation of the oblique force.

Transducer Characterization. The piezoelectric transducers used in the work reported here are primarily sensitive to normal velocity. They are also small in size such that a point receiver model is appropriate. Thus, we assume that the output of the amplifier $a(t)$ can be expressed as,

$$a(t) = T(t) * \frac{du_3(\mathbf{x}, t)}{dt} = T(t) * \dot{u}_3(t) \tag{10}$$

In this equation, $T(t)$ is the transfer function for the transducer. It characterizes not only the transducer but the coupling of the transducer to the structure and the recording

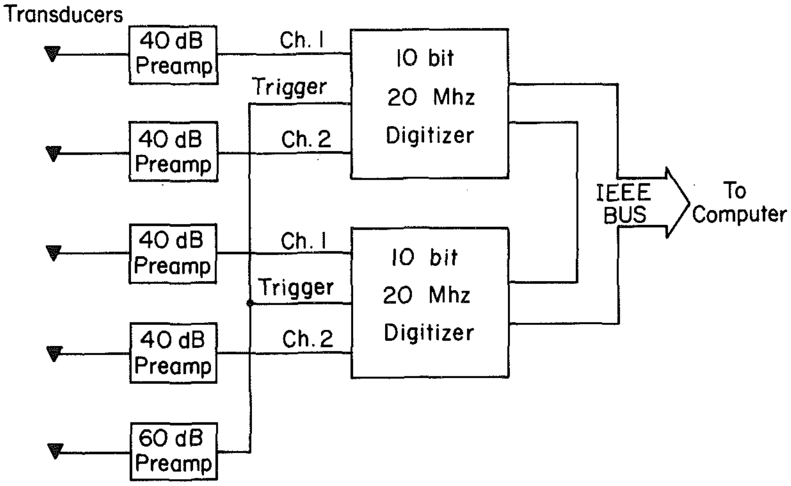


Fig. 2 Data acquisition equipment for oblique force measurements

equipment (cabling, amplifiers, transient recorders). It is further assumed that $T(t)$ is of finite duration.

The first step in the characterization procedure is to determine the transfer function $T(t)$ for each transducer. This is done by measuring the response to the fracture of a glass capillary. This dynamic source is initiated by slowly pressing vertically against a capillary tube until it fractures. As shown by Breckenridge et al. [1967], this source is a concentrated vertical force that has the time dependency of a step-like unloading function. If the finite rise time is neglected, the time dependence of the source is represented by a step function $H(t)$, and we have,

$$\begin{aligned} \mathbf{f} &= -\mathbf{e}_3 \\ s(t) &= H(t) \end{aligned} \qquad (11)$$

The resulting normal velocity at $\mathbf{x} = (r, \theta, h)$ is,

$$\begin{aligned} v_c(t) &= -\frac{d}{dt}\{G_{33}(\mathbf{x}, t; \mathbf{0}) * H(t)\} \\ &= -G_{33}(\mathbf{x}, t; \mathbf{0}) \end{aligned} \qquad (12)$$

If $a_c(t)$ denotes the signal generated by breaking the capillary, equation (10) then yields

$$a_c(t) = T(t) * v_c(t) \qquad (13)$$

Since $a_c(t)$ is measured and $v_c(t)$ is known from calculation of the Green's function G_{33}, the transfer function $T(t)$ can be evaluated by deconvolution.

Now let $a(t)$ be the measured signal from the unknown source. From this signal, the iterative deconvolution procedure recovers a source time function $S(t)$. This $S(t)$ is the convolution of $\dot{s}(t)$, the derivative of the source time function, with the transducer transfer function $T(t)$.

$$S(t) = T(t) * \dot{s}(t) \qquad (14)$$

Since both $S(t)$ and $T(t)$ are known, $\dot{s}(t)$ can be calculated by deconvolution. It must be numerically integrated to obtain the source time function $s(t)$. Therefore, the source time function can be recovered using a transducer that is not a displacement sensor, but that is sensitive to vertical velocity.

If the transfer function is band-limited, those frequency components of $s(t)$ not present in $T(t)$ cannot be recovered by deconvolution. For example, if the transducer is not sensitive to high frequencies, the recovered $s(t)$ will be missing high frequency information and fast rise times cannot be accurately recovered.

3 Experimental Methods

Experiments were performed by generating oblique forces with step-like time functions on the surface of a glass plate. Signals were recorded and analyzed to determine the orientation and time dependence of several forces.

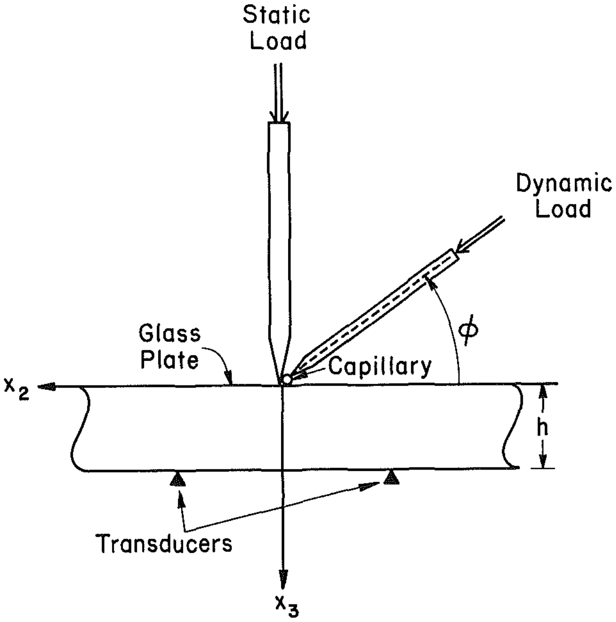


Fig. 3 Experimental setup for generating oblique forces

Specimens and Equipment. A single glass plate was used for all experimental measurements. The plate was approximately 150 mm × 150 mm in extent, and was 18.46 mm thick ($h = 18.46$ mm). The longitudinal and shear wave speeds were measured with a pulse-overlap technique, and were 5.81 mm/μs and 3.46 mm/μs, respectively.

The transducers used to record the wave motion in the plate contained circular piezoelectric crystals 1.35 mm in diameter. The frequency response of these transducers had an upper limit of approximately 1 Mhz. Since the high frequency response of the transducers was negligible, rise times faster than about 1 μs could not be accurately recovered.

The transducer voltage signals were amplified with a gain of 40 dB and then digitized and stored. An additional transducer was used to trigger the digitizers for each channel, as shown in the equipment diagram in Fig. 2. The sampling frequency was 20 Mhz ($\Delta t = 0.05$ μs), and the data were digitized with a resolution of 10 bits. Each recorded signal consisted of 201 points, which corresponds to a time window of 10 μs. This was the longest possible time window that could be recorded with no reflections from the edge of the plate.

Table I Coordinates of transducers for measurements at different radii

Transducer Number	r	θ	z
1	1.03h	0°	1h
2	2.31h	117°	1h
3	3.10h	180°	1h
4	2.06h	256°	1h

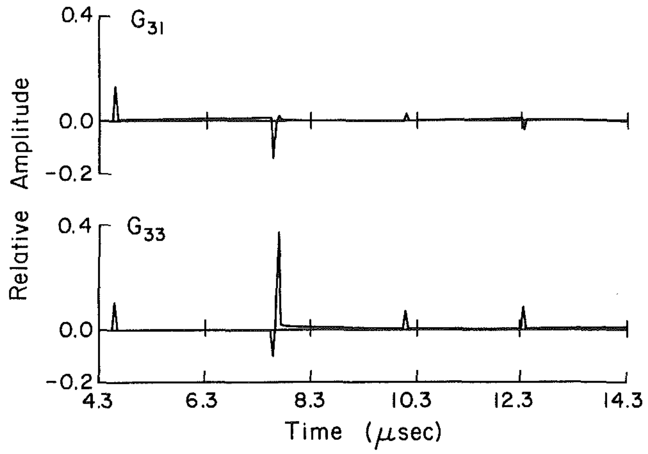


Fig. 4(*a*) Green's functions G_{31} and G_{33} at $x(r, \theta, z)$ = (1.03 *h*, 0, *h*) for a source at x^o = (0, 0, 0)

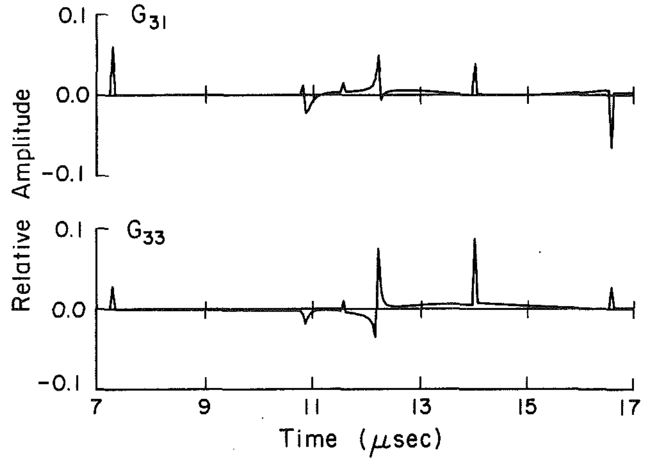


Fig. 4(*b*) Green's functions G_{31} and G_{33} at $x(r, \theta, z)$ = (2.31 *h*, 0, *h*) for a source at x^o = (0, 0, 0)

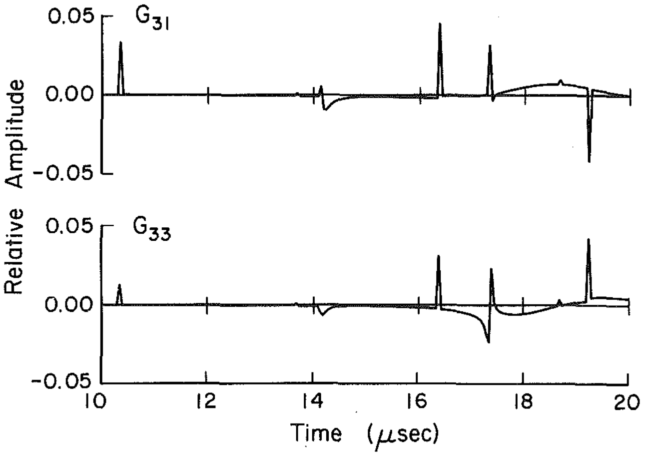


Fig. 4(c) Green's functions G_{31} and G_{33} at $x(r, \theta, z)$ = (3.10 *h*, 0, *h*) for a source at x^o = (0, 0, 0)

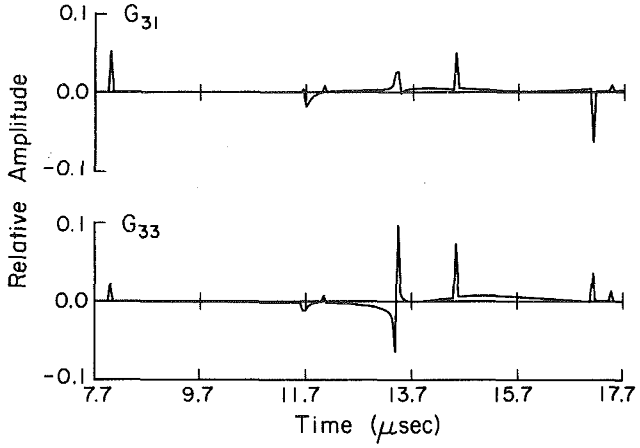


Fig. 4(*d*) Green's functions G_{31} and G_{33} at $x(r, \theta, z)$ = (2.06 *h*, 0, *h*) for a source at x^o = (0, 0, 0)

Oblique Force Generation. The experimental setup for generating oblique forces on the top surface of the plate is shown in Fig. 3. Note that the transducers are located on the bottom surface of the glass plate. The vertical rod, which was 3.18 mm in diameter and had a taper angle at the tip of about 33 degrees, was statically located onto the top plate surface. A glass capillary tube with 0.08 mm o.d. and 0.05 mm i.d. was placed in the corner between the rod tip and the plate along the x_1 axis. A second rod, which was 2.38 mm in diameter and had a taper angle of about 21 degrees, was held at an angle ϕ such that its tip was in contact with the capillary. This rod was slowly loaded along its axis until the capillary fractured. Therefore, the time dependence of this source is a step-like function, and the orientation of the force is along the axis of the second rod but in the opposite direction of the applied force, since the fracture of the capillary unloads the plate.

$$\begin{aligned} \mathbf{f} &= -\cos\phi \mathbf{e}_2 - \sin\phi \mathbf{e}_3 \\ s(t) &= H(t) \end{aligned} \qquad (15)$$

Since the capillary is physically very small, it is quite difficult to position the tip of the second rod exactly on the apex of the capillary. Therefore, the measured angle ϕ may differ from the exact angle of the oblique force by as much as 5 to 10 degrees.

4 Results

Experiments were performed with four transducers that were located at different radial and angular locations. The coordinates of the transducers are listed in Table I. Since the radii are different for each transducer, the Green's functions G_{31} and G_{33} are also different. These Green's functions are shown in Figures 4(*a*)-(*d*). The ordinate is relative displacement, and is consistent for the four functions.

The Green's functions were calculated by numerically differentiating the displacement responses to forces with the time dependence of a Heaviside unit step function. Therefore, G_{31} and G_{33} are not truly the impulse responses, but the responses to a rectangular pulse of unit amplitude and width $\Delta t = 0.05$ μs. Since the width of the pulse is the same as the sampling interval, G_{31} and G_{33} may be treated as though they were impulse responses.

The step function responses were calculated by a computer program developed at Cornell University by R. Gajewski and A. Ceranoglu. The program is based upon the generalized ray theory (Pao and Gajewski, 1977), which obtains the transient displacement signals for a given time in the form of a finite series of integrals. Each integral corresponds to a particular ray path in the plate, and is evaluated numerically by Cagniard's method. The displacement signals obtained by this program are very accurate, and have been experimentally

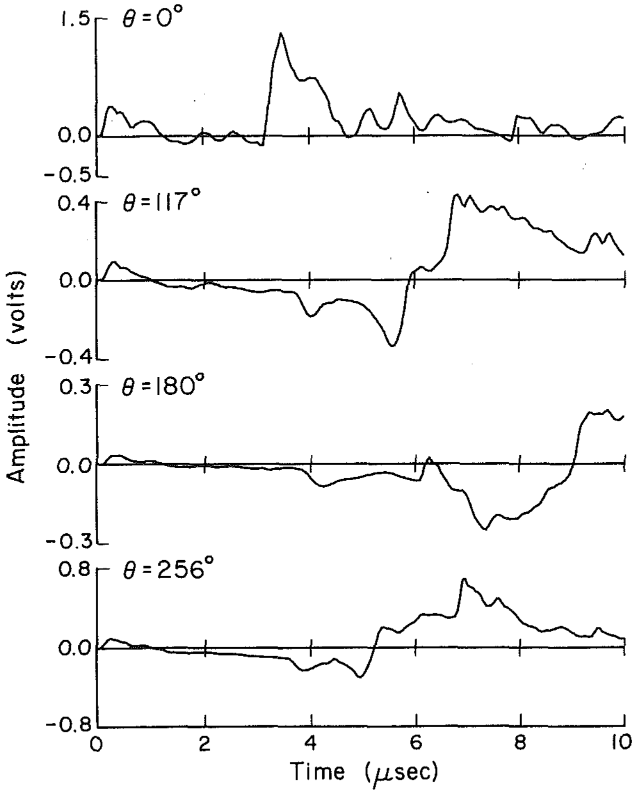


Fig. 5 Recorded waveforms from calibration source

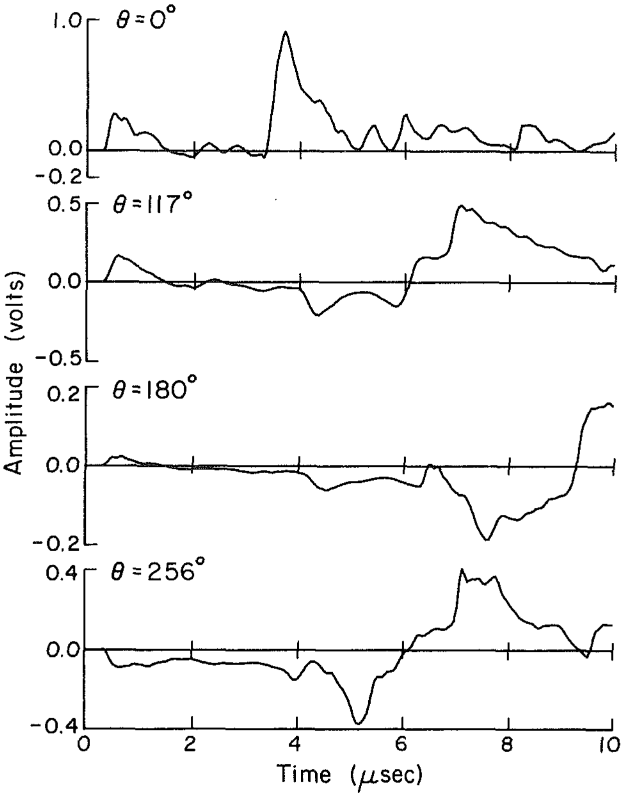


Fig. 7 Recorded waveforms for oblique force at 55 deg

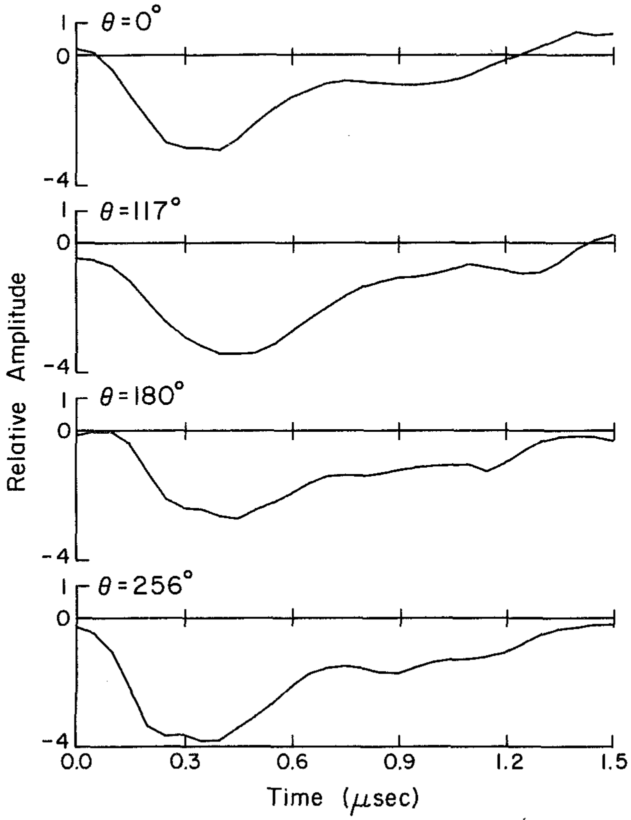


Fig. 6 Transducer transfer functions

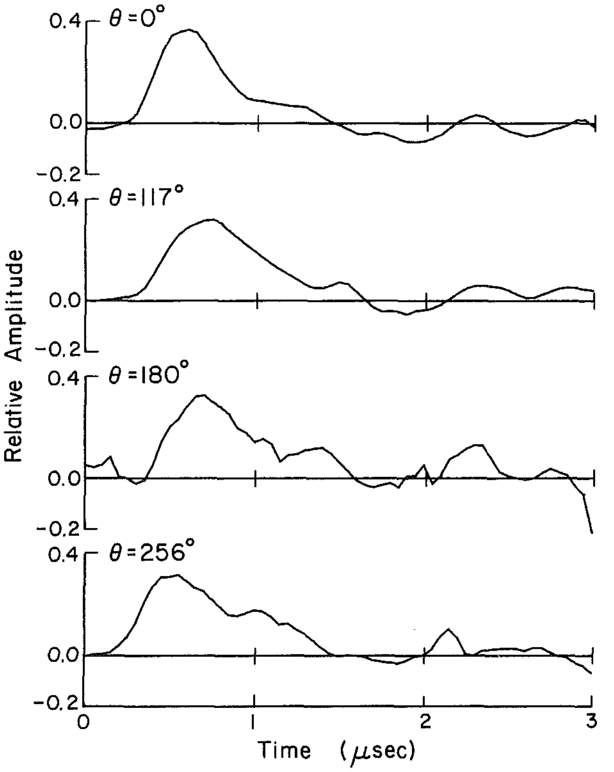


Fig. 8 Convolved functions $S(t) = \dot{s}(t) * T(t)$ recovered from 55 deg oblique force data

verified for vertical forces (Sachse and Ceranoglu, 1979, and Procter et al., 1983).

The four transducers were characterized by velocity transfer functions according to equation (10). A glass capillary 0.08 mm o.d. and 0.05 mm i.d. was broken at $\mathbf{x}^o = (0,0,0)$, and the resulting voltage signals for each transducer were digitized and stored. The average of three signals for each transducer is shown in Fig. 5. These signals were deconvolved to obtain transfer functions for each transducer, which are shown in Fig. 6. These functions are negative in sign because of an inverting amplifier.

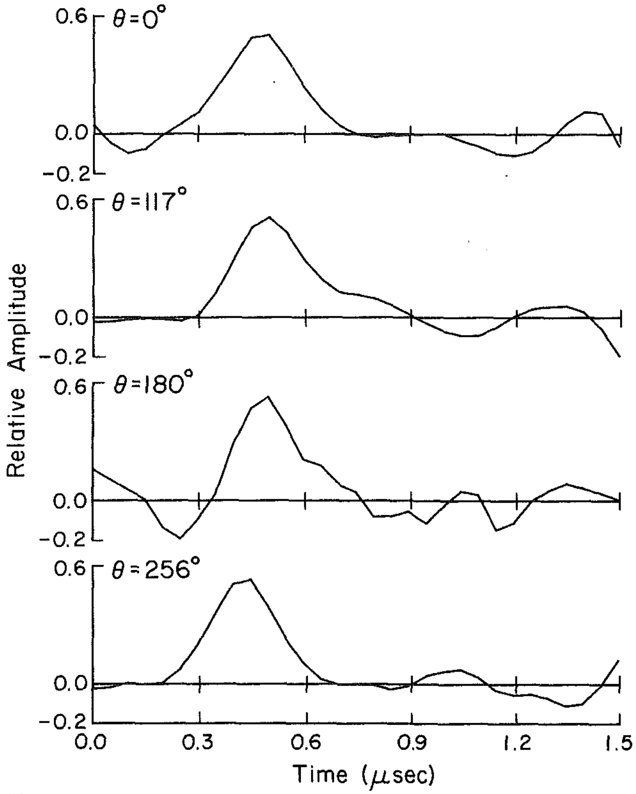


Fig. 9 Differentiated source time functions $\dot{s}(t)$ recovered from 55 deg oblique force data

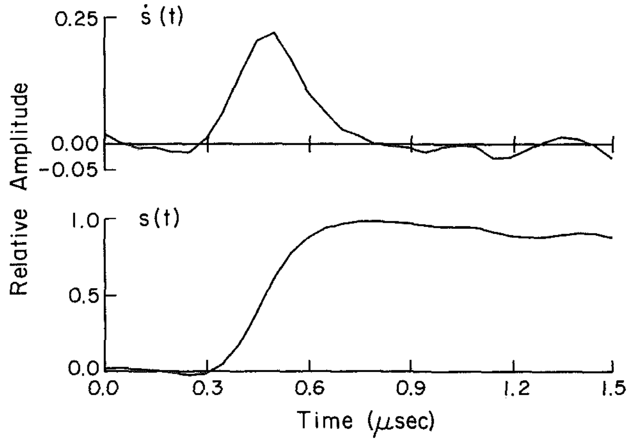


Fig. 10 Averaged function $\dot{s}(t)$ and source time function $s(t)$ recovered from 55 deg oblique force data

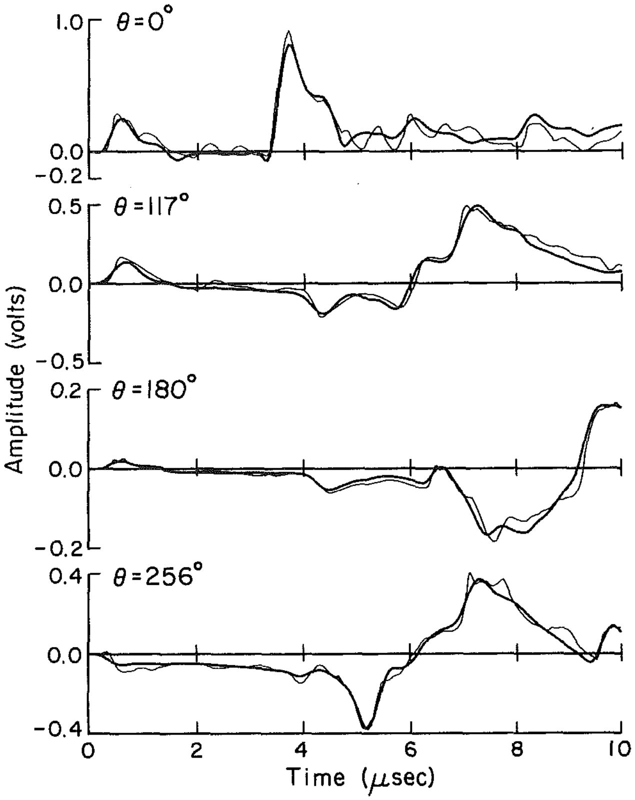


Fig. 11 Fitted transducer signals for 55 deg oblique force data

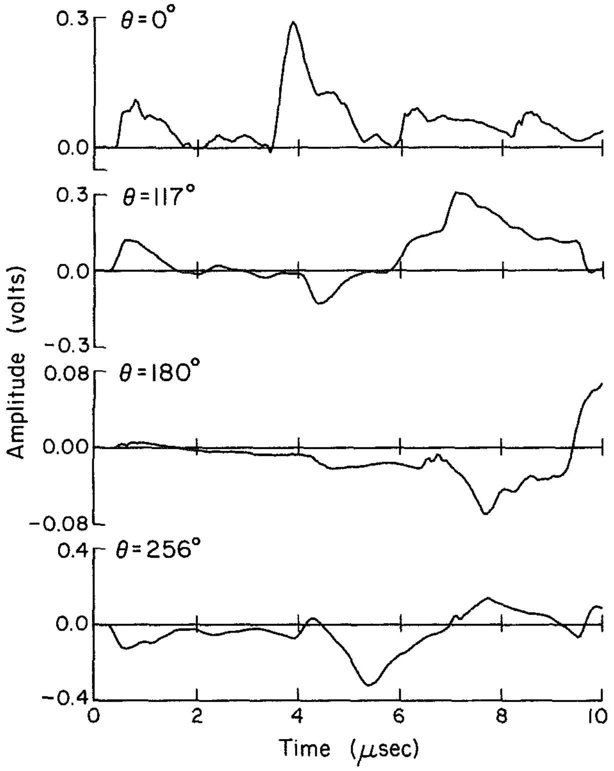


Fig. 12 Recorded waveforms for oblique force at 32 deg

Oblique Force at 55 Deg. The first oblique force was generated at a nominal angle of $\phi = 55$ deg. This corresponds to an orientation of,

$$\mathbf{f} = (0,\ -0.5736,\ 0.8192) \tag{16}$$

The measured signals from each of the transducers are shown in Fig. 7.

These signals were analyzed according to the iterative deconvolution procedure described in Section 2. The recovered time functions are the convolution of the differentiated source time function $\dot{s}(t)$ with the transfer functions $T(t)$, and are shown in Fig. 8. The functions $\dot{s}(t)$ at the four transducer locations were determined by deconvolution, and are shown in Fig. 9. The final estimate of $\dot{s}(t)$ was found by first averaging the signals at the four locations, and then rescaling to correct for the normalization of $\mathbf{f}$. It is shown in Fig. 10 along with $s(t)$, which was obtained by numerical integration. The recovered $s(t)$ is a step-like function with a rise time of approximately 0.3 μs. The actual rise time could be less because of the limited high frequency response of the transducers. The amplitude scale is not absolute but is relative to the magnitude of the calibration source.

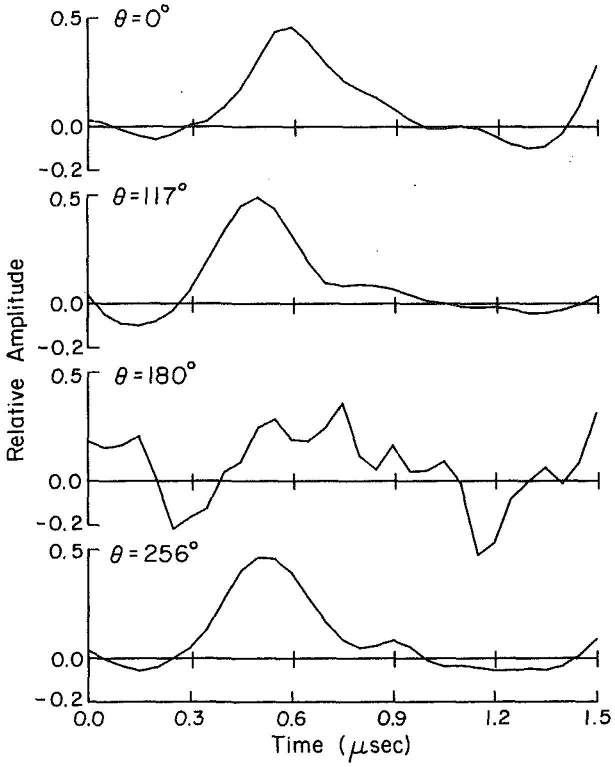


Fig. 13 Differentiated source time functions $\dot{s}(t)$ recovered from 32 deg oblique force data

The recovered orientation vector $\hat{\mathbf{f}}$ after normalization to a unit vector is,

$$\hat{\mathbf{f}}=(-0.0501,\ -0.5658,\ -0.8230) \qquad (17)$$

The angular error between $\mathbf{f}$ and $\hat{\mathbf{f}}$, as calculated by equation (9), is 2.87 deg.

To evaluate how well the recovered $\hat{\mathbf{f}}$ and $s(t)$ model the data, they were used to calculate fitted transducer signals by convolution according to equations (5) and (10). These calculated signals are shown in Fig. 11. They may be compared to the measured signals in Fig. 7, which are shown as light lines in Fig. 11 for comparison. The calculated signals closely match the measured data except for some high frequency information that is not modeled by the recovered parameters.

Oblique Force at 32 Deg. The second oblique force was generated at an angle of $\phi=32$ deg such that,

$$\mathbf{f}=(0,\ -0.8480,\ -0.5299) \qquad (18)$$

The measured signals are shown in Fig. 12. Note that the signal at 180 deg is quite small in amplitude and somewhat noisy compared to the other three signals. This is because the transducer for this signal is located at a null in the radiation field for the horizontal component, and it is also at the largest distance from the source ($r=3.10\ h$).

These signals were analyzed to obtain $\dot{s}(t)$ at each transducer, with results shown in Fig. 13. These $\dot{s}(t)$ are all quite similar except for the one obtained at 180 deg, which is very noisy and bears little resemblance to the others. This is because the measured signal at 180 deg is small in amplitude and has a lower signal-to-noise ratio than the other three. Thus, the final estimate of $\dot{s}(t)$ is the average of the signals from the first, second and fourth transducers only, and is shown in Fig. 14. Also shown in Fig. 14 is $s(t)$, which was obtained by numerical integration. It is a step-like function with a rise time of about 0.5 μs, and is very similar to the source time function obtained for the 55 deg force.

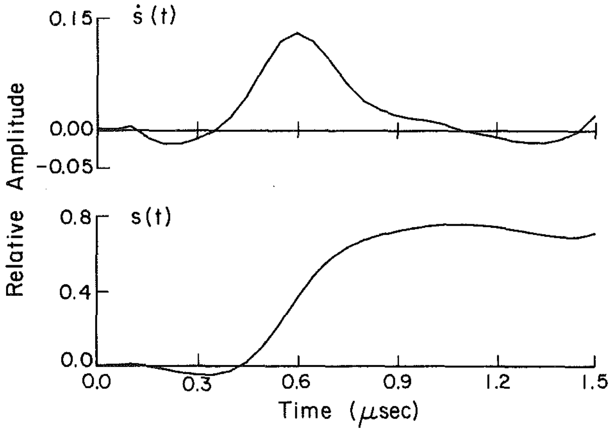


Fig. 14 Averaged function $\dot{s}(t)$ and source time function $s(t)$ recovered from 32 deg oblique force data

The recovered orientation vector $\hat{\mathbf{f}}$ after normalization to a unit vector is,

$$\hat{\mathbf{f}}=(-0.0555,\ -0.8923,\ 0.4480) \qquad (19)$$

The angular error between $\mathbf{f}$ and $\hat{f}$ is 6.25 deg.

5 Summary and Conclusions

In this paper we have presented results that experimentally confirm an inverse method for determining the orientation and time history of dynamic oblique forces. In previous work, time histories of forces with known orientation have been determined by deconvolution techniques. Here, we have solved the problem of simultaneously determining the orientation as well as the time history of an oblique force. The required data are signals recorded at a minimum of two receivers that are sensitive to normal motion.

An important part of the successful demonstration of this inverse method was the development of an experimental procedure to generate controlled oblique forces. The procedure consists of fracturing a glass capillary with a load slowly applied at a known angle to the specimen surface. The resulting dynamic unloading force has a step-like time function and controlled orientation.

The key step in the inversion method is the determination of a source time function $s(t)$ and coefficients c_m of a linear combination of Green's functions, as represented by equation (6). This procedure is not limited to the characterization of oblique forces, but can be applied to many dynamic sources that are separable in time and space. The particular problem that motivated this study is that of characterizing cracks or earthquakes from recorded transient signals. The parameters to be determined for these sources are the time history and moment tensor components of the crack or earthquake. Another problem to which this method can be applied is that of recovering the spatial distribution of a separable extended source of known orientation, as was done by Chung and Sachse [1985] for synthetic data.

Acknowledgments

The authors would like to thank Dr. Kwang-Yul Kim for providing the mechanical fixture for holding the glass plate and the transducers and Dr. Masahiko Hirao for measuring the shear and longitudinal wave speeds of the glass plate. The first author (JEM) would like to acknowledge the support of IBM Corporation in the form of a fellowship to Cornell University, and the second author (YHP) would like to acknowledge the support of the Office of Naval Research (ONR/SRO III Program).

References

Breckenridge, F. R., Tschiegg, C. E., and Greenspan, M. J., 1967, "Acoustic Emission: Some Applications of Lamb's Problem," *J. Acoust. Soc. Am.*, Vol. 57, pp. 1393–1408.

Ceranoglu, A. N., and Pao, Y.-H., 1981, "Propagation of Elastic Pulses and Acoustic Emission in a Plate, Parts I, II, III," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 125–147.

Chang, C., and Sachse, W., 1985, "Analysis of Elastic Wave Signals from an Extended Source in a Plate," *J. Acoust. Soc. Am.*, Vol. 77, pp. 1335–1341.

Goodier, J. N., Jahsman, W. E., and Ripperger, E. A., 1959, "An Experimental Surface-Wave Method for Recording Force-Time Curves in Elastic Impacts," ASME JOURNAL OF APPLIED MECHANICS, Vol. 26, pp. 3–7.

Hsu, N. N., Simmons, J. A., and Hardy, S. C., 1977, "An Approach to Acoustic Emission Signal Analysis," *Materials Evaluation*, Vol. 35, No. 10, pp. 100–106.

Michaels, J. E., Michaels, T. E., and Sachse, W., 1981, "Applications of Deconvolution to Acoustic Emission Signals Analysis," *Materials Evaluation*, Vol. 39, No. 11, pp. 1032–1036.

Michaels, J. E., 1984, "An Inverse Source Problem for Elastic Waves," Ph.D. Thesis, Cornell University, Ithaca, New York.

Michaels, J. E., and Pao, Y.-H., 1985, "The Inverse Problem for an Oblique Force on an Elastic Plate," *J. Acoust. Soc. Am.*, Vol. 77, pp. 2005–2011.

Pao, Y.-H., and Gajewski, R. R., 1977, "The Generalized Ray Theory and Transient Responses of Layered Elastic Solids," *Physical Acoustics*, Vol. 13, Mason, W. P., and Thurston, R. N., eds., Academic Press, New York, pp. 183–265.

Procter, T. M., Breckenridge, F. R., and Pao, Y.-H., 1983, "Transient Waves in an Elastic Plate: Theory and Experiment Compared," *J. Acoust. Soc. Am.*, Vol. 74, pp. 1905–1906.

Sachse, W., and Ceranoglu, A. N., 1979, "Experiments with a Well-Characterized Acoustic Emission System," *Ultrasonics International 1979: Conference Proceedings*, IPC Science and Technology Press, Guildford, U.K., pp. 138–145.

N. Ghosh
Assistant Professor,
Department of Mathematics,
Pennsylvania State University,
University Park, PA

H. Rajiyah
Graduate Student,
Department of Theoretical and Applied Mechanics,
Cornell University,
Ithaca, NY

S. Ghosh
Graduate Student,
Department of Mechanical Engineering and Applied Mechanics,
University of Michigan,
Ann Arbor, MI

S. Mukherjee
Associate Professor,
Department of Theoretical and Applied Mechanics,
Cornell University,
Ithaca, NY
Mem. ASME

A New Boundary Element Method Formulation for Linear Elasticity

A new boundary element formulation for linear elasticity problems is presented in this paper. The standard formulation for planar problems uses two kernels – one of which is logarithmic singular and the other is 1/r *singular, where* r *is the distance between a source and a field point. The new formulation avoids the use of the strongly singular kernel so that both kernels are now only logarithmic singular. The new formulation has several potential advantages over the standard one, the most significant of which is that it delivers stresses accurately at internal points which are extremely close to the boundary of a body. Numerical results for sample problems, from each of the formulations, are presented and compared here.*

Introduction

The boundary element method is based upon classical integral equation formulations of boundary value problems. Although such formulations were originally thought to be primarily of theoretical interest, the contributions of energetic researchers together with the explosive advance of computers have developed the method to the extent that it is today a powerful general purpose procedure for obtaining numerical solutions to many practical problems in science and engineering. In the field of solid mechanics, starting with applications in torsion (Jawson and Ponter, 1963) and then linear elasticity (Rizzo, 1967, Cruse, 1969), the method has been applied to a wide range of problems with material nonlinearities (see, for example, Mukherjee, 1982) and recently, even to elastic-viscoplastic problems with both material and geometrical nonlinearities (Mukherjee and Chandra, 1984). The BEM, today, is a strong competitor to the more widely used finite element method (FEM). Some user friendly commercial codes capable of solving general linear elastic and other problems, based on the BEM, are now available in the market.

As mentioned in the last paragraph, Rizzo (1967) and Cruse (1969) were the first researchers to solve linear elasticity problems by the BEM. Rizzo obtained numerical solutions to two-dimensional and Cruse to three-dimensional problems. Their original BEM formulation for linear elasticity, which is now a standard approach, is based on the equation

$$u_i(p)=\int_{\partial B}[U_{ij}(p,Q)\tau_j(Q)-T_{ij}(p,Q)u_j(Q)]ds_Q \qquad (1)$$

where the displacement and traction components are u_i and τ_i, respectively, with p a source and q a field point (with upper case letters denoting points on the surface ∂B of a body B and lower case letters denoting points in the interior). Further, the kernels U_{ij} and T_{ij}, which are obtained from Kelvin's singular solutions due to point loads in elastic bodies of infinite extent, are singular two point functions. These are available in many references (e.g., Mukherjee (1982)) for both three-dimensional as well as planar (plane strain and plane stress) problems. It should be mentioned here that body forces are assumed to be absent in equation (1), but these can be easily included in the formulation. Also, equation (1) is valid for simply as well as multiply connected domains B. The surface ∂B must, of course, include the outer as well as inner boundaries of B in the latter case.

Equation (1) contains both the boundary tractions and boundary displacements over the entire surface ∂B. Only half of these are prescribed in a well posed problem of linear elasticity. Boundary integral equations for the rest can be obtained by taking the limit of equation (1) as an internal point p approaches a boundary point P. This gives

---

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February 21, 1985; final revision, June 29, 1985.

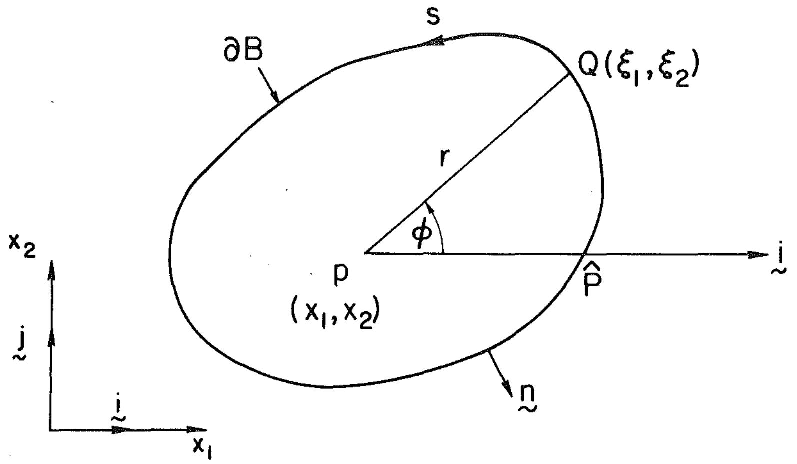


Fig. 1

$$C_{ij}(P)u_j(P)=\int_{\partial B}[U_{ij}(P,Q)\tau_j(Q)-T_{ij}(P,Q)u_j(Q)]ds_Q \tag{2}$$

The new coefficients C_{ij} depend on the local geometry of ∂B at P. It simply equals $(1/2)\delta_{ij}$ (wherre δ_{ij} is the Kronecker delta) if the boundary is locally smooth at P. Otherwise, it can be obtained in closed form for two dimensional problems (see, for example, Mukherjee, 1982)[1] but direct evaluation of C_{ij} appears difficult in three-dimensional situations. A convenient indirect approach, however, is possible in such cases.

In the interest of what is to follow, the explicit forms of the kernels, for plane strain problems, (with $i,j=1,2$) are given below:

$$U_{ij}=\frac{-1}{8\pi(1-\nu)G}[(3-4\nu)\ln r\delta_{ij}-r,_i r,_j] \tag{3}$$

$$T_{ij}=\frac{-1}{4\pi(1-\nu)r}\left[\{(1-2\nu)\delta_{ij}+2r,_i r,_j\}\frac{\partial r}{\partial n}\right.$$
$$\left.-(1-2\nu)(r,_i n_j-r,_j n_i\right] \tag{4}$$

In the above, G is the shear modulus and ν is the Poisson's ratio. Also, r is the Euclidean distance between p and q and a comma denotes a derivative with respect to the corresponding coordinate of a field point. Finally, the components of a unit outward normal to ∂B, at a point on it, are denoted by n_i and $\partial r/\partial n$ is the derivative of r with respect to the normal at a field point.

The above formulation for linear elasticity, as mentioned before, is a standard one and forms the basis for many BEM analyses and computer programs. The singularity of the kernel $T_{ij}-1/r$ for two-dimensional and $1/r^2$ for three-dimensional problems – is, however, strong. Also, using (Fig. 1)

$$\frac{\partial}{\partial s}=-n_2\frac{\partial}{\partial x_1}+n_1\frac{\partial}{\partial x_2} \tag{5}$$

it can be demonstrated that, at any point P on ∂B,

$$r,_j n_i-r,_i n_j=\epsilon_{ij}\frac{\partial r}{\partial s} \tag{6}$$

where the boundary ∂B of a two-dimensional simply connected body is considered for the moment. This boundary, as shown in Fig. 1, has a unit outward normal **n** with components n_i at a point on it and s is the distance, increasing in the counterclockwise direction, measured along ∂B from some fixed point on it. Further, the range of indices in equation (5) is 1,2 and $\epsilon_{11}=\epsilon_{22}=0$, $\epsilon_{12}=-\epsilon_{21}=1$.

The tangential derivative of $\ln r$ is mathematically ill behaved as $r\rightarrow 0$. The strong singularity in T_{ij}, coupled with the presence of the tangential derivative of $\ln r$ in it, is known to cause some numerical problems in the standard formulation. Perhaps the most serious of these is the rapid degeneration in numerical accuracy of the stress and strain components, as one samples these quantities at an internal point very close to the boundary. This phenomenon is sometimes called the "boundary layer" effect.

It is the main purpose of this paper, then, to reformulate the elasticity problem so that the cumbersome kernel T_{ij} is replaced by a kernel with a weaker singularity. This kernel, called W_{ij}, has the same singularity as U_{ij}. The new formulation has tractions and tangential derivatives of the displacements, τ_i and $\partial u_i/\partial s$, as primary variables. It is shown that this new formulation completely eliminates the "boundary layer" effect in some numerical examples, and is expected to do so in general. Also, it is expected that numerical results based upon the new formulation will, in general, be more accurate than those obtained from the standard one. Only two-dimensional plane strain problems are considered in the rest of the paper and the range of indices in subsequent equations, unless otherwise specified, is 1,2. It is possible to carry through these ideas into three-dimensional formulations but the calculations there become more involved.

New BEM Formulation

Simply Connected Region. The plane strain problem of linear elasticity is considered here. The cross-section of the solid has the boundary ∂B in the (x_1,x_2) plane. Given a source point, a corresponding reference point is defined to be an arbitrary point $\hat{P}$ on ∂B. To be specific, for an internal source point p, $\hat{P}$ is here chosen to be the point of intersection of ∂B with the line parallel to the global x_1-axis through p. In case there is more than one point of intersection, the point which is farthest from p is chosen as $\hat{P}$. If, instead of an internal source point p, a boundary source point P is considered, the corresponding reference point $\hat{P}$, for the sake of simplicity, is chosen to coincide with P. The angle ϕ is the angle between the global x_1-axis and the line joining the source and the field points (Fig. 1). It is important to note that ϕ has a discontinuity across $\hat{P}$ with a jump of 2π.

Referring to Fig. 1, it can be shown that

$$\frac{1}{r}\frac{\partial r}{\partial n}=\frac{(\xi_1-x_1)}{r^2}n_1+\frac{(\xi_2-x_2)}{r^2}n_2 \tag{7}$$

and

---

[1] equation $(5.10)_2$ p. 52 in Mukherjee (1982) should read $c_{12}=c_{21}=\frac{\sin 2\gamma\sin\beta}{4\pi(1-\nu)}$

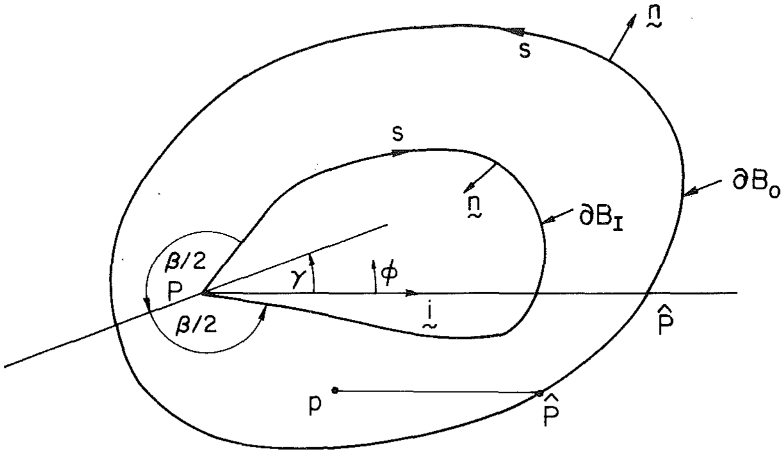


Fig. 2

$$\tan\phi = \frac{\xi_2 - x_2}{\xi_1 - x_1} \tag{8}$$

Differentiating equation (8) with respect to s and obtaining the expression for $\partial\phi/\partial s$, it is revealed that

$$\frac{1}{r}\frac{\partial r}{\partial n} = \frac{\partial\phi}{\partial s} \tag{9}$$

Further, noting that $r,_1 = \cos\phi$ and $r,_2 = \sin\phi$ one can obtain

$$\frac{1}{r}\frac{\partial r}{\partial n} r,_i r,_j = \frac{\partial f_{i,j}{}^{(\phi)}}{\partial s} \tag{10}$$

where

$$f_{11} = \frac{\phi}{2} + \frac{\sin 2\phi}{4};\; f_{12} = f_{21} = \frac{-\cos 2\phi}{4};\; f_{22} = \frac{\phi}{2} - \frac{\sin 2\phi}{4}$$

Using equations (9), (10) and (6) in the expression for T_{ij} (equation (4)), one finds

$$T_{ij} = \frac{d}{ds} R_{ij} \tag{11}$$

where

$$R_{ij} = \frac{-1}{4\pi(1-\nu)}\Big[2(1-\nu)\phi\delta_{ij} + \frac{1}{2}\epsilon_{ij} + \epsilon_{jk} r,_i r,_k$$
$$+ (1-2\nu)\epsilon_{ij} \ln r\Big]$$

Therefore, by partial integration,

$$\oint_{\partial B} T_{ij} u_j ds = \oint_{\partial B} \frac{d}{ds}(R_{ij}) u_j ds = [R_{ij} u_j]_{\phi=0}^{\phi=2\pi}$$
$$- \oint_{\partial B} R_{ij} \frac{\partial u_j}{\partial s} ds$$

Defining

$$W_{ij} = -R_{ij} - \frac{\epsilon_{ij}}{8\pi(1-\nu)},$$

this gives

$$\oint_{\partial B} T_{ij} u_j ds = -u_i(\hat{P}) + \oint_{\partial B} W_{ij} \frac{\partial u_j}{\partial s} ds$$

Therefore, the standard BEM formulation, equation (1), can be recast as

$$u_i(p) - u_i(\hat{P}) = \oint_{\partial B} \left[U_{ij}(p,Q)\tau_j(Q) - W_{ij}(p,Q)\frac{\partial u_j}{\partial s}(Q)\right] ds_Q \tag{12}$$

where the new kernel

$$W_{ij} = \frac{1}{4\pi(1-\nu)}[2(1-\nu)\phi\delta_{ij} + \epsilon_{jk} r,_i r,_k + (1-2\nu)\epsilon_{ij}\ln r] \tag{13}$$

It is important to note that W_{ij} has a discontinuity across $\hat{P}$. Also, W_{ij} has the singularity $\ln r$, as has U_{ij} and the primary variables in equation (12) are the traction components τ_i and the trangential derivatives of displacement components, $\partial u_i/\partial s$.

A boundary integral equation for the new formulation can be obtained by treating equation (2) of the standard formulation in the same manner as above (Appendix 1). The resulting boundary integral equation is

$$\oint_{\partial B} U_{ij}(P,Q)\tau_j(Q) ds_Q = \oint_{\partial B} W_{ij}(P,Q)\frac{\partial u_j}{\partial s}(Q) ds \tag{14}$$

As proved in Appendix 1, the above equation is valid even for a point P lying at a corner of ∂B.

Some potential advantages of the new formulation will be noted here. First, the boundary tensor C_{ij} is absent here. Second, and perhaps more important, a consequence of $u_i(p) - u_i(\hat{P})$ appearing on the left hand side of equation (12) suggests that this difference would be obtained very accurately from this formulation. This should be particularly important in obtaining very accurate displacements at a point p when it is an internal point very close to $\hat{P}$. This issue is discussed further in the section on numerical results which is presented later in the paper.

Multiply Connected Region. The plane strain problem in a multiply connected region is considered next. The situation is depicted in Fig. 2. The cross-section, of course, can have an arbitrary number of cutouts. Only one, with boundary ∂B_I, is shown here. The outer boundary is ∂B_o. The reference point $\hat{P}$ is always chosen to lie on ∂B_o.

The convention used here is that a unit normal must always point away from the body B. Also, the direction of integration is chosen in such a way that the body always lies on the left, i.e, it is counterclockwise on ∂B_o and clockwise on each ∂B_I. Thus, s increases in a counterclockwise direction on ∂B_o and in a clockwise direction on each ∂B_I.

$$\oint_{\partial B} \equiv \oint_{\partial B_O} + \oint_{\partial B_I}$$

Using the above convention, it is easy to show that $T_{ij} = d/ds(R_{ij})$, as before, at every point on ∂B.

The derivation of an equation for an internal point p proceeds as before. The result is

$$u_i(p) - u_i(\hat{P}) = \oint_{\partial B} \Big[U_{ij}(p,Q)\tau_j(Q)$$
$$- W_{ij}(p,Q)\frac{\partial u_j}{\partial s}(Q)\Big] ds_Q \text{ for } p\epsilon B \tag{15}$$

The situation, however, is different for the two cases – one in which $P \in \partial B_o$ and another where $P \in \partial B_I$. The resulting equations are

$$0 = \oint_{\partial B} \left[U_{ij}(P,Q)\tau_j(Q) - W_{ij}(P,Q)\frac{\partial u_j}{\partial s}(Q) \right] ds_Q \text{ for } P \in \partial B_o \tag{16}$$

where $\hat{P}$ on ∂B_o is chosen to coincide with P, and

$$u_i(P) - u_i(\hat{P}) = \oint_{\partial B} \left[U_{ij}(P,Q)\tau_j(Q) - W_{ij}(P,Q)\frac{\partial u_j}{\partial s}(Q) \right] ds_Q \text{ for } P \in \partial B_I \tag{17}$$

where $\hat{P}$ is as in Fig. 2.

Equations (16) and (17) are valid even if P lies at a corner of ∂B_o or ∂B_I. The derivation of equation (17) is given in Appendix 2. Equations (15) and (16) can be derived in an analogous manner.

At first glance, the boundary integral equations (16) and (17) for the multiply connected region problem appear underdetermined since displacements at boundary points appear in these equations in addition to τ_i and $\partial u_i/\partial s$ on ∂B. One must, however, not lose sight of constraint equations of the type

$$u_i(P_2) - u_i(P_1) = \int_{P_1}^{P_2} \frac{\partial u_i}{\partial s} ds$$

between two points P_1 and P_2 lying on the same outer or inner boundary. Inclusion of such constraint equations allow a well posed problem to be solved.

Another interesting observation pertains to rigid body modes. It is well known that if all the tractions over ∂B are prescribed, the standard formulation (equations (1), (2)) delivers a displacement field which is unique only within rigid body translations and rotations. Similarly, the pure traction problem for the new formulation delivers a tangential derivative of the displacement field to within a rigid body rotation. Rigid body translations for the pure traction problem do not affect the values of the tangential derivatives of the displacement components, hence the appearance of displacement differences in equations (12), (15) and (17) of the new formulation. Therefore, the new formulation has an advantage in that it has less indeterminancy relative to the standard one.

Strains and Stresses. The strain field throughout the body must be obtained by differentiating the displacement field. A convenient approach is to analytically differentiate equation (15) at an internal source point p in order to obtain the displacement gradient there. This gives

$$u_{i,L}(p) = \oint_{\partial B} \left[U_{ij,L}(p,Q)\tau_j(Q) - W_{ij,L}(p,Q)\frac{\partial u_j}{\partial s}(Q) \right] ds_Q \tag{18}$$

An upper case letter following a comma in the above equation denotes differentiation with respect to a source point. Using the identity

$$r,_L = -r,_l$$

the differentiated kernels can be written in terms of field point derivatives as follows

$$U_{ij,L} = -U_{ij,l} = \frac{1}{8\pi(1-\nu)Gr}[-r,_i\delta_{jl} - r,_j\delta_{li} + (3-4\nu)r,_l\delta_{ij} + 2r,_ir,_jr,_l] \tag{19}$$

$$W_{ij,L} = -W_{ij,l} = \frac{-1}{4\pi(1-\nu)r}[2(1-\nu)r\phi,_l\delta_{ij} - 2\epsilon_{jk}r,_ir,_kr,_l + (1-2\nu)r,_l\epsilon_{ij} + \epsilon_{jk}(r,_k\delta_{il} + r,_i\delta_{kl})] \tag{20}$$

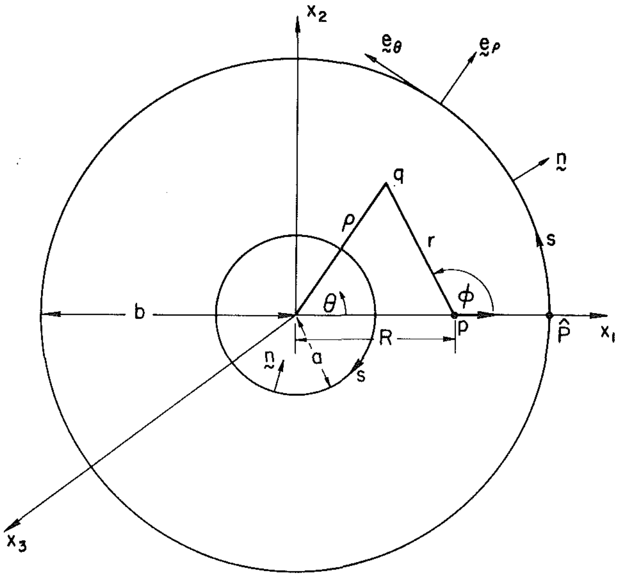


Fig. 3 Geometry of the cylinder problem

The stress at an internal point can then be easily obtained using Hooke's law.

The fact that equation (18) only involves kernels with singularity of the type $1/r$ results in very accurate numerical determination of displacement gradients – even when the point p lies very close to the boundary ∂B. Thus, the so called "boundary layer" essentially vanishes if this new formulation is used. This aspect is discussed further in a later section with numerical results for some sample problems.

The stress components at a boundary point where ∂B is locally smooth can be most conveniently determined from the boundary data, i.e., τ_i and $\partial u_i/\partial s$ at that point, together with Hooke's law. This approach is essentially the same as that described by Rizzo and Shippy (1968) except that there is no longer the need for the numerical evaluation of $\partial u_i/\partial s$. Since numerical differentiation is intrinsically an unstable procedure and generates large round-off errors, the new formulation has an enormous advantage over the standard one in this respect. In actual computations with the standard approach, shape functions for displacements are differentiated with respect to s. This results in lower order shape functions for $\partial u_i/\partial s$ and reduces accuracy. Thus, in any case, it is of great advantage to obtain $\partial u_i/\partial s$ directly.

An Illustrative Analytical Example

The analytical example considered next is the plane strain problem for a long cylinder, of internal radius a and external radius b, subjected to internal and external pressures p_i and p_o, respectively.

The cross-section of the cylinder is shown in Fig. 3. The location of a general field point is given by polar coordinates (ρ,θ). Due to axisymmetry and plane strain, it is sufficient to locate the source point at (R,O). The displacement vector is written as

$$\mathbf{u} = \mathbf{e}_R u_R(R) = \mathbf{e}_R u(R) \tag{21}$$

A unit load is applied at p in the x_1 direction as shown in Fig. 3. The integral formula, equation (15), for an internal point p, becomes

$$u(R) - u(b) = \oint_{\substack{\text{cylinder} \\ p=a}} \left(U_{1\rho}\tau_\rho + \frac{W_{1\theta}}{\rho} u_\rho \right) ds + \oint_{\substack{\text{cylinder} \\ p=a}} \left(U_{1\rho}\tau_\rho - \frac{W_{1\theta}}{\rho} u_\rho \right) ds \tag{22}$$

Table 1 Integrals for cylinder problem

$f(\theta)$	$\int_o^{2\pi} f(\theta)d\theta$		
	$\rho > R$	$\rho = R$	$\rho < R$
$\cos\theta \ln r$	$-\frac{\pi R}{\rho}$	$-\pi$	$\frac{-\pi\rho}{R}$
$\frac{(\rho\cos\theta - R)(\rho - R\cos\theta)}{r^2}$	$\frac{-\pi R}{\rho}$	$-\pi$	$\frac{-\pi\rho}{R}$
$\phi\sin\theta$	$-2\pi+\frac{\pi R}{\rho}$	$-\pi$	$\frac{-\pi\rho}{R}$

On the outer circle, $\rho = b$, $\partial/\partial s = (1/\rho)(\partial/\partial\theta)$, $\partial\mathbf{u}/\partial s = (\mathbf{e}_\theta/\rho)u_\rho$, and $W_{1j}(\partial u_j/\partial s) = (W_{1\theta}/\rho)u_\rho$. But on the inner circle, $\rho = a$, $\partial/\partial s = (-1/\rho)(\partial/\partial\theta)$ so that $W_{1j}(\partial u_j/\partial s) = (-W_{1\theta}/\rho)u_\rho$. Therefore, the signs on the second terms are different in the two integrals in equation (22).

Now, on the outer circle $s = b\theta$ but on the inner circle $s = a(2\pi - \theta)$. The limits on s are 0 to $2b\pi$ on the outer circle and 0 to $2a\pi$ on the inner circle. Using these, equation (22) becomes

$$\begin{aligned} u(R) - u(b) \;&=\; ap_i \int_o^{2\pi} U_{1\rho}(R;\, a,\, \theta)d\theta \\ &- \; bp_o \int_o^{2\pi} U_{1\rho}(R;\, b,\, \theta)d\theta \\ &+ \; u(a) \int_o^{2\pi} W_{1\theta}(R;\, a,\, \theta)d\theta \\ &- \; u(b) \int_o^{2\pi} W_{1\theta}(R;\, b,\, \theta)d\theta \end{aligned} \tag{23}$$

where

$$U_{1\rho}(R;\rho,\theta) = \frac{-1}{8\pi(1-\nu)G}\left[(3-4\nu)\cos\theta\ln r - \frac{(\rho\cos\theta - R)(\rho - R\cos\theta)}{r^2}\right. \tag{24}$$

and

$$W_{1\theta}(R;\rho,\theta) = \frac{1}{4\pi(1-\nu)}\left[-2(1-\nu)\phi\sin\theta + (1-2\nu)\cos\theta\ln r - \frac{(\rho\cos\theta - R)(\rho - R\cos\theta)}{r^2}\right] \tag{25}$$

with $\mathbf{r}^2 = \rho^2 + R^2 - 2\rho R\cos\theta$.

In order to solve for $u(R)$, equation (23) is first used with $R = a$ and then with $R = b$. The required integrals are shown in Table 1. The resulting equations are

$$bu(a) - au(b) = \frac{ab}{2G}(p_i - p_o) \tag{26}$$

$$au(a) - bu(b) = \left(\frac{1-2\nu}{2G}\right)(b^2p_o - a^2p_i) \tag{27}$$

These equations are solved for $u(a)$ and $u(b)$, and then the boundary displacements are used in equation (23) for an internal points, with $a < R < b$. This yields the well known Lame solution (Timoshenko and Goodier, 1970)

$$u(R) = AR + B/R \tag{28}$$

where $A = (1-2\nu)(b^2p_o - a^2p_i)/2G(a^2 - b^2)$ and $B = a^2b^2(p_o - p_i)/2G(a^2 - b^2)$.

Numerical Results

Numerical implementation of equations (15)-(17) is carried out in standard fashion. The boundary ∂B must be discretized into boundary elements. Suitable shape functions must now be chosen for the variation of tractions and tangential derivatives of displacements on the boundary elements. The numerical results presented next are obtained by using straight boundary elements with piecewise linear representations of τ_i and $\partial u_i/\partial s$ on these elements. Also, the kernels, for both the standard and new formulations, are integrated in closed form over boundary elements.

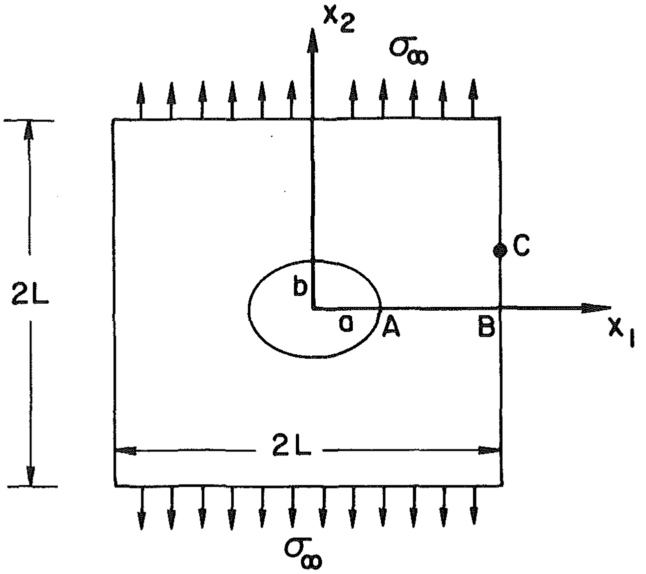


Fig. 4 Square plate with elliptical cutout

One way to model jumps in normals at corners of ∂B in the standard formulation is to put double nodes at corners (see, for example Mukherjee, 1982). This procedure cannot be used in the new formulation as it would lead to singular matrices in the system of algebraic equations obtained by discretization of equation (16). The corner problem, however, can be easily taken care of by placing sampling points away from a corner itself in the boundary elements that meet at a corner.

Numerical results for a sample plane strain problem of a square cross-section with an elliptical hole (Fig. 4) have been obtained by the standard and new BEM formulations. The loading is tensile in the x_2 direction with stress σ_∞. Using quarter symmetry, a quarter of the region is modeled here so that the region involved becomes a simply connected one. The boundary mesh used here is very similar to that shown in Fig. 5.8 of Mukherjee (1982). The same mesh has been used for both formulations. Piecewise linear representations for u_i and τ_i (for the standard formulation) and $\partial u_i/\partial s$ and τ_i (for the new formulation) have been employed here.

The calculated variations of dimensionless displacement u_2 and stress σ_{22} along the line AB, for various values of a/b, obtained from the two methods, are shown in Figs. (5-8). In these examples, $L/a = 10$ and $\nu = 0.3$. The results from the two formulations agree well. Both formulations model the stress concentrations at A, for the various cases, adequately. The new formulation requires about 23 percent more computer time than the standard one.

It has been mentioned earlier that since $W_{ij,l}$ in equation (20) is only $1/r$ singular, while $T_{ij,l}$ in the standard formulation has a $1/r^2$ singularity, it is expected that the new formulation should be able to overcome the so-called "boundary layer effect" and deliver accurate stresses at an internal point very close to the point C, with coordinates (10,2), in Fig. 4. The results are summarized in Table 2. It is seen, in dramatic fashion, that the boundary layer effect vanishes in the new formulation. In fact, the point with $x_1/a = 10 - 5\times10^{-12}$, $x_2/a = 2$, for the case $a/b = 1$ has been tested also and the new formulation yields $\sigma_{22}/\sigma_\infty = 0.95303$. It should be noted that

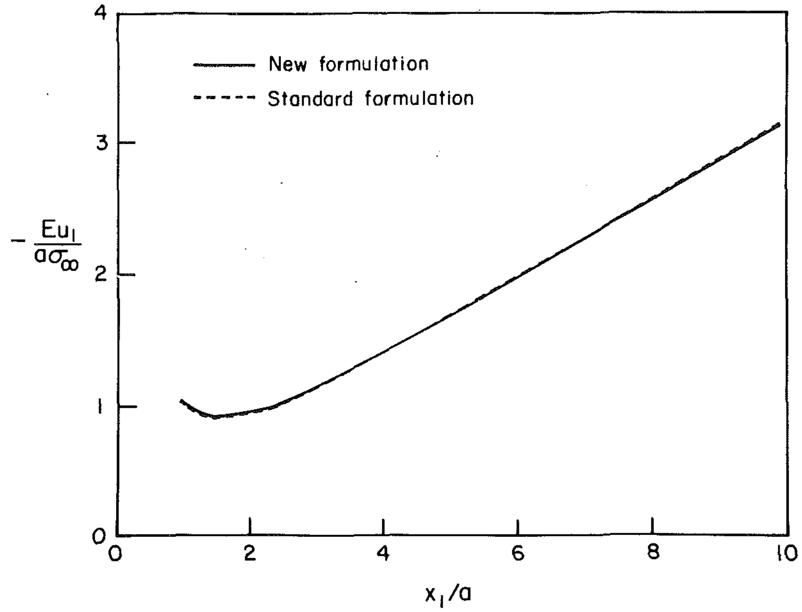


Fig. 5 Displacement variation along *AB* from two BEM formulations; *a/b* = 4

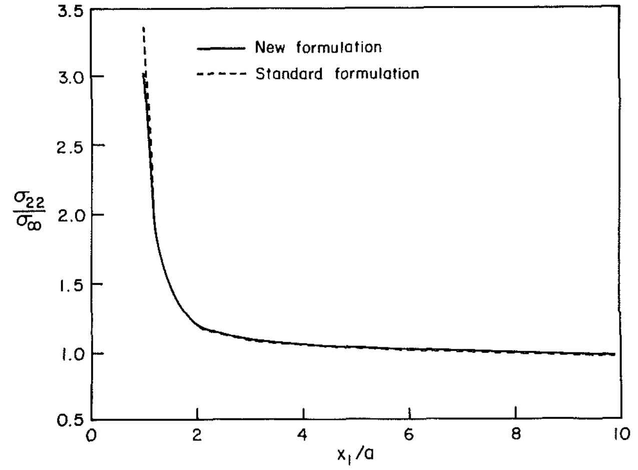


Fig. 6 Stress variation along *AB* from two BEM formulations; *a/b* = 1

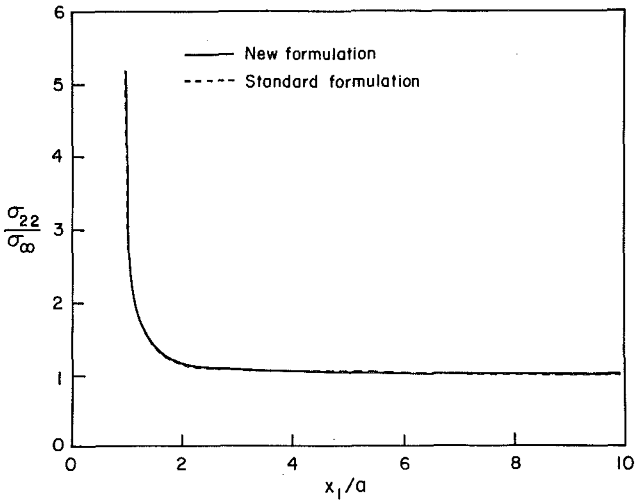


Fig. 7 Same situation as in Fig. 6; *a/b* = 2

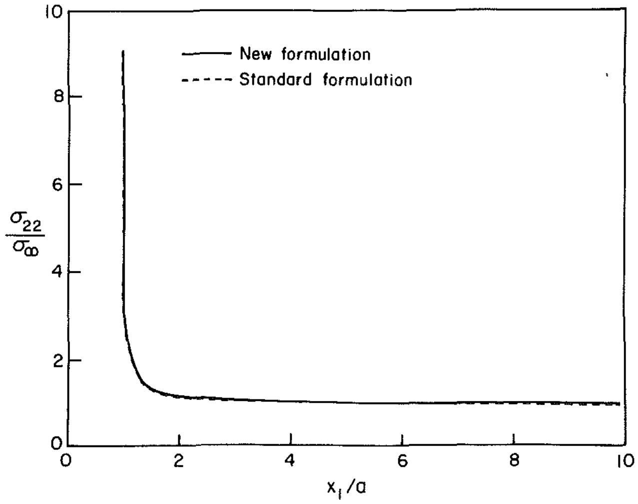


Fig. 8 Same situation as in Fig. 6; *a/b* = 4

Table 2 Stresses near point *C* in Fig. 4 from two BEM formulations

x_1/a	x_2/a	$\sigma_{22}/\sigma_\infty$ (New formulation)	$\sigma_{22}/\sigma_\infty$ (Standard formulation)
		$a/b=1$	
9.95	2.0	0.95303	0.97411
9.995	2.0	0.95303	0.96831
9.9995	2.0	0.95303	0.96278
9.99995	2.0	0.95303	0.95728
9.999995	2.0	0.95303	0.1037×10^6
9.9999995	2.0	0.95303	0.1037×10^7
		$a/b=2$	
9.95	2.0	0.95953	0.98369
9.995	2.0	0.95953	0.97951
9.9995	2.0	0.95953	0.97552
9.99995	2.0	0.95953	0.97155
9.999995	2.0	0.95953	0.1009×10^6
9.9999995	2.0	0.95953	0.1009×10^7
		$a/b=4$	
9.95	2.0	0.95662	0.98875
9.995	2.0	0.95662	0.982237
9.9995	2.0	0.95662	0.98541
9.99995	2.0	0.95662	0.98223
9.999995	2.0	0.95662	0.99364×10^5
9.9999995	2.0	0.95662	0.9937×10^6

the boundary layer effect in the standard formulation occurred despite the fact that the kernels were integrated analytically over boundary elements.

Conclusions

The new BEM formulation, presented in this paper, has several potential advantages relative to the standard one. Some of these are:

(*a*) The strongly singular kernel T_{ij} is avoided here so that both kernels are only logarithmic singular. Since T_{ij} is a strongly singular kernel, the resulting integral

$$\int_{\partial B} T_{ij}(P,Q)u_j(Q)dS_Q$$

is a singular integral in the sense of Calderon-Zygmund (1952), where it is assumed that the Cauchy principal value of the integral is taken. Since, in the numerical implementation, it is common to take nonuniform subdivisions of the boundary, the value obtained is generally different from the Cauchy principal value and this could result in substantial errors. The boundary integral equation (14) involves only integrable kernels and this problem does not arise.

(*b*) It should be noted that the kernel U_{ij} is such that if some $\tau_i(P)$ is unknown, the corresponding row in the coefficient matrix for the BEM, although not being diagonally dominant, has the property that the diagonal term is large, (in absolute value), relative to other terms in that row. Such, however, is not true for the kernel T_{ij}, where the presence of the strong singulary in $1/r\ \partial r/\partial s$ causes some off-diagonal elements to blow up as the mesh size goes to zero. This undesirable property of T_{ij} can negate any advantage derived from the presence of a free term in the standard formulation. In contrast, the kernel W_{ij}, like U_{ij}, only produces bounded elements in the BEM coefficient matrix, although, the presence of ϵ_{ij} shifts the largest terms off the main diagonal. Thus, although neither BEM formulation produces a diagonally dominant coefficient matrix, it is expected that the new formulation will produce a more well-conditioned coefficient matrix relative to the standard one.

(*c*) Another important consequence of the above is that the kernels for stresses are only $1/r$ singular. This fact is demonstrated, in an illustrative numerical example, to overcome the "boundary layer problem" and deliver the stresses accurately at internal points that are extremely close to the boundary of a body. It is expected that the new formulation will be able to overcome the "boundary layer problem" in all cases.

(*d*) The tractions τ_i and displacement derivatives $\partial u_i/\partial s$ are the primary variables in the new formulation. Choice of a particular shape function for $\partial u_i/\partial s$ is essentially equivalent to using a higher order shape function for u_i at no extra cost.

(*e*) The corner tensor C_{ij} is absent in the new formulation and need not be calculated.

The new formulation, like the standard one, can be easily extended to solve problems with material nonlinearities. In fact, since all that is involved here is partial integration of the term $T_{ij}u_j$, the domain integrals associated with plasticity (Mukherjee, 1982) remain unaltered. The new formulation can also be extended to three-dimensional problems. The calculations for the three-dimensional case, however, become more involved.

Acknowledgments

The contributions of H. Rajiyah, S. Ghosh and S. Mukherjee to this research have been supported by Grant No. CEE-8206344 of the US National Science Foundation to Cornell University.

References

Calderon, A. P., and Zygmund, A., 1952, "On the Existence of Certain Singular Integrals," *Acta Mathematica*, Vol. 88, pp. 85–139.

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

Jaswon, M. A., and Ponter, A. R., 1963, "An Integral Equation Solution of the Torsion Problem," *Proceedings of the Royal Society*, London, Series A, Vol. 273, pp. 237–246.

Mukherjee, S., 1982, *Boundary Element Methods in Creep and Fracture*, Elsevier Applied Science Publishers, Barking, Essex, England.

Mukherjee, S., and Chandra, A., 1984, "Boundary Element Formulation for Large Strain-Large Deformation Problems of Plasticity and Viscoplasticity," *Developments in Boundary Element Methods—3*, Banerjee, P. K., and Mukherjee, S., eds., Elsevier Applied Science Publishers, Barking, Essex, England, pp. 27–58.

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

Rizzo, F. J., and Shippy, D. J., 1968, "A Formulation and Solution Procedure for the General Nonhomogeneous Elastic Inclusion Problem," *International Journal of Solids and Structures*, Vol. 4, pp. 1161–1179.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd ed., McGraw-Hill.

APPENDIX 1

Proof of Equation (14)

Referring to Fig. 9 and considering P at a corner of the outer boundary ∂B_o (the angles β and γ are defined in Mukherjee (1982) p. 52),

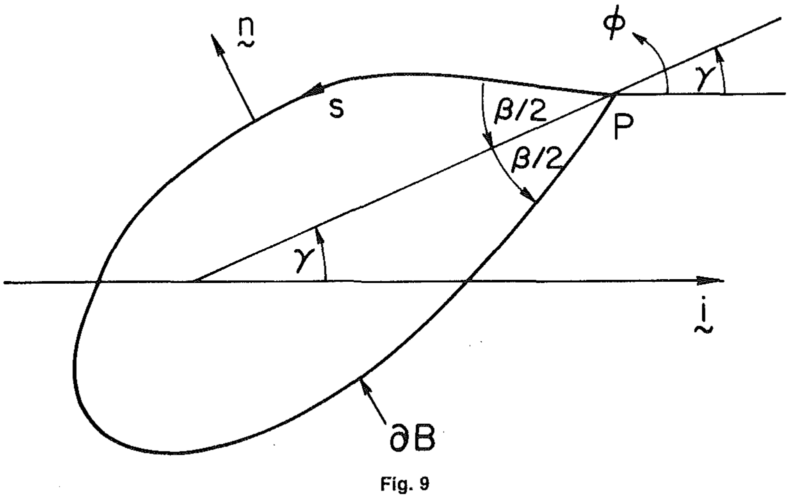


Fig. 9

$$\oint_{\partial B} T_{ij}u_j ds = \Big[R_{ij}u_j\Big]_P^P - \oint_{\partial B} R_{ij}\frac{\partial u_j}{\partial s}ds$$

The boundary term is

$$\Big[R_{ij}u_j\Big]_P^P = \frac{-1}{4\pi(1-\nu)}\Big[2(1-\nu)\phi u_i + \epsilon_{jk}r,_i r,_k u_j\Big]_{\phi=\gamma+\pi-\beta/2}^{\phi=\gamma+\pi+\beta/2}$$

For $i=1$, the above expression becomes

$$\begin{aligned}\Big[R_{1j}u_j\Big]_P^P &= -\Big[\frac{\beta}{2\pi}+\frac{\cos 2\gamma\sin\beta}{4\pi(1-\nu)}\Big]u_1(P)\\ &\quad -\frac{\sin 2\gamma\sin\beta}{4\pi(1-\nu)}u_2(P) = -C_{11}(P)u_1(P)\\ &\quad -C_{12}(P)u_2(P) \qquad (4)\end{aligned}$$

Similarly, for $i=2$

$$\begin{aligned}\Big[R_{2j}u_j\Big]_P^P &= -\frac{\sin 2\gamma\sin\beta}{4\pi(1-\nu)}u_1(P) - \Big[\frac{\beta}{2\pi}\\ &\quad -\frac{\cos 2\gamma\sin\beta}{4\pi(1-\nu)}\Big]u_2(P) = -C_{21}(P)u_1(P)\\ &\quad -C_{22}(P)u_2(P)\end{aligned}$$

Thus, $\Big[R_{ij}u_j\Big]_P^P = -C_{ij}(P)u_j(P)$ and $\oint_{\partial B} T_{ij}(P,Q)u_j(Q)ds_Q$

$$= -C_{ij}(P)u_j(P) + \oint_{\partial B} W_{ij}(P,Q)\partial u_j/\partial s(Q)ds_Q$$

This equation, together with equation (2), immediately gives equation (14).

APPENDIX 2

Proof of Equation (17)

Referring to Fig. 2 and considering P at a corner of the cutout ∂B_I,

$$\begin{aligned}\oint_{\partial B} T_{ij}u_j ds &= \oint_{\partial B_O} T_{ij}u_j ds + \oint_{\partial B_I} T_{ij}u_j ds\\ &= \Big[R_{ij}u_j\Big]_{\hat P}^{\hat P} + \Big[R_{ij}u_j\Big]_P^P - \oint_{\partial B_O} R_{ij}\frac{\partial u_j}{\partial s}ds\\ &\quad -\oint_{\partial B_I} R_{ij}\frac{\partial u_j}{\partial s}ds\end{aligned}$$

The boundary terms are

$$\Big[R_{ij}u_j\Big]_{\hat P}^{\hat P} = \Big[R_{ij}u_j\Big]_{\phi=0}^{\phi=2\pi} = -\delta_{ij}u_j(\hat P)$$

$$\Big[R_{ij}u_j\Big]_P^P = -\frac{1}{4\pi(1-\nu)}\Big[2(1-\nu)\phi u_i + \epsilon_{jk}r,_i r,_k u_j\Big]_{\phi=\gamma+\pi-\beta/2}^{\phi=\gamma-\pi+\beta/2}$$

This time, for $i=1, 2$ (see Appendix 1)

$$\Big[R_{1j}u_j\Big]_P^P = u_1(P) - C_{11}(P)u_1(P) - C_{12}(P)u_2(P)$$

$$\Big[R_{2j}u_j\Big]_P^P = -C_{21}(P)u_1(P) - C_{22}(P)u_2(P) + u_2(P)$$

so that $\Big[R_{ij}u_j\Big]_P^P = \delta_{ij}u_j(P) - C_{ij}u_j(P)$

Thus,

$$\begin{aligned}\oint_{\partial B} T_{ij}u_j ds &= -\delta_{ij}u_j(\hat P) + \delta_{ij}u_j(P) - C_{ij}u_j(P)\\ &\quad -\oint_{\partial B} R_{ij}\frac{\partial u_j}{\partial s}ds\end{aligned}$$

This equation, together with equation (2), gives equation (17).

An Analysis of Large Strain Viscoplasticity Problems Including the Effects of Induced Material Anisotropy

A. Chandra
Assistant Professor,
Aerospace and Mechanical
Engineering Department,
University of Arizona,
Tucson, AZ 85721
Assoc. Mem. ASME

S. Mukherjee
Associate Professor,
Department of Theoretical and
Applied Mechanics,
Cornell University,
Ithaca, NY 14853
Mem. ASME

This paper examines the modeling of large shearing of solids that exhibit induced anisotropy during inelastic deformation. The "traditional" approach uses integration of material rates of certain tensors which are obtained from Jaumann rates of these tensors delivered by a material constitutive model. This leads to erroneous results (spurious oscillations) in a simple shear example. Several previous authors have suggested resolutions to this dilemma based on modification of the constitutive model – usually based upon changing the interpretation of the tensor rates delivered by a constitutive model. This paper draws attention to another aspect of the modeling process – that of obtaining the components of tensors such as the Cauchy stress in a global, space-fixed basis, from the objective rates of these tensors as delivered by the material constitutive model. In essence, it is suggested here that the elastic rotation rather than the spin should be used to achieve the above objective. The rotation idea is first discussed in the context of a simple shear example. This philosophy is then incorporated in a general purpose two-dimensional boundary element method (BEM) formulation and computer program. Numerical results for the simple shear problem, using the rotation idea, are obtained both by direct integration and from the general BEM computer program.

Introduction

The subject of this paper is a close examination of the proper way of generalizing conventional small strain inelastic constitutive models for materials exhibiting induced anisotropy, to cases where large strains and rotations are present. It is clear that the rates of certain tensors, such as that of the back stress in a small strain, small rotation kinematically hardening plasticity constitutive model, must be interpreted as suitable objective rates in a generalized large-strain, large-rotation version of that material model. This is necessary in order to maintain rotational frame indifference of the generalized model.

The "traditional" approach, for such problems, has been the identification of such an objective rate as the Jaumann rate ($\overset{*}{T}_{ij}$) of the corresponding tensor T_{ij}. The next step has been to obtain the material rate from the Jaumann rate by taking account of the rotations, i.e., from the equation

$$\dot{T}_{ij} = \overset{*}{T}_{ij} - T_{ik}\omega_{kj} + \omega_{ik}T_{kj} \tag{1}$$

where the material rate

$$\frac{DT_{ij}}{Dt} = \dot{T}_{ij} = \frac{\partial T_{ij}}{\partial t} + \frac{\partial T_{ij}}{\partial x_k} v_k. \tag{2}$$

In the above, x_k are the coordinates of a material point in the current configuration; v_k are its velocity components and

$$\omega_{ij} = (1/2)(v_{i,j} - v_{j,i}) \tag{3}$$

is the spin of a differential material element containing the point.

The final step in the "traditional" approach involves the integration of the material rate $\dot{T}_{ij}$ with time, in order to determine the evolution of the tensor T_{ij} in a global, spatially fixed coordinate frame.

In a recent paper, Nagteegal and de Jong (1982) used the above approach to numerically analyze a very simple boundary value problem – the case of simple shear for large plastic deformations using a Mises-type kinematic hardening constitutive model. Their calculations lead to the surprising result that use of a kinematic hardening rule of the Prager-Ziegler (1959) type results in the prediction of an oscillatory shearing stress in response to a monotonically increasing shearing strain. In fact, in an earlier paper, Dienes (1979) had observed the same phenomenon using a hypo-elastic material model where T_{ij} in equations (1) and (2) was the Cauchy stress rather than the back stress.

Several researchers such as Lee et al. (1983), Defalias (1983), Dienes (1979), Atluri (1984) and Johnson and Bamman (1984) have examined the above situation. A recent

---

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March 4, 1985; final revision, September 5, 1985.

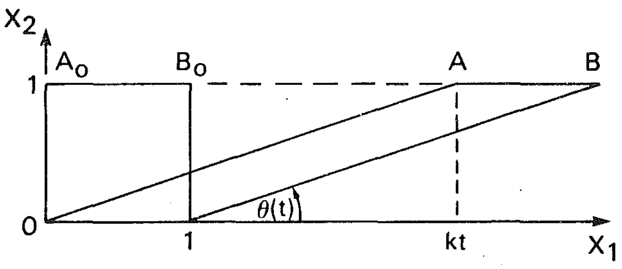


Fig. 1 Simple shear in the x_1 direction

ASME publication (Willam, 1984), which contains papers on this subject by several of these and other authors, summarizes the current debate in excellent fashion. In essence, in the simple shear example (Fig. 1), while the angular velocity of a line of material points depends only on its current orientation angle θ (see Fig. 1) and is given by

$$\dot{\theta} = -k\sin^2\theta \tag{4}$$

(with $v_1 = kX_2$), the spin tensor ω_{ij} has constant nonzero components

$$\omega_{12} = k/2 \text{ and } \omega_{21} = -k/2. \tag{5}$$

As noted by Lee et al. (1983), $k/2$ is the magnitude of the angular velocity of the material lines at $\theta = \pi/4$ and $\theta = 3\pi/4$ which instantaneously coincide with the principal directions of the deformation rate tensor d_{ij}. This is also the average of the angular velocities over all directions in the current configuration.

The oscillations in stresses, according to Lee et al. (1983), arise since the constant spin terms in equation (1) (with $T_{ij} = \alpha_{ij}$, the back stress), generate a tensor α which rotates with an angular velocity of $k/2$ and consequently causes the components of the stress tensor to oscillate with angular frequency k. Thus, while the maximum possible rotation of any material line in the simple shear example (Fig. 1) is π, the "traditional" approach allows unlimited rotation of the back stress tensor (by virtue of equations (1) and (5)). This phenomenon, which is clearly physically unacceptable, is of concern, of course, not only in simple shear but whenever large shearing strains occur in a material exhibiting induced anisotropy.

The solution to this problem, suggested by Dienes (1979), Lee, et al. (1983) and Dafalias (1983), involves modification of the rate equation (1) by replacing ω_{ij} by some other $\hat{\omega}_{ij}$. Lee et al. (1983) demonstrate that using the actual angular velocity of a line of material points (from equation (4)) eliminates the oscillatory stresses in the simple shear example. For more general cases, they suggest a modified Jaumann derivative associated with a spin tensor $\hat{\omega}_{ij}$ based on the spin of lines of material elements carrying the major influence of the back stress α. Dafalias (1983) and Dienes (1979), on the other hand, suggest using the spin associated with the antisymmetric tensor $\dot{\mathbf{R}}\bullet\mathbf{R}^T$ (where R is the rotation matrix from the polar decomposition of the deformation gradient $\mathbf{F} = \mathbf{R}\bullet\mathbf{U}$).

Atluri (1984), while still suggesting modification of the constitutive models involved, takes a somewhat different approach. He starts with a lengthy *hypo-elastic* constitutive model, initially leaving open the exact choice of an objective stress rate. Later, based on ideas of a "complete" *hypo-elastic law*, he modifies the rate of the back stress equation for the case of a *rigid* – kinematic hardening plastic model. The relevance of elasticity considerations, to the modeling of rigid-plastic behavior, is not clear from this paper.

The authors of this present paper feel that while there is considerable merit in the ideas of the researchers cited above, the present situation falls far short of a complete resolution of the above problem, and further investigation is necessary. Proper understanding of finite plastic strains and rotations in the presence of anisotropy will require careful experimentation and possible development of appropriate micromechanical models before a satisfactory phenomenological model is obtained. Clearly, mathematical formalism should not dictate the mechanics concepts, but, rather, a mathematical model should be sought which most clearly expresses the mechanics involved.

It is not the purpose of this paper, then, to offer a complete resolution of the dilemma discussed in the preceding paragraphs. Instead, attention is being drawn here to another aspect of the modeling process which is very important but has not received much attention so far. This aspect relates to the process of proceeding from the intrinsic objective rate of tensors as delivered by the constitutive model to the desired goal of obtaining the components of the Cauchy stress at a moving material point in a spatially fixed coordinate frame. The suggestion here is to reexamine the usual process of converting the objective tensor rates obtained from the constitutive model to a material rate through the use of an appropriate spin $\hat{\omega}_{ij}$, and then integration of these rates in a spatially fixed basis. The essential idea here is to use the elastic rotation instead of the spin in order to attain the above goal. This is motivated by the observation that, in the simple shear problem, integration of the constant angular velocity leads to unbounded rotation, while, as mentioned previously, the actual maximum rotation of a material line is bounded.

If material behavior is *hypo-elastic*, the elastic rotation $\mathbf{R}^{(e)}$ is equal to the total rotation $\mathbf{R}$. In this case, the new approach suggested here is equivalent to use of the Dienes (1979) rate (also called the Green-Naghdi rate by Johnson and Bamman (1984)). Also, under certain restrictive conditions, the new algorithm becomes equivalent to the formulation of Rolph and Bathe (Willam, 1984) using the logarithmic or Hencky strain. Thus, careful examination of the proposed idea, for the case of a simple hypoelastic material, provides a great deal of insight.

The elastic-plastic case is considered next. This is followed by a very brief discussion of a state variable type constitutive model which includes induced anisotropy and presentation of numerical results for the simple shear problem obtained by direct integration, from the "traditional" (Nagtegaal and deJong, 1982) as well as the present point of view. It is further realized that since this issue is not limited, of course, to the problem of simple shear alone, it is important to demonstrate the implementation of the rotation idea in a general computational context. The rotation idea is implemented in a general purpose two-dimensional BEM computer program (Chandra and Mukherjee, 1983, 1985) and the same simple shear example is solved using this program.

A Hypoelastic Example

Hooke's law for isotropic linear elasticity is of the form

$$\sigma = \phi(\epsilon) = \lambda(tr\,\epsilon)\mathbf{I} + 2G\epsilon \tag{6}$$

where σ is the Cauchy stress, ϵ the strain, λ and $\mu = G$ are Lame constants and tr denotes the trace of the corresponding tensor. A simple hypoelastic generalization of the above equation can be written as

$$\overset{\square}{\sigma} = \phi(\mathbf{D}) = \lambda(tr\,\mathbf{D})\mathbf{I} + 2G\mathbf{D} \tag{7}$$

where $\overset{\square}{\sigma}$ is an, as yet, unspecified but objective rate of the stress and $\mathbf{D}$ is the deformation rate whose components have the form

$$d_{ij} = (1/2)(v_{i,j} + v_{j,i}) \tag{8}$$

The proposal here is to proceed from $\phi(\mathbf{D})$ to the Cauchy stress in a desired global spatially fixed basis by using the equation

$$\sigma(t) = R(t)\sigma_0 R^T(t) + R(t)\left[\int_0^t R^T(\tau)\phi(D)R(\tau)d\tau\right]R^T(t) \tag{9}$$

where $F = \partial \mathbf{x}/\partial \mathbf{X} = \mathbf{R}\cdot\mathbf{U}$; σ_o is the value of the Cauchy stress at time zero and $\mathbf{R}^T$ denotes the transpose of $\mathbf{R}$. The symbols in the above equation are matrices corresponding to the appropriate tensors.

The rationale behind this proposal is to observe that the expression $R^T\phi(D)R$ delivers the components of $\phi(\mathbf{D})$ in a local basis which is rotating with respect to the fixed global basis with R the measure of this rotation. Integration is then carried out with respect to an observer in the rotating basis. Finally, premultiplication by R and postmultiplication by R^T, at any time, delivers the Cauchy stress components in the desired global basis.

Relationship With the Dienes Rate (Dienes, 1979). Defining the Cauchy stress $\bar{\sigma}$ in the rotating basis as (Johnson and Bamman, 1984) and using equation (9)

$$\bar{\sigma} = R^T\sigma R = \sigma_0 + \int_0^t R^T\phi(D)R\,d\tau \tag{10}$$

Proceeding in a manner similar to Dienes (1979), one differentiates equation (10) with respect to time. Comparing these expressions and taking note of equation (7) results in the equation

$$\phi(D) = \overset{\square}{\sigma} = \dot{\sigma} - \Omega\sigma + \sigma\Omega \tag{11}$$

so that $\overset{\square}{\sigma}$, for this hypoelastic model, is the Dienes rate of the Cauchy stress, with $\Omega = \dot{R}R^T$.

Thus, equation (9) can be regarded as an integral form of equation (7) and helps to clarify the physical interpretation of the Dienes rate. It has recently been brought to the authors' attention that Goddard and Miller (1966) have presented an equation similar to (9) for the inverse of the Jaumann derivative.

Relationship With Rolph and Bathe (Willam, 1984). Starting from the relationship

$$2\mathbf{D} = \dot{\mathbf{F}}\cdot\mathbf{F}^{-1} + \mathbf{F}^{-T}\cdot\dot{\mathbf{F}}^T \tag{12}$$

and using polar decomposition of $\mathbf{F}$, it is easy to show that

$$2\mathbf{D} = \mathbf{R}\cdot(\dot{\mathbf{U}}\cdot\mathbf{U}^{-1} + \mathbf{U}^{-1}\cdot\dot{\mathbf{U}})\cdot\mathbf{R}^T. \tag{13}$$

This time a special situation is considered in which the *directions of principal stretches remain fixed in the body* during deformation. Thus, one may decompose $U = Q\,\Lambda(t)\,Q^T$, where $\Lambda(t)$ is diagonal and Q is orthogonal but *independent of time.* In this case, one can show that

$$R^TDR = \frac{D}{Dt}\,(lnU) \tag{14}$$

where $lnU = Qln\Lambda Q^T$, so that, from equation (7)

$$R^T\phi(D)R = \frac{D}{Dt}\,[\lambda tr(lnU)I + 2GlnU] \tag{15}$$

From equation (10), the left hand side of the above equation equals $\dot{\bar{\sigma}}$. Hence,

$$\sigma = \lambda(trE)I + 2GE \tag{16}$$

where $E = R\,lnU\,R^T$ is the logarithmic or Hencky Strain. The above equation is the model of Rolph and Bathe (Willam, 1984). Thus, it has been shown that for an isotropic linear hypoelastic material, for deformations in which the directions of principal stretches remain fixed in the body, equations (7), (9) and (16), are all equivalent. Equations (7) and (9), of course, are equivalent under more general conditions as shown before.

Elasto-Plasticity With Finite Rotations

Elasto-plastic problems in materials exhibiting induced anisotropy typically involve tensors such as the back stress α. A typical evolution equation for small-strain small rotation elasto-plasticity might be of the form (Lee, et al., 1983)

$$\dot{\alpha} = g[\dot{\bar{\epsilon}}^{(p)}]\dot{\epsilon}^{(p)} \tag{17}$$

where $\dot{\epsilon}^{(p)}$ is the plastic strain rate and $\dot{\bar{\epsilon}}^{(p)}$ is a suitable invariant of $\dot{\epsilon}^{(p)}$. Large strain generalizations of equation (17) usually involve replacing $\dot{\alpha}$ with a suitable objective rate of α and $\dot{\epsilon}^{(p)}$ with $\mathbf{D}^{(p)}$, the plastic part of the rate of deformation tensor. Thus, the evolution of such tensors must be considered in addition to that of the stress. Also, in general, the function $\phi(\mathbf{D})$ in the hypoelastic law (7) might involve the stress as well.

It is proposed that for such problems, a modified form of equation (9) with elastic rotations $\mathbf{R}^{(e)}$ be adopted. Thus, for a small time step Δt, one may write

$$\bar{T}_{t+\Delta t} = \bar{T}_t + [R^{(e)T}HR^{(e)}]_t\Delta t \tag{18}$$

$$T_{t+\Delta t} = [R^{(e)}\,\bar{T}R^{(e)T}]_{t+\Delta t} \tag{19}$$

where the tensor $\mathbf{T}$ can be the Cauchy stress, the back stress or some other suitable tensor internal variable and $\mathbf{H}$ is a function such that for *small strain small rotation* problems $\dot{\mathbf{T}} = \mathbf{H}$.

The function $\mathbf{H}$ must be suitably interpreted by replacing $\dot{\epsilon}^{(e)}$ by $D^{(e)}$ etc. Also, Δt is a small time increment. The above equations are, strictly speaking, correct in the limit $\Delta t \to 0$. Operationally, of course, one must use small, finite time increments Δt. Also, these equations must be used in a march forward time integration procedure.

The need for the use of elastic rotations arises from the nature of elastic and plastic deformations. As with elastic strains, elastic rotations with respect to a virgin configuration are remembered by the solid. It is common to assume that for plasticity analysis it is not necessary to use variables involving the virgin configuration of the material prior to plastic flow (Lee, et al., 1983).

Recently, researchers such as Dafalias (Willam, 1984) and Aifantis (1984) have been looking into constitutive descriptions for plastic and elastic rotations. This research is still at an early stage. Thus, *for the purpose of the rest of this work*, it is assumed that $\mathbf{R} = \mathbf{R}^e$.

Inelastic Constitutive Model

The first constitutive assumption made here is that the deformation rate tensor can be linearly decomposed into an elastic and a nonelastic part

$$d_{ij} = d_{ij}^{(e)} + d_{ij}^{(n)} \tag{20}$$

A hypoelastic law, similar to equation (7), is here assumed to relate the stress rate to the elastic part of the rate of deformation tensor

$$\overset{\square}{\sigma}_{ij} = \lambda d_{kk}^{(e)}\delta_{ij} + 2Gd_{ij}^{(e)} \tag{21}$$

where the new symbol δ_{ij} is the Kronecker delta.

The nonelastic strain rates $d_{ij}^{(n)}$ must be specified in terms of the Cauchy stress, and possibly other variables, through appropriate inelastic constitutive equations. Combined creep-plasticity constitutive models with state variables are of interest in this work. In the interest of brevity, the reader is referred to Mukherjee (1982) for a discussion of such constitutive models.

The particular constitutive model that has been used to generate the numerical results presented in the next sections is, due to Hart (1976), generalized to the case of large strains and displacements (Chandra and Mukherjee, 1983, 1985). The model has a scalar state variable Σ (the hardness, called σ^* in previous publications) and a tensor state variable, the anelastic strain $\epsilon_{ij}^{(a)}$, which is responsible for induced material anisotropy. For this work, it is assumed that

$$\overset{\square}{\epsilon}{}_{ij}^{(a)} = d_{ij}^{(n)} - d_{ij}^{(p)} \tag{22}$$

where $d_{ij}^{(p)}$ is the irrecoverable portion of the deformation rate. Given the Cauchy stress and the state variables at any time, this constitutive model delivers, among other things, $d_{ij}^{(n)}$ and the state variable rates $\dot{\Sigma}$ and $\overset{\square}{\epsilon}{}_{ij}^{(a)}$ at that time.

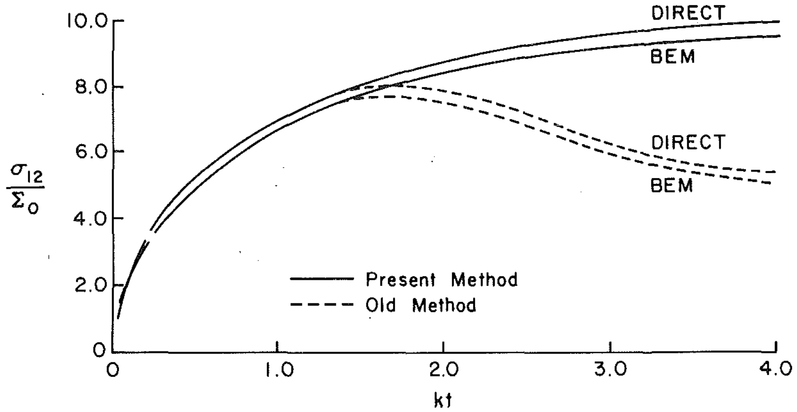


Fig. 2 Shearing stress as a function of shearing strain from two methods. Direct integration and BEM solutions.

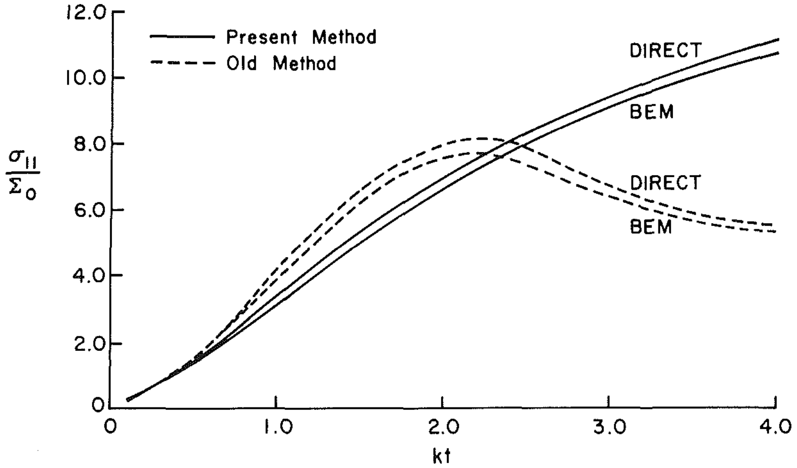


Fig. 3 Longitudinal stress as a function of shearing strain from two methods. Direct integration and BEM solutions.

Numerical Results by Direct Integration

Numerical results for the simple shear problem (Fig. 1), obtained by direct integration for two different approaches, are presented here. The first of these is the "traditional" approach (Nagtegaal and deJong, 1982) and the second is the new one.

The simple shear problem, with (Dienes, 1979, Lee, et al., 1983, Dafalias, 1983)

$$x_1 = X_1 + ktX_2, x_2 = X_2, x_3 = X_3 \tag{23}$$

yields the following nonzero components of deformation rates and spin

$$d_{12} = d_{21} = k/2, \omega_{12} = -\omega_{21} = k/2. \tag{24}$$

Using the "traditional" approach, the nonzero material rates of the components of Cauchy stress, using equations (1, 21 with σ^*_{ij}, $d^{(n)}_{kk} = 0$ and 24), becomes

$$\dot{\sigma}_{11} = k\sigma_{12} - 2Gd^{(n)}_{11} \qquad \dot{\sigma}_{22} = -k\sigma_{12} - 2Gd^{(n)}_{22} \tag{25}$$
$$\dot{\sigma}_{12} = \frac{k}{2}(\sigma_{22} - \sigma_{11}) + Gk - 2Gd^{(n)}_{12}.$$

The objective rates of the Cauchy stress and the tensor state variable $\epsilon^{(a)}_{ij}$ are identified with the Jaumann rate in this "traditional" formulation.

The new formulation requires the elastic rotation matrix $R^{(e)}$. For the time being, this is taken to be the rotation matrix R from the polar decomposition theorem. Following Dienes (1979), the nonzero components of R become $R_{11} = R_{22} = \cos\beta$, $R_{12} = -R_{21} = \sin\beta$, $R_{33} = 1$, where $\tan\beta = kt/2$.

Starting from equation (21)

$$\overset{\square}{\sigma}_{11} = -2Gd^{(n)}_{11}, \overset{\square}{\sigma}_{22} = -2Gd^{(n)}_{22}$$
$$\text{and} \qquad \qquad \tag{26}$$
$$\overset{\square}{\sigma}_{12} = Gk - 2Gd^{(n)}_{12}$$

and similar equations for $\overset{\square}{\epsilon}_{ij}$ from Hart's constitutive model, it is a simple matter to compute $\sigma_{ij}(t)$ and $\epsilon^{(a)}_{ij}(t)$ in a spatially fixed frame through the use of R_{ij} and equations (18 and 19) with $h_{11} = -2Gd^{(n)}_{11}$ for $\boldsymbol{\sigma} = \mathbf{T}$ etc.

Numerical results for the two cases, obtained by direct integration by marching forward in time, are shown in Figs. 2 and 3. The material parameters used are representative of 304 stainless steel at 400°C (Mukherjee, 1982). It is clearly evident that use of constant spin components from equation (24) leads to a prediction of oscillation of stress components in response to a montonically increasing shearing strain kt, while the use of rotation rather than spin leads to a physically acceptable situation in this example, namely monotonic increase in stress components as functions of monotonically increasing shearing strain.

Implementation in BEM

It is clear that the importance of the issue discussed in the previous sections is not limited to the simple shear example, but would be of consequence in problems whenever the local shearing strain at a point or points in a structure becomes large during deformation of the structure. While most of the researchers referred to earlier have not looked into this more general situation, the subject of this concluding section of this paper is an investigation of the implementation of the rotation idea in a general context. The rotation idea has been implemented in a general two-dimensional BEM computer program (Chandra and Mukherjee, 1983, 1985) capable of solving elastic-viscoplastic problems in the presence of large nonelastic strains and rotations. The original BEM formulation has been published before (Chandra and Mukherjee, 1983, 1985) and only some key points and the final BEM equations are presented here.

The original BEM formulation, published previously by Chandra and Mukherjee (1983, 1985), is based on use of the Jaumann rate in the hypoelasticity equation (7), together with the updated Lagrangian approach. A key equation in this formulation is the relationship between the material rate of the Lagrangian (or nominal) stress **S** and the Jaumann rate of the Cauchy stress. In the updated Lagrangian framework, this relationship is (Mukherjee and Chandra, 1985)

$$\dot{\mathbf{S}} = \overset{*}{\sigma} - \sigma \cdot \omega - \mathbf{D} \cdot \sigma \tag{27}$$

where the deformation is assumed to be incompressible and, in general,

$$\mathbf{S} = (\det \mathbf{F})\mathbf{F}^{-1} \cdot \sigma. \tag{28}$$

Further, it can be shown in general that (Atluri, 1984)

$$\boldsymbol{\Omega} = \omega - \frac{1}{2}\,\mathbf{R} \cdot [\dot{\mathbf{U}} \cdot \mathbf{U}^{-1} - \mathbf{U}^{-1} \cdot \dot{\mathbf{U}}] \cdot \mathbf{R}^T \tag{29}$$

so that in the updated Lagrangian approach, with $\mathbf{R} = \mathbf{U} = \mathbf{I}$, one obtains $\boldsymbol{\Omega} = \omega$.

Thus, the BEM equation for the velocity field, from Chandra and Mukherjee (1983 and 1985), still remains valid for the present formulation with the Dienes rate. It must be remembered, however, that once the velocity field is obtained throughout the body, further calculations are necessary before, for example, the Cauchy stress history is obtained in a reference frame. These subsequent manipulations are carried out here using equations (18, 19) rather than by integrating the material rates of the relevant tensors.

It has been shown before (Chandra and Mukherjee, 1983, 1985) that the velocity at a point p inside a structure undergoing small elastic strains, but large viscoplastic strains and rotations can be written as

$$\begin{aligned} v_j(p) = {} & \int_{\partial B^o} [U_{ij}(p,Q)\dot{\tau}_i(Q) - T_{ij}(p,Q)v_i(Q)]dS_Q^o \\ & + \int_{B^o} \rho_o U_{ij}(p,q)\dot{F}_i^o(q)dV_q^o \\ & + \int_{B^o} 2GU_{ij,k}(p,q)d_{ik}^{(n)}(q)dV_q^o \\ & + \int_{B^o} U_{ij,m}(p,q)G_{mikl}(q)v_{k,l}(q)dV_q^o. \end{aligned} \tag{30}$$

In the above, U_{ij} and T_{ij} are the usual two point kernels of small deformation elasticity that are given in many references (e.g., Mukherjee, 1982); G_{ijkl} is a function of components of the Cauchy stress (Mukherjee and Chandra, 1985); τ_i is the traction at a point on the boundary; and ρ_o and $\dot{F}_i^o$ are the mass density and body force rate, respectively, in the reference configuration. Also, p (or P) is a source point and q (or Q) is a field point, with lower case letters denoting points inside the reference volume B_o and P and Q denoting points on its boundary ∂B_o. Finally, a comma denotes differentiation at a field point q.

The traction rate $\dot{\tau}_i$ in the above equation (using equation (27)), is

$$\dot{\tau}_i = n_j \dot{s}_{ji} = \dot{t}_i - n_j \sigma_{jk} \omega_{ki} - n_j d_{jk} \sigma_{ki} \tag{31}$$

where

$$\dot{t}_i = n_j \overset{\square}{\sigma}_{ji}. \tag{32}$$

In the above equation, $\dot{t}_i$ can be interpreted as the component of the rate of the prescribed follower force, per unit deformed surface area, on the deforming boundary. The follower force moves with a boundary point and rotates with the normal to the boundary at that point.

The next steps require the determination of a boundary integral equation by taking the limit of equation (30) as $p \rightarrow P$ and the stress rate equation by differentiating equation (30) at a source point. These equations are given in previous papers (e.g., Chandra and Mukherjee, 1983). The plane strain and plane stress versions of these equations are also available in the above-mentioned paper.

Solution Strategy. The solution strategy follows the same approach as discussed in previous publications (e.g., Chandra and Mukherjee, 1983). The boundary integral equation is first discretized to obtain appropriate algebraic equations of the form

$$[A]\{v\} + [B]\{\dot{\tau}\} = \{b\}. \tag{33}$$

A march forward time integration scheme is then used with suitable updating of the configuration of the body. The presence of velocity gradients in the boundary traction rates and in the last domain integral in equation (30) requires iterations within each step. The essential differences between this new approach and the "standard" one (Chandra and Mukherjee, 1983) are stated below.

a) Once the elastic problem is solved at zero time, the displacement gradients are used to form the deformation gradient matrix F.

b) This matrix F is decomposed according to the polar decomposition theorem $F = RU$ and R is obtained.

c) The iterations are completed and $v_j(p)$, $v_{j,l}(p)$, and $\overset{\square}{\sigma}_{ji}$ and $\overset{\square}{\epsilon}{}_{ij}^{(a)}$ are calculated in the fixed global basis at zero time. The material derivative of F is determined from the equation $\dot{F} = LF$ where $L_{ij} = v_{i,j}$.

d) Time integration is performed next. An explicit Euler type scheme with proper time step controls (Kumar, et al., 1980) is used to find the relevant quantities including F at time Δt. F is decomposed into RU at time Δt and R is obtained at Δt.

e) The objective rates of the Cauchy stress and the anelastic strain are integrated in time through equations (18) and (19). Here

$$\begin{aligned} [\bar{\sigma}]_{t=\Delta t} &= [\bar{\sigma}]_{t=0} + ([R]_{t=\Delta t}^T [\overset{\square}{\sigma}]_{t=0} [R]_{t=\Delta t})\Delta t \\ [\sigma]_{t=\Delta t} &= ([R][\bar{\sigma}][R]^T)_{t=\Delta t} \end{aligned} \tag{34}$$

and similarly for $[\epsilon^{(a)}]$.

Thus, the relevant quantities (displacements, displacement gradients, stresses, anelastic strains, etc.) are found at $t = \Delta t$. The time histories of various quantities are then obtained by marching forward in time and suitable updating of the geometry and the kernels.

Numerical Results. The BEM formulation, described earlier, has been implemented in a computer program. This program can, of course, be used to solve general two-dimensional elastic-viscoplastic problems in the presence of large strains and rotations. It has been validated by solving the simple shear problems discussed earlier. The BEM numerical results in Figs. 2 and 3 agree, within about 5 percent, with the direct numerical results for the same problem. This exercise demonstrates the feasibility of implementation of the rotation idea in a general purpose two-dimensional computer program.

Conclusions

This paper suggests tht elastic rotations should be used in the modeling of large shearing of solids exhibiting induced anisotropy during inelastic deformation. It is demonstrated here that direct integration of objective rates of tensors, obtained from a constitutive model, and subsequent correction for the rotation of these tensors, overcomes problems of unbounded rotations that could emanate from the usual procedure of using spins to relate Jaumann rates to material rates. Use of rotations eliminates spurious oscillations in a simple shear example. It is further demonstrated that the rotation idea can be efficiently incorporated into a general purpose analysis formulation and computer program. This is done by adapting a two-dimensional BEM formulation and computer program to reflect the use of rotations rather than spins. Numerical results from the general BEM program, for the simple shear example, agree well with a solution for the same problem obtained by direct numerical integration. The rotation idea can also be incorporated into the finite element or other general numerical methods.

Acknowledgments

The research of S. Mukherjee was supported by contract No. CEE-8206344 of the National Science Foundation with Cornell University.

References

Aifantis, E. C., 1984, "On the Microstructural Origin of Certain Inelastic Models," ASME *Journal of Engineering Materials and Technology*, Vol. 106, pp. 326-330.

Atluri, S. N., 1984, "On Constitutive Relations at Finite Strain: Hypo-Elasticity and Elasto-Plasticity with Isotropic or Kinematic Hardening," *Computer Methods in Applied Mechanics and Engineering*, Vol. 43, pp. 137-171.

Chandra, A., and Mukherjee, S., 1983, "Applications of the Boundary Element Method to Large Strain-Large Deformation Problems of Viscoplasticity," *Journal of Strain Analysis*, Vol. 8, pp. 261-270.

Chandra, A., and Mukherjee, S., 1985, "A Boundary Element Formulation for Sheet Metal Forming," *Applied Mathematical Modeling*, Vol. 9, pp. 175-182.

Dafalias, Y. F., 1983, "Corotational Rates for Kinematic Hardening at Large Plastic Deformations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 561-565.

Dienes, J. K., 1979, "On the Analysis of Rotation and Stress Rate in Deforming Bodies," *Acta Mechanica*, Vol. 32, pp. 217-232.

Goddard, J. D., and Miller, C., 1966, "An Inverse of the Jaumann Derivative and Some Applications in the Rheology of Viscoelastic Fluids," *Rheologica Acta*, Vol. 5, pp. 177-184.

Hart, E. W., 1976, "Constitutive Relations for the Non-Elastic Deformation of Metals," ASME *Journal of Engineering Materials and Technology*, Vol. 98, pp. 193-202.

Johnson, G. C., and Bamman, D. J., 1984, "A Discussion of Stress Rates in Finite Deformation Problems, *International Journal of Solids and Structures*, Vol. 20, pp. 725-737.

Kumar, V., Morjaria, M., and Mukherjee, S., 1980, "Numerical Integration of Some Stiff Constitutive Models of Inealstic Deformation," ASME *Journal of Engineering Materials and Technology*, Vol. 102, pp. 92-96.

Lee, E. H., Mallett, R. L., and Wertheimer, T. B., 1983, "Stress Analysis for Anisotropic Hardening in Finite Deformation Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 554-560.

Mukherjee, S., 1982, *Boundary Element Methods in Creep and Fracture*, Applied Science Publishers, London, U.K.

Mukherjee, S., and Chandra, A., 1985, "Nonlinear Solid Mechanics," *Boundary Element Methods in Mechanics*, Beskos, D. E., ed., North Holland, (expected).

Nagtegaal, J. C., and de Jong, J. E., 1982, "Some Aspects of Non-Isotropic Workhardening in Finite Strain Plasticity," *Proceedings of the Workshop on Plasticity of Metals at Finite Strain*, Lee, E. H., and Mallet, R. L., eds., Division of Applied Mechanics, Stanford University, Stanford, CA, and Department of Mechanical Engineering, Aeronautical Engineering and Mechanics, R.P.I., Troy, NY, pp. 65-102.

Willam, K., ed., 1984, *Constitutive Equations: Macro and Computational Aspects*, ASME Publication, New York.

Ziegler, H., 1959, "A Modification of Prager's Hardening Rule," *Quarterly of Applied Mathematics*, Vol. 17, pp. 55-65.

Radial Dependence of Near-Tip Continuum Fields for Plane Strain Tensile Crack Growth in Elastic-Ideally Plastic Solids

W. J. Drugan
Department of Engineering Mechanics,
University of Wisconsin,
Madison, WI 53706
Assoc. Mem. ASME

This paper is an extension of work by Drugan et al. (1982) who derive the stress and deformation fields at the tip of a plane strain tensile crack that grows quasi-statically, under general nonsteady conditions, in an elastic-ideally plastic solid. Here I perform a higher-order analysis of the near-tip fields for this growing crack problem. My principal objectives are to determine the radial variation of the near-tip stress field and elucidate the structure of the deformation fields in the 90-deg sector ahead of the growing crack; this information was not provided by the lowest-order solution of Drugan et al. (1982). I also derive a crucial asymptotic expression for the normal radial component of the deformation rate tensor in a moving "centered fan" plastic sector, which was given without complete proof by Rice (1982). The analysis presented herein differs from typical perturbation analyses in that I am able to derive *the higher-order structure of the continuum fields rather than having to* assume *expansions for them. Among the results, normal polar components of deviatoric stress are shown to vary as* $(\ln r)^{-1}$*, while the in-plane polar shear component varies as* $(\ln r)^{-2}$*, for small* $r > 0$ *in moving "centered fan" plastic sectors,* r *denoting distance from the (moving) crack tip. Further, in-plane strains proportional to* $\ln|\ln r|$ *as* $r \to 0$ *appear not to be precluded in the 90-deg sector ahead of the growing crack.*

1 Introduction

A knowledge of the stress and deformation fields associated with the presence of a growing crack in a body is of great importance for the formulation of fracture criteria and for the investigation of interactions between cracks and other geometrical, or material, inhomogeneities in a solid. Drugan et al. (1982) give an exact asymptotic solution (exact in the limit of $r \to 0$, where r is distance from the crack tip) of the stress and deformation fields at the tip of a quasi-statically growing plane strain tensile crack under well-contained yielding conditions. Their analysis treats an isotropic elastic-ideally plastic solid obeying the Huber-Mises yield condition, for the general case of Poisson ratio $\nu < 0.5$, i.e., elastic compressibility. The Drugan et al. (1982) solution is a *lowest-order* solution to the equations governing stress and deformation as $r \to 0$: the stress components are determined as the functions of angle about the crack tip to which they reduce at $r=0$, and explicit expressions for the deformation fields at small r result for all angles about the crack tip *except* for a 90-deg sector centered directly ahead of the growing crack, where the solution simply bounds the order of the deformation fields as $r \to 0$. Other investigators have dealt with more specialized versions of the elastic-ideally plastic crack growth problem (e.g., fully incompressible material, steady-state crack growth), and their work is referenced and discussed in the Drugan et al. (1982) study. In particular, Gao (1983) has attacked the problem for compressible material via a steady-state formulation, but we have serious disagreement with his solution, as detailed in Drugan et al. (1982).

The present work supplements the Drugan et al. (1982) solution by extending it to higher order in radial dependence, in near-tip sectors of principal plastic loading. This is accomplished by employing the more explicit asymptotic forms of the governing continuum equations derived recently by Drugan (1985). These reduced forms are specialized to the present problem in Section 2. As in the Drugan et al. (1982) analysis, the effects of geometry changes on the formulation of stress and stress-rate measures and on the equilibrium equations are neglected. Finite element solutions by McMeeking (1977) for a *stationary* crack experiencing plane strain, Mode I, small scale yielding conditions show finite strain effects to be important only within a radius of approximately 2 to 3 times the crack tip opening displacement; this radius is expected to be considerably smaller for a *growing* crack due to the far weaker crack tip strain singularity in this case. These

---

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October 30, 1984; final revision, July 2, 1985.

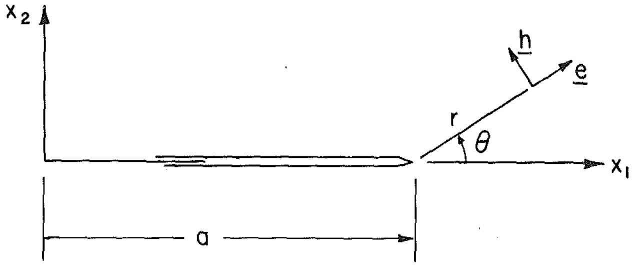


Fig. 1 Cartesian coordinates x_1, x_2, x_3 are fixed in the body; polar coordinates *r*, *θ* are centered at the tip and move with it through the material as the crack grows

and related issues are clearly explained in the excellent review article by Hutchinson (1982).

The lowest-order solution of Drugan et al. (1982) is discussed in Section 3. This discussion includes a new derivation of the lowest-order deformation-rate fields in moving "centered fan" plastic sectors, providing verification of the results of Rice (1982) (employed in the Drugan et al. (1982) solution) whose derivation is not sufficient, as explained. Section 4 then presents a higher-order analysis of the stress field, showing that the lowest-order solution, when coupled with the general governing equations, *specify* the radial behavior of the stress field. Knowledge of the higher-order stress field behavior in the 90-deg sector centered directly ahead of the growing crack facilitates determination of the heretofore elusive structure of the deformation field in this sector, as shown in Section 5.

The geometry of the problem, illustrated in Fig. 1, is identical to that considered by Drugan et al. (1982). A Cartesian coordinate system, x_1, x_2, x_3, is fixed in the body, with x_1 pointing in the direction of crack growth, a being the measure of crack length, and x_3 lying parallel to the crack front. A polar coordinate system, r, θ, lies in the x_1, x_2 plane, is centered at the crack tip and moves with it through the material as the crack grows; θ is measured from the line ahead of the crack. The unit vectors **e** and **h** correspond to the radial and angular directions, respectively, of this translating polar coordinate system. Therefore,

$$\partial r/\partial x_i = e_i, \quad \partial\theta/\partial x_i = h_i/r \tag{1.1}$$

where

$$e_1 = h_2 = \cos\theta, \quad e_2 = -h_1 = \sin\theta, \quad e_3 = h_3 = 0. \tag{1.2}$$

The convention to be followed throughout the paper is that Greek indices α,β have range 1,2 only, while Latin indices i,j,k,l have range 1,2,3; both types of index will follow the summation convention and will indicate Cartesian (but never polar) components of tensors.

2 Plane Strain Governing Equations and Asymptotic Forms

The analysis to follow is based on the continuum equations that govern plane strain deformation of a Prandtl-Reuss material, assuming small displacement-gradients, i.e., neglecting the effects of deformation in the formulation of stress and stress-rate measures, and equilibrium equations. These continuum equations are summarized here, and the forms to which they reduce for small r are given by employing the results of Drugan (1985). For plane strain deformation, $u_3 = 0$, $\epsilon_{3i} = 0$ and $\sigma_{3\alpha} = 0$, where u_i, ϵ_{ij} and σ_{ij} are components of the displacement vector, the infinitesimal strain tensor and the stress tensor, respectively.

2.1 Equilibrium. For three-dimensional equilibrium, the stress tensor must be symmetric, $\sigma_{ij} = \sigma_{ji}$, and must satisfy the equations

$$\partial\sigma_{ij}/\partial x_j + f_i = 0 \tag{2.1}$$

where f_i are components of the body force vector. These equations can be rephrased in terms of the crack tip polar coordinate system for the plane problem to be studied by regarding $\sigma_{ij} = \sigma_{ij}\ (r, \theta, t)$, with t denoting time, and by using (1.1):

$$(\partial\sigma_{\alpha\beta}/\partial\theta)(h_\beta/r) + (\partial\sigma_{\alpha\beta}/\partial r)e_\beta + f_\alpha = 0. \tag{2.2}$$

Equations (2.2) govern Cartesian stress components; for polar components (2.2) become

$$\partial\sigma_{rr}/\partial r + (1/r)(\partial\sigma_{r\theta}/\partial\theta) + (\sigma_{rr} - \sigma_{\theta\theta})/r + f_r = 0 \tag{2.3a}$$

$$\partial\sigma_{r\theta}/\partial r + (1/r)(\partial\sigma_{\theta\theta}/\partial\theta) + 2\sigma_{r\theta}/r + f_\theta = 0. \tag{2.3b}$$

Drugan (1985) proved that if the following conditions are all met: (*i*) *deviatoric* components of stress are bounded (as they are in the present elastic-ideally plastic constitutive model); (*ii*) the body is in equilibrium with bounded body forces f_i; (*iii*) the quantities $r[\ln(R/r)]\partial\sigma_{ij}/\partial r$ *exist* in the limit as $r \to 0$ (the equilibrium equations (2.3) require this for $r\partial\sigma_{rr}/\partial r$ and $r\partial\sigma_{r\theta}/\partial r$ if $\partial\sigma_{r\theta}/\partial\theta$, $\partial\sigma_{\theta\theta}/\partial\theta$, σ_{rr}, $\sigma_{\theta\theta}$, $\sigma_{r\theta}$ are all presumed to exist as $r \to 0$); then

$$\partial\sigma_{ij}/\partial r = o\{[r\ln(R/r)]^{-1}\} \quad \text{as} \quad r \to 0 \tag{2.4}$$

where R is an undetermined constant having length dimensions. Using this result together with the bounded body force assumption, the equilibrium equations adopt simplified asymptotic forms; for example, (2.2) become

$$(\partial\sigma_{\alpha\beta}/\partial\theta)h_\beta + o\{[\ln(R/r)]^{-1}\} = 0 \quad \text{as} \quad r \to 0. \tag{2.5}$$

Here and throughout the text, I employ standard order symbols; a clear discussion of these and gauge functions is given by Van Dyke (1975).

2.2 Yield Condition. The Huber-Mises yield condition

$$f(\sigma_{ij}) = s_{ij}s_{ij}/2 - k^2 = 0 \tag{2.6}$$

is assumed, where $s_{ij} \equiv \sigma_{ij} - \delta_{ij}\sigma_{kk}/3$ are components of the deviatoric stress tensor, δ_{ij} is the Kronecker delta, and k is shear strength. In the analysis to follow, it is often more convenient to employ two differential forms of (2.6),

$$s_{ij}(\partial\sigma_{ij}/\partial\theta) = 0 \tag{2.7a}$$

$$s_{ij}(\partial\sigma_{ij}/\partial r) = 0 \tag{2.7b}$$

which must both hold in all plastically deforming regions of the body.

2.3 Elastic-Plastic Stress-Strain Relations. The rate of deformation D_{ij} is defined as

$$D_{ij} \equiv \dot{\epsilon}_{ij} = \frac{1}{2}(\partial v_i/\partial x_j + \partial v_j/\partial x_i) \tag{2.8}$$

where v_i are components of the material velocity vector and a superposed dot denotes time rate at a material point. The incremental Prandtl-Reuss theory is adopted in rate form, so that for material currently experiencing elastic-plastic deformation,

$$D_{ij} = D^e_{ij} + D^p_{ij} = \frac{1+\nu}{E}\dot{\sigma}_{ij} - \frac{\nu}{E}\delta_{ij}\dot{\sigma}_{kk} + \Lambda s_{ij} \tag{2.9}$$

where E is Young's modulus, ν is Poisson's ratio, $\Lambda \geq 0$ is an undetermined parameter, and superscripts e and p denote elastic and plastic components, respectively. For material experiencing solely elastic current deformation, whether or not it has been previously strained plastically, (2.9) applies with $\Lambda \equiv 0$.

A convenient expression for the stress rate at a material point, $\dot{\sigma}_{ij}$, results from applying the chain rule to the assumed functional form $\sigma_{ij} = \sigma_{ij}\ (r, \theta, t)$ (Rice, 1982):

$$\begin{aligned}\dot{\sigma}_{ij} &= (\partial\sigma_{ij}/\partial\theta)\dot{\theta} + (\partial\sigma_{ij}/\partial r)\dot{r} + \partial\sigma_{ij}/\partial t \\ &= (\partial\sigma_{ij}/\partial\theta)\dot{a}\sin\theta/r - (\partial\sigma_{ij}/\partial r)\dot{a}\cos\theta + \partial\sigma_{ij}/\partial t,\end{aligned} \tag{2.10}$$

having employed $\dot{\theta} = \dot{a}\sin\theta/r$ and $\dot{r} = -\dot{a}\cos\theta$, which result from the translation of the crack tip polar coordinate system with the growing crack. Two asymptotic forms of (2.10) were derived by Drugan (1985):

$$\dot{\sigma}_{ij} = (\partial\sigma_{ij}/\partial\theta)\dot{a}\sin\theta/r - (\partial\sigma_{ij}/\partial r)\dot{a}\cos\theta + 0(1) \quad \text{as} \quad r \to 0 \qquad (2.11a)$$

$$\dot{\sigma}_{ij} = (\partial\sigma_{ij}/\partial\theta)\dot{a}\sin\theta/r + o\{[r\ln(R/r)]^{-1}\} \quad \text{as} \quad r \to 0; \qquad (2.11b)$$

these facilitate employment of (2.9) in the asymptotic analysis.

2.4 Plane Strain Condition. Since I assume $\epsilon_{ij} = \epsilon^e_{ij} + \epsilon^p_{ij}$, the plane strain condition $\epsilon_{33} = 0$ requires

$$\epsilon^p_{33} = -\epsilon^e_{33} = -[\sigma_{33} - \nu(\sigma_{11} + \sigma_{22})]/E. \qquad (2.12)$$

This shows ϵ^p_{33} to be bounded at the crack tip since (Drugan, 1985) equilibrium requires *all* stress components to be bounded at the crack tip if deviatoric stress components are. In the form $D_{33} = 0$, the plane strain requirement permits Λ to be determined from (2.9) whenever $s_{33} \neq 0$, viz.

$$\Lambda = -[\dot{\sigma}_{33} - \nu(\dot{\sigma}_{11} + \dot{\sigma}_{22})]/(Es_{33}). \qquad (2.13)$$

2.5 Compatibility. I assume the displacement field u_i to be continuous near the growing crack tip (except, of course, across the crack flank) since a moving surface of displacement discontinuity would correspond to an unbounded plastic work rate in any subregion of the body traversed by such a surface. Drugan and Rice (1984) showed that a continuous displacement field implies a continuous velocity field near a growing crack tip for the constitutive model employed here, *except* across a quasi-statically propagating surface whose stress state meets highly restrictive conditions which they derived. In all regions where the velocity field is thrice differentiable, plane strain compatibility of the deformation rate field will be satisfied if

$$\partial^2 D_{22}/\partial x_1^2 + \partial^2 D_{11}/\partial x_2^2 - 2\partial^2 D_{12}/\partial x_1 \partial x_2 = 0, \qquad (2.14)$$

which may be restated in terms of the crack tip polar coordinate system as

$$\left(-\frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)D_{rr} - \left(\frac{2}{r}\frac{\partial^2}{\partial r\partial\theta} + \frac{2}{r^2}\frac{\partial}{\partial\theta}\right)D_{r\theta} + \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)D_{\theta\theta} = 0. \qquad (2.15)$$

3 Lowest-Order Solution for Crack Tip Fields

3.1 Summary. Drugan et al. (1982) derive a lowest-order (in r) solution for the stress and deformation fields near the tip of a Mode I (tensile) crack that grows quasi-statically under plane strain, well-contained yielding conditions in an isotropic elastic-ideally plastic solid. The stress field is thus an exact solution to the governing equations (Section 2) at $r = 0$ and is only a function of θ. The Drugan et al. (1982) analysis shows the crack tip field to divide into five angular sectors of four different types on each side of the crack symmetry plane, for the general $\nu < 0.5$ case. This configuration is illustrated in Fig. 2, in which: Sectors A and B can be described, for $r \to 0$, in slip-line terminology as "constant stress" and "centered fan" plastic sectors, respectively; Sectors C and E are plastically deforming with non-constant stress fields and have no asymptotic slip-line analogy; and Sector D contains elastically unloading material. Drugan et al. (1982) show that $\theta = \pi/4$ is the boundary (at $r = 0$) between Sectors A and B for general ν, while θ_1, θ_2, θ_3 are found to depend on the specific ν value. For the case $\nu = 0.3$, they find

$$\theta_1 = 110.26°, \quad \theta_2 = 123.13°, \quad \theta_3 = 160.38°. \qquad (3.1)$$

The lowest-order solution of Drugan et al. (1982) is not sufficient to specify the deformation field in Sector A of Fig. 2 except to show that $D_{ij} = o(1/r)$ as $r \to 0$ there and thus that Sector A produces only bounded contributions to ϵ^p_{ij} as $r \to 0$,

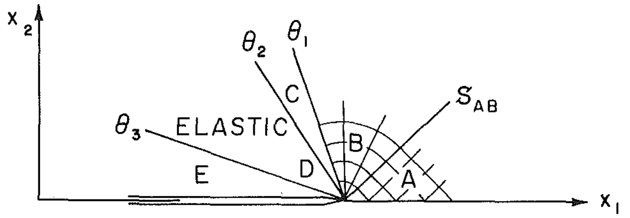


Fig. 2 Exact solution configuration, directly at the crack tip, for a growing crack in material with $\nu < 0.5$ (from Drugan et al., 1982). The boundary S_{AB} is known to be $\theta = \pi/4$ at $r = 0$.

except possibly on $\theta = 0$, where $\epsilon^p_{ii} = o(\ln r)$ as $r \to 0$. Their solution does provide explicit expressions for the deformation field as $r \to 0$ in each of the other near-tip sectors.

3.2 New Derivation of Asymptotic Deformation-Rate Field. Rice (1982) derived from (2.9) and (2.7*a*) two restrictions on the deformation-rate field (specialized here to plane strain response of a Prandtl-Reuss material) which can be fruitfully employed in moving "centered fan" plastic sectors:

$$D_{rr} + D_{\theta\theta} = [(1-2\nu)/E]\dot{\sigma}_{kk} \qquad (3.2)$$

$$D_{\alpha\beta}(\partial\sigma_{\alpha\beta}/\partial\theta) = \frac{1+\nu}{E}\,\dot{\sigma}_{ij}(\partial\sigma_{ij}/\partial\theta) - \frac{\nu}{E}\,\dot{\sigma}_{kk}(\partial\sigma_{ii}/\partial\theta). \qquad (3.3)$$

These restrictions have the following asymptotic forms:

$$D_{rr} + D_{\theta\theta} = [(1-2\nu)/E](\partial\sigma_{kk}/\partial\theta)\dot{a}\sin\theta/r + o[1/(r\ln r)] \quad \text{as} \quad r \to 0 \qquad (3.4)$$

$$rD_{rr}(\partial\sigma_{rr}/\partial\theta - 2\sigma_{r\theta}) = (1/E)[(1+\nu)(\partial\sigma_{ij}/\partial\theta)(\partial\sigma_{ij}/\partial\theta) - \nu(\partial\sigma_{kk}/\partial\theta)^2]\dot{a}\sin\theta + o(1) \text{ as } r \to 0. \qquad (3.5)$$

Equation (3.4) is obtained simply by employing (2.11*b*) in (3.2) as $r \to 0$.

The demonstration that (3.3) reduces to (3.5) as $r \to 0$ in a moving "centered fan" plastic sector is more involved. Although Rice (1982) arrives at a form equivalent to (3.5), his derivation is not sufficient (Pan, 1982), since he considers only the lowest-order *Cartesian* terms in the sum $D_{\alpha\beta}$ $(\partial\sigma_{\alpha\beta}/\partial\theta)$ as $r \to 0$. These terms cancel, however, necessitating retention of higher-order terms. Thus I give below a new derivation of (3.5), since in addition to being fundamental to the earlier solutions cited, it is crucial to the higher-order analysis performed here.

Use of (2.11*b*) permits rewriting (3.3) as

$$rD_{\alpha\beta}(\partial\sigma_{\alpha\beta}/\partial\theta) = (1/E)[(1+\nu)(\partial\sigma_{ij}/\partial\theta)(\partial\sigma_{ij}/\partial\theta) - \nu(\partial\sigma_{kk}/\partial\theta)^2]\dot{a}\sin\theta + o(1) \quad \text{as } r \to 0, \qquad (3.6)$$

having multiplied through by r. The left side of (3.6) is expressible as

$$rD_{\alpha\beta}(\partial\sigma_{\alpha\beta}/\partial\theta) = rD_{rr}(\partial\sigma_{rr}/\partial\theta - 2\sigma_{r\theta}) + 2rD_{r\theta}[-r(\partial\sigma_{rr}/\partial r + f_r)] + rD_{\theta\theta}[-r(\partial\sigma_{r\theta}/\partial r + f_\theta)], \qquad (3.7)$$

where (2.3) have been employed. What remains is to show that $rD_{rr}(\partial\sigma_{rr}/\partial\theta - 2\sigma_{r\theta})$ dominates the other two terms on the right side of (3.7) as $r \to 0$. That it dominates the third term is easily observed by substituting for $D_{\theta\theta}$ from (3.4) and using (2.4), having noted from Drugan et al. (1982) that in Sector B of Fig. 2,

$$\partial\sigma_{rr}/\partial\theta - 2\sigma_{r\theta} = -4k + o(1) \quad \text{as } r \to 0.$$

In other words, use of these facts reveals that the first and third terms of the right side of (3.7) combine to give $[-4k(rD_{rr}) + o(rD_{rr})]$ as $r \to 0$.

To assess the second term on the right side of (3.7), we recognize that a "centered fan" plastic sector moving with a growing tensile crack tip cannot share its *rear* boundary with a "constant stress" plastic sector. This is true since by invoking the requirement that the full stress tensor be continuous across quasi-statically moving surfaces in the material being analyzed (Drugan and Rice, 1984), it is easily proved that v_r would have to undergo a jump across such a boundary (the proof is very similar to one presented in Section 5.2 of Drugan et al. (1982)), and this jump would produce negative plastic work since $\sigma_{r\theta} = +k + o(1)$ as $r \to 0$ and $v_r > 0$ as $r \to 0$ in a "centered fan" sector at a *tensile* crack tip. Thus a moving "centered fan" sector must be bordered at its trailing edge by either a currently elastic sector or by a plastic sector with $s_{33} \neq 0$ at $r = 0$; in both of these sector types, the asymptotic analysis of Rice (1982) specifies that $v_\alpha = 0(\ln r)$ as $r \to 0$. Now since the results of Drugan and Rice (1984) require continuous v_θ across asymptotic (radial) sector boundaries, we conclude that $v_\theta = 0(\ln r)$ as $r \to 0$ in a moving "centered fan" sector. This fact together with (3.4), (3.6), (3.7) and (2.4) require that $v_r = 0(\ln r)$ in a moving "centered fan" sector. Thus,

$$D_{r\theta} \equiv (1/2r)[\partial v_r/\partial\theta - v_\theta + r\partial v_\theta/\partial r] = 0[(\ln r)/r] \quad \text{as} \quad r \to 0$$

in a moving "centered fan" sector. This conclusion coupled with (2.4) shows that

$$2rD_{r\theta}[-r(\partial\sigma_{rr}/\partial r + f_r)] = o(1) \quad \text{as} \quad r \to 0.$$

Therefore, $rD_{rr}(\partial\sigma_{rr}/\partial\theta - 2\sigma_{r\theta})$ dominates the other two terms on the right side of (3.7) as $r \to 0$ (since it cannot vanish as $r \to 0$ because the right side of (3.6) is $0(1)$ as $r \to 0$), and thus (3.5) is demonstrated.

Drugan et al. (1982) show how (3.4) and (3.5) are employed to determine the complete lowest-order deformation-rate and velocity fields in Sector B of Fig. 2; here I reproduce only their results for D^p_{ij}, which will be needed in the ensuing analysis:

$$\begin{aligned} D^p_{rr} &= D^p_{\theta\theta} = -(1-2\nu)(k/E)(\dot a/r)\sin\theta + o(1/r) \quad \text{as} \quad r \to 0 \\ D^p_{r\theta} &= [(5-4\nu)/(2\sqrt{2})](k/E)(\dot a/r)\ln(R/r) + o[(\ln r)/r] \quad \text{as} \quad r \to 0 \\ D^p_{33} &= 2(1-2\nu)(k/E)(\dot a/r)\sin\theta + o(1/r) \quad \text{as} \quad r \to 0 \\ D^p_{3\alpha} &= 0. \end{aligned} \tag{3.8}$$

4 Radial Dependence of the Stress Field

I shall now show that the radial behavior of the stress field at small but finite distances from the growing crack tip is directly determinable from the lowest-order solution of Drugan et al. (1982) coupled with the general governing equations of Section 2. In particular, explicit higher-order expressions will be derived for the stress field in Sectors A and B, the sectors in which the principal plastic deformation occurs. (The only other location for singular plastic strain accumulation is as $r \to 0$ along the crack flank, $\theta = \pm\pi$.)

4.1 Sector *B*. The deviatoric stress field in the "centered fan" plastic sector, Sector B of Fig. 2, was determined by Drugan et al. (1982) to be

$$\left.\begin{aligned} s_{rr} &= s_{\theta\theta} = s_{33} = 0 \\ s_{r\theta} &= k \end{aligned}\right\} \quad \text{at} \quad r = 0. \tag{4.1}$$

The manner in which this stress field alters as r increases from zero in Sector B can be derived from the lowest-order solution for the rates of plastic deformation, (3.8), together with the plastic part of the Prandtl-Reuss flow rule,

$$D^p_{ij} = \Lambda s_{ij}. \tag{4.2}$$

First, the parameter Λ is specified to lowest order in Sector B via the $r\theta$ components of (3.8), (4.1) and (4.2):

$$\begin{aligned} \Lambda &= \frac{D^p_{r\theta}}{s_{r\theta}} = \frac{[(5-4\nu)/(2\sqrt{2})](k/E)(\dot a/r)\ln(R/r) + o[(\ln r)/r]}{k + o(1)} \quad \text{as} \quad r \to 0 \\ &= [(5-4\nu)/(2\sqrt{2})](1/E)(\dot a/r)\ln(R/r) + o[(\ln r)/r] \quad \text{as} \quad r \to 0. \end{aligned} \tag{4.3}$$

Then, employing (4.3) and (3.8) in (4.2), one obtains

$$\begin{aligned} s_{rr} &= \frac{D^p_{rr}}{\Lambda} = \frac{-(1-2\nu)(k/E)\sin\theta(\dot a/r) + o(1/r)}{[(5-4\nu)/(2\sqrt{2})](1/E)(\dot a/r)\ln(R/r) + o[(\ln r)/r]} \quad \text{as} \quad r \to 0 \\ &= -2\sqrt{2}[(1-2\nu)/(5-4\nu)]k\sin\theta\left(\ln\frac{R}{r}\right)^{-1} + o[(\ln r)^{-1}] \quad \text{as} \quad r \to 0 \end{aligned} \tag{4.4a}$$

$$s_{\theta\theta} = \frac{D^p_{\theta\theta}}{\Lambda} = s_{rr} + o[(\ln r)^{-1}] \quad \text{as} \quad r \to 0 \tag{4.4b}$$

$$s_{33} = \frac{D^p_{33}}{\Lambda} = -2s_{rr} + o[(\ln r)^{-1}] \quad \text{as} \quad r \to 0. \tag{4.4c}$$

To determine the r-dependence of $s_{r\theta}$ for small r in Sector B, write

$$s_{r\theta} = k + g(r,\theta,t) \quad \text{as} \quad r \to 0 \tag{4.5}$$

where $g(r, \theta, t) \to 0$ as $r \to 0$. Next, rewrite the Huber-Mises yield condition (2.6) as

$$3s_{33}^2/4 + (s_{rr} - s_{\theta\theta})^2/4 + s_{r\theta}^2 = k^2 \tag{4.6}$$

and substitute from (4.4) and (4.5) to obtain

$$3\left\{32[(1-2\nu)/(5-4\nu)]^2k^2\sin^2\theta\left(\ln\frac{R}{r}\right)^{-2} + o[(\ln r)^{-2}]\right\}/4 + \{o[(\ln r)^{-2}]\}/4 + [k^2 + 2kg + g^2] = k^2.$$

Evidently this specifies g to $0[(\ln r)^{-2}]$, so (4.5) becomes

$$s_{r\theta} = k - 12[(1-2\nu)/(5-4\nu)]^2k\sin^2\theta\left(\ln\frac{R}{r}\right)^{-2} + o[(\ln r)^{-2}] \quad \text{as} \quad r \to 0. \tag{4.7}$$

That this deviatoric stress field satisfies equilibrium to the appropriate order in r is easily verified by rewriting (2.3) as

$$-\partial\sigma_{rr}/\partial\left(\ln\frac{R}{r}\right) = s_{\theta\theta} - s_{rr} - \partial s_{r\theta}/\partial\theta - rf_r \tag{4.8a}$$

$$\partial\sigma_{\theta\theta}/\partial\theta = -2s_{r\theta} + \partial s_{r\theta}/\partial\left(\ln\frac{R}{r}\right) - rf_\theta \tag{4.8b}$$

and assuming f_r, f_θ to be bounded as $r \to 0$. Substitution of (4.4) and (4.7) into (4.8) shows that (4.8) are satisfied to $0[(\ln r)^{-2}]$ as $r \to 0$ provided that the yet-unspecified hydrostatic stress component is chosen properly, as reflected in the derivation of (4.9) below. The deviatoric stress field given by (4.4) and (4.7) is thus the solution in Sector B to the order expressed, since it was shown to satisfy equilibrium (2.3) and yield (2.6) to $0[(\ln r)^{-2}]$ as $r \to 0$, and because it is derived from the lowest-order deformation rate field, which is in turn derived from a continuous velocity field and hence is compatible. Integration of the equilibrium equations at small r, using (4.4) and (4.7), shows that the stress field in Sector B must have the form

$$\begin{aligned} \sigma_{rr} &= C_f - 2k\theta + 6M^2k(2\theta + \sin 2\theta)\left(\ln\frac{R}{r}\right)^{-2} + f(r,t)/r + o[(\ln r)^{-2}] \quad \text{as} \quad r \to 0 \end{aligned} \tag{4.9a}$$

$$\begin{aligned} \sigma_{\theta\theta} &= C_f - 2k\theta + 6M^2k(2\theta - \sin 2\theta)\left(\ln\frac{R}{r}\right)^{-2} + \partial f(r,t)/\partial r + o[(\ln r)^{-2}] \quad \text{as} \quad r \to 0 \end{aligned} \tag{4.9b}$$

$$\sigma_{r\theta} = k - 12M^2 k \sin^2\theta \left(\ln \frac{R}{r}\right)^{-2} + o[(\ln r)^{-2}] \quad \text{as} \quad r \to 0 \tag{4.9c}$$

$$\sigma_{33} = C_f - 2k\theta + 6\sqrt{2}Mk\sin\theta \left(\ln \frac{R}{r}\right)^{-1} + [\partial f(r,t)/\partial r + f(r,t)/r]/2 + o[(\ln r)^{-1}] \quad \text{as} \quad r \to 0 \tag{4.9d}$$

where $M \equiv (1-2\nu)/(5-4\nu)$ and $f(r, t)$ is a function of integration having the properties that $\partial f(r, t)/\partial r = o(1)$ as $r \to 0$ and $f(r, t)/r = o(1)$ as $r \to 0$.

By *assuming* an Airy stress function to have an expansion of the form

$$\phi = r^2 \sum_{n=0}^{\infty} f_n(\theta) \left(\ln \frac{R}{r}\right)^{-n},$$

Gao (1980) arrives at expressions having the same functional forms in a moving centered fan sector as (4.9) when specialized to steady-state conditions; a complete comparison is not possible since all of his higher-order terms contain unspecified constants. See the discussion by Drugan et al. (1982) of the Gao (1980) study.

4.2 Sector A. The results of the preceding subsection were derived strictly for the "centered fan" plastic sector (Sector B of Fig. 2). However, since the stress field near a growing crack tip in a solid satisfying the governing equations of Section 2 must be fully continuous (Drugan and Rice, 1984), (4.9) imply the asymptotic form of the radial dependence of the stress field at all angles about the crack tip. In particular, full stress continuity across S_{AB} of Fig. 2, the boundary between Sectors A and B, implies that the stress field in Sector A has the form

$$\sigma_{ij} = \sigma_{ij}^{(0)}(\theta) + \sigma_{ij}^{(1)}(\theta,t)\xi^{-1} + \sigma_{ij}^{(2)}(\theta,t)\xi^{-2} + o(\xi^{-2}) \quad \text{as} \quad r \to 0 \tag{4.10}$$

where $\xi \equiv \ln(R/r)$; the adoption of (4.10) embodies the assumption that $f(r, t)$ of (4.9) is expressible as

$$f(r,t) = \alpha_1(t) r\xi^{-1} + \alpha_2(t) r\xi^{-2} + o(r\xi^{-2}) \quad \text{as} \quad r \to 0.$$

Similarly, we represent the deviatoric stress field in Sector A as

$$s_{ij} = s_{ij}^{(0)}(\theta) + s_{ij}^{(1)}(\theta,t)\xi^{-1} + s_{ij}^{(2)}(\theta,t)\xi^{-2} + o(\xi^{-2}) \quad \text{as} \quad r \to 0. \tag{4.11}$$

Substituting (4.10) into (2.2) and equating terms of like order in r, while employing the assumption of bounded body forces, gives the equilibrium requirements (to $0(\xi^{-2}/r)$ as $r \to 0$):

$$\sigma_{\alpha\beta}^{(0)\prime} h_\beta = 0 \tag{4.12a}$$

$$\sigma_{\alpha\beta}^{(1)\prime} h_\beta = 0 \tag{4.12b}$$

$$\sigma_{\alpha\beta}^{(2)\prime} h_\beta + \sigma_{\alpha\beta}^{(1)} e_\beta = 0 \tag{4.12c}$$

where the superscript $'$ denotes $\dfrac{\partial}{\partial\theta}$.

The most convenient form of the yield condition to apply here is the differential form (2.7b), which, upon substitution of (4.10) and (4.11), requires (to $0(\xi^{-3}/r)$ as $r \to 0$)

$$s_{ij}^{(0)}\sigma_{ij}^{(1)} = 0 \tag{4.13a}$$

$$2s_{ij}^{(0)}\sigma_{ij}^{(2)} + s_{ij}^{(1)}\sigma_{ij}^{(1)} = 0. \tag{4.13b}$$

Note that the other differential form of the yield condition, (2.7a), must also be satisfied in Sector A; however, a short calculation reveals that (2.7a) results in the same restrictions as do (4.13), but in less illuminating forms.

An asymptotic expression for the stress rate in Sector A for general, non-steady crack growth is obtained by substituting (4.10) into the result (2.11a), viz.

$$\dot{\sigma}_{ij} = \sigma_{ij}^{(0)\prime}\sin\theta(\dot{a}/r) + \sigma_{ij}^{(1)\prime}\sin\theta(\dot{a}/r)\xi^{-1} + [\sigma_{ij}^{(2)\prime}\sin\theta - \sigma_{ij}^{(1)}\cos\theta](\dot{a}/r)\xi^{-2} + o[\xi^{-2}/r] \quad \text{as} \quad r \to 0. \tag{4.14}$$

Drugan et al. (1982) determined the lowest-order stress field in Sector A to be

$$\sigma_{ij}^{(0)\prime} = 0 \tag{4.15a}$$

$$\sigma_{12}^{(0)} = 0 \tag{4.15b}$$

$$\sigma_{33}^{(0)} = (\sigma_{11}^{(0)} + \sigma_{22}^{(0)})/2 \tag{4.15c}$$

$$\sigma_{22}^{(0)} - \sigma_{11}^{(0)} = 2k \tag{4.15d}$$

$$\sigma_{11}^{(0)} + \sigma_{22}^{(0)} = \text{known constant for a specific } \nu \text{ value} \tag{4.15e}$$

which of course satisfy equilibrium (4.12a) and the yield condition (2.6) at $r = 0$.

The $\sigma_{ij}^{(0)}$ of (4.15) are employed in the higher-order equations of equilibrium (4.12), yield (4.13) and compatibility (2.15) (written in terms of stresses via (2.9), (2.13) and (4.14)). The resulting system of differential equations, when augmented by the Mode I symmetry conditions and the boundary conditions on Sector A, permits specification of the higher-order coefficients $\sigma_{ij}^{(1)}(\theta, t)$ and $\sigma_{ij}^{(2)}(\theta, t)$ of (4.10). The derivation is lengthy but fairly straightforward and is thus omitted here. Making use of these results, the stress field in Sector A of Fig. 2 is expressible to $0[(\ln r)^{-2}]$ as

$$\sigma_{11} = P_{11} + Q_{11}(t)\left(\ln \frac{R}{r}\right)^{-1} + \left\{(Q_{11}(t)/2)\ln(\cos 2\theta) + R_{11}(t)\right\}\left(\ln \frac{R}{r}\right)^{-2} + o[(\ln r)^{-2}] \quad \text{as} \quad r \to 0 \tag{4.16a}$$

$$\sigma_{22} = P_{22} + Q_{11}(t)\left(\ln \frac{R}{r}\right)^{-1} + \left\{(Q_{11}(t)/2)\ln(\cos 2\theta) + R_{11}(t) - 12M^2k\right\}\left(\ln \frac{R}{r}\right)^{-2} + o[(\ln r)^{-2}] \quad \text{as} \quad r \to 0 \tag{4.16b}$$

$$\sigma_{12} = -(Q_{11}(t)/2)\ln[\tan(\pi/4+\theta)]\left(\ln \frac{R}{r}\right)^{-2} + o[(\ln r)^{-2}] \quad \text{as} \quad r \to 0 \tag{4.16c}$$

$$\sigma_{33} = P_{33} + [Q_{11}(t) + 6kM]\left(\ln \frac{R}{r}\right)^{-1} + \left\{[Q_{11}(t)/2 + 3kM]\ln(\cos 2\theta) + R_{33}(t)\right\}\left(\ln \frac{R}{r}\right)^{-2} + o[(\ln r)^{-2}] \quad \text{as} \quad r \to 0, \tag{4.16d}$$

where P_{11}, P_{22}, P_{33} are known constants for a specific ν value (i.e., $P_{ij} = \sigma_{ij}^{(0)}$ of (4.15)), and $Q_{11}(t)$, $R_{11}(t)$ and $R_{33}(t)$ are undetermined functions of integration. ($Q_{11}(t)$ and $R_{11}(t)$, being proportional to the $0[(\ln r)^{-1}]$ and $0[(\ln r)^{-2}]$ stress triaxiality coefficients, respectively, can be determined only by complete solution of the higher-order fields at *all* angles about the crack tip.)

Because all stress components are required to be bounded near the crack tip in the elastic-ideally plastic constitutive model employed (Drugan, 1985), *the results (4.16) show that the boundary of Sector A must curve for $r > 0$.* That is, representing this boundary for small $r \geq 0$ as

$$\theta = \pi/4 + m(r,t) \tag{4.17}$$

where $m(r, t) \to 0$ as $r \to 0$ in view of the lowest-order solution, (4.16) require that $m(r, t) < 0$ for $r > 0$. This is evident in spite of the fact that the value of $Q_{11}(t)$ is undetermined, since (4.16) show that at least one component of σ_{ij} becomes infinite as $\theta \to \pi/4$ for *any* value of $Q_{11}(t)$. A more explicit restriction on the asymptotic behavior of $m(r, t)$ is obtained by enforcing continuity of $\sigma_{r\theta}$ to $0[(\ln r)^{-2}]$ across S_{AB}, which via (4.9*c*), (4.16) and (4.17) can be shown to require

$$m(r,t) = o\left[\left(\ln \frac{R}{r}\right)^{-1}\right] \quad \text{as} \quad r \rightarrow 0.$$

5 Deformation Field in the Constant Stress Sector

Drugan et al. (1982) show that in Sector A of Fig. 2

$$D_{ij} = o(1/r) \quad \text{as} \quad r \rightarrow 0 \tag{5.1}$$

but their lowest-order solution does not specify D_{ij} further. The results of Section 4.2 permit the structure of D_{ij} in Sector A to be determined, and from this the velocity field and plastic strain field in this sector may be derived.

For $r > 0$ in Sector A (so that $s_{33} \neq 0$), Λ is given by (2.13), so that the Prandtl-Reuss flow rule (2.9) becomes

$$ED_{ij} = (1+\nu)\dot{\sigma}_{ij} - \nu\delta_{ij}\dot{\sigma}_{kk} - [\dot{\sigma}_{33} - \nu(\dot{\sigma}_{11} + \dot{\sigma}_{22})](s_{ij}/s_{33}). \tag{5.2}$$

Employing the Sector A stress field results (4.16), together with the asymptotic stress rate expression (4.14), (5.2) provides D_{ij} in Sector A to higher order:

$$\begin{aligned} D_{ij} &= D^p_{ij} + 0[(\ln r)^{-2}/r] \\ &= (1/4E)[(5-4\nu)Q_{11}(t) + 6k](\delta_{2i}\delta_{2j} - \delta_{1i}\delta_{1j}) \\ &\quad \times (\cos\theta/\cos 2\theta)(\dot{a}/r)\left(\ln \frac{R}{r}\right)^{-1} + 0[(\ln r)^{-2}/r] \text{ as } r \rightarrow 0. \end{aligned} \tag{5.3}$$

To determine the asymptotic material velocity field, we begin by integrating the *rr* component of (5.3) for small r, recalling that $D_{rr} \equiv \partial v_r/\partial r$, to obtain

$$\begin{aligned} v_r = {} & (\dot{a}/4E)[(5-4\nu)Q_{11}(t) + 6k]\cos\theta \ln\left(\ln \frac{R}{r}\right) \\ & + \partial F(\theta,t)/\partial\theta + 0[(\ln r)^{-1}] \quad \text{as} \quad r \rightarrow 0 \end{aligned} \tag{5.4}$$

where $\partial F(\theta, t)/\partial\theta$ is an undetermined function of integration. Next, the $\theta\theta$ component of (5.3) is employed, using $D_{\theta\theta} \equiv (1/r)(\partial v_\theta/\partial\theta + v_r)$. Substituting from (5.4) and integrating with respect to θ gives

$$\begin{aligned} v_\theta = {} & -(\dot{a}/4E)[(5-4\nu)Q_{11}(t) + 6k]\sin\theta \ln\left(\ln \frac{R}{r}\right) \\ & - F(\theta,t) + G(r,t) + 0[(\ln r)^{-1}] \quad \text{as} \quad r \rightarrow 0, \end{aligned} \tag{5.5}$$

where $G(r, t)$ is an undetermined function of integration.

Now, Mode I symmetry requires that

$$\partial v_r/\partial\theta = v_\theta = 0 \quad \text{on} \quad \theta = 0,$$

which via (5.4) and (5.5) yield the restrictions

$$\left.\begin{aligned} &\partial^2 F(\theta,t)/\partial\theta^2 = 0 && (5.6a) \\ &G(r,t) = F(\theta,t) + 0[(\ln r)^{-1}] \quad \text{as} \quad r \rightarrow 0. && (5.6b) \end{aligned}\right\} \text{on} \quad \theta = 0$$

Finally, (5.4) and (5.5) must satisfy the $r\theta$ component of (5.3), expressed in terms of velocities via $D_{r\theta} \equiv (1/2r)[\partial v_r/\partial\theta - v_\theta + r\partial v_\theta/\partial r]$; this will be satisfied to $0(1/r)$ as $r \rightarrow 0$ provided that

$$\partial^2 F(\theta,t)/\partial\theta^2 + F(\theta,t) - F(0,t) = 0.$$

Thus,

$$F(\theta,t) = B(t)\sin\theta + C(t)$$

where $B(t)$ and $C(t)$ are undetermined functions, and the velocity field to $0(1)$ as $r \rightarrow 0$ in Sector A specifies to

$$v_r = (\dot{a}/4E)[(5-4\nu)Q_{11}(t) + 6k]\cos\theta \ln\left(\ln \frac{R}{r}\right) + B(t)\cos\theta$$

$$v_\theta = -(\dot{a}/4E)[(5-4\nu)Q_{11}(t) + 6k]\sin\theta \ln\left(\ln \frac{R}{r}\right) - B(t)\sin\theta.$$

The plastic strain accumulated by a material point as it passes through Sector A can be determined by integrating (5.3) at that material point, i.e.,

$$\epsilon^p_{ij} = \int_{t_o}^{t} D^p_{ij} d\tau, \tag{5.7}$$

where t_o is the time at which the material point enters the plastically deforming region. Drugan et al. (1982) showed that when a is increasing continuously with applied loading, if $D^p_{ij} = 0(\dot{a}/r)$ as $r \rightarrow 0$ in a plastically deforming sector, then the contribution to the corresponding ϵ^p_{ij} from that sector will be bounded for $\theta \neq 0$. Thus we immediately deduce from (5.3) that since $D^p_{ij} = o(\dot{a}/r)$ as $r \rightarrow 0$,

$$\epsilon^p_{ij} = 0(1) \quad \text{as} \quad r \rightarrow 0 \quad \text{when} \quad \theta \neq 0.$$

However, (5.3) shows that a material point located directly on the crack line has

$$\begin{aligned} D^p_{ij} \equiv \dot{\epsilon}^p_{ij} = {} & (1/4E)[(5-4\nu)Q_{11}(t) + 6k](\delta_{2i}\delta_{2j} \\ & - \delta_{1i}\delta_{1j})(\dot{a}/r)\left(\ln \frac{R}{r}\right)^{-1} + 0[(\ln r)^{-2}/r] \\ & \text{as} \quad r \rightarrow 0 \quad \text{along} \quad \theta = 0. \end{aligned} \tag{5.8}$$

When a is increasing continuously with applied loading, we employ the relationship $\dot{r} = -\dot{a}$ (valid on $\theta = 0$) to substitute for $\dot{a}$ in (5.8) and then integrate to find

$$\begin{aligned} \epsilon^p_{11} = -\epsilon^p_{22} = {} & (1/4E)[(5-4\nu)Q_{11}(t) + 6k]\ln\left(\ln \frac{R}{r}\right) \\ & + 0(1) \quad \text{as} \quad r \rightarrow 0 \quad \text{along} \quad \theta = 0. \end{aligned}$$

This reveals that two components of the plastic strain at a material point on the crack line may (depending on the value of $Q_{11}(t)$) become infinite as the crack tip approaches.

Acknowledgments

Support of this work by the U.S. National Science Foundation, Solid Mechanics Program, under Grant MEA-8312348, and by the US Department of Energy through Contract DE-AC02-80ER/10556 is gratefully acknowledged. I thank Professor J. R. Rice, Harvard University, for valuable discussions, and I am grateful to a conscientious reviewer for detailed, helpful comments.

References

Drugan, W. J., 1985, "On the Asymptotic Continuum Analysis of Quasi-Static Elastic-Plastic Crack Growth and Related Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 601-605.

Drugan, W. J., and Rice, J. R., 1984, "Restrictions on Quasi-Statically Moving Surfaces of Strong Discontinuity in Elastic-Plastic Solids," in *Mechanics of Material Behavior: The D. C. Drucker Anniversary Volume*, G. J. Dvorak and R. T. Shield, eds., Elsevier Science Publishers, Amsterdam, pp. 59-73.

Drugan, W. J., Rice, J. R., and Sham, T.-L., 1982, "Asymptotic Analysis of Growing Plane Strain Tensile Cracks in Elastic-Ideally Plastic Solids," *Journal of the Mechanics and Physics of Solids*, Vol. 30, pp. 447-473; Vol. 31, 1983, p. 191.

Gao, Y.-C., 1980, "Elastic-Plastic Fields at Crack-Tips in Perfectly Plastic Medium," presented at 15th ICTAM, Toronto, Aug. 1980; also, 1981, *Acta Mechanica Sinica*.

Gao, Y.-C., 1983, "Influence of Compressibility on the Elastic-Plastic Field of a Growing Crack," *Elastic-Plastic Fracture: Second Symposium, Volume I—Inelastic Crack Analysis, ASTM STP 803*, C. F. Shih and J. P. Gudas, eds., American Society for Testing and Materials, Philadelphia, pp. 176-190.

Hutchinson, J. W., 1982, "Crack-Tip Singularity Fields in Nonlinear Fracture Mechanics: A Survey of Current Status," in *Advances in Fracture Research*, D. Francois et al., eds., Pergamon Press, Oxford, pp. 2669-2684.

McMeeking, R. M., 1977, "Finite Deformation Analysis of Crack-Tip Opening in Elastic-Plastic Materials and Implications for Fracture," *Journal of the Mechanics and Physics of Solids*, Vol. 25, pp. 357-381.

Pan, Hao, 1982, private communication.

Rice, J. R., 1982, "Elastic-Plastic Crack Growth," in *Mechanics of Solids: The R. Hill 60th Anniversary Volume*, H. G. Hopkins and M. J. Sewell, eds., Pergamon Press, Oxford, pp. 539-562.

Van Dyke, M., 1975, *Perturbation Methods in Fluid Mechanics*, The Parabolic Press, Stanford, p. 23.

Three-Dimensional Stress Singularities at Conical Notches and Inclusions in Transversely Isotropic Materials

Nihal Somaratna
Engineering Technology Associates,
Southfield, MI 48076
Assoc. Mem. ASME

T. C. T. Ting
Professor of Applied Mechanics.
Department of Civil Engineering,
Mechanics and Metallurgy,
University of Illinois at Chicago,
Chicago, IL 60680
Fellow ASME

This study examines analytically the possible existence of stress singularities of the form $\sigma = \rho^\delta f(\theta,\phi)$ *at the apex of axisymmetric conical boundaries in transversely isotropic materials.* (ρ, θ, ϕ) *refer to spherical coordinates with the origin at the apex. The problems of one conical boundary and of two conical boundaries with a common apex are considered. The boundaries are either rigidly clamped or traction free. Separation of variables enables the general solution to be expressed in terms of Legendre functions of the first and second kind. Imposition of boundary conditions leads to an eigenequation that would determine possible values of* δ. *The degenerate case that arises when the eigenvalues of the elasiticity constants are identical is also discussed. Isotropic materials constitute only a particular case in this class of degenerate materials and previously reported eigenequations corresponding to isotropic materials are shown to be recoverable from the present results. Numerical results corresponding to a few selected cases are also presented to illustrate the solution procedure.*

1 Introduction

The occurrence of stress singularities in three-dimensional elasticity problems (e.g., at the apex of notches and inclusions; along crack fronts; and interfaces between dissimilar materials) assumes importance in studies of the fracture behavior of materials as well as in general stress analysis (Sih, 1971; Kassir and Sih, 1975; Erdogan, 1983). Most common is the power type singularity where the stress σ assumes the form $\sigma = \rho^\delta f(\theta, \phi)$, where (ρ, θ, ϕ) refer to a spherical coordinate system with the origin at the point of singularity and $\mathrm{Re}(\delta) < 0$, so that the stresses become unbounded at the origin. In fracture mechanics it is desirable to know the order of the singularity δ. As far as three-dimensional stress analysis is concerned a complete analytical solution for a problem containing a crack or a notch is very difficult to obtain and numerical methods such as finite element techniques have to be employed. In the vicinity of stress singularities special numerical methods are preferred as they improve accuracy substantially. One of the often used methods is to employ a "special finite element" which can adequately represent the singular nature of the stresses in that region (Pian et al., 1972; Tong et al., 1973; Lin and Mar, 1976; Rhee and Atluri, 1982). This requires primarily a knowledge of the value of δ and, if possible, of $f(\theta, \phi)$ as well.

In the case of a crack, a notch or any other such "source" of a singularity of arbitrary geometry in a three-dimensional medium, the task of obtaining an analytical solution for δ seems mathematically intractable for isotropic materials and even more so for anisotropic materials. For isotropic materials, semi-analytical methods have been proposed by Benthem (1977, 1980) and Kawai et al. (1977) while numerical methods have been used by Bazant (1974) and Bazant and Estenssoro (1977, 1983). For anisotropic materials a numerical procedure has been recently proposed by Somaratna and Ting (1986). For any numerical technique it is desirable to have analytical solutions to at least a few simplified cases so that they may serve as bench-mark tests for the numerical method and provide useful checks on its accuracy. In the case of anisotropic materials the simplifications are achieved by making the material constitutive relation and the geometry of the problem simpler. In the particular case of assuming the geometry to be axisymmetric and the material to be transversely isotropic, the problem can in fact be made mathematically two-dimensional. In the present paper we investigate the occurrence of stress singularities in that simplified problem.

The material is assumed to be linearly elastic and transversely isotropic with respect to the z axis in the cylindrical coordinates (r, ϕ, z). In the spherical coordinate system (ρ, θ, ϕ) the material is bounded by either one (Fig. 1*a*) or two (Fig. 1*b*) conical surfaces defined by $\theta = \theta_1$ and $\theta = \theta_2$. On the boundary or boundaries the surface is either traction free

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami, Fla., November 17-21, 1985 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February 21, 1985; final revision, May 31, 1985. Paper No. 85-WA/APM-25.

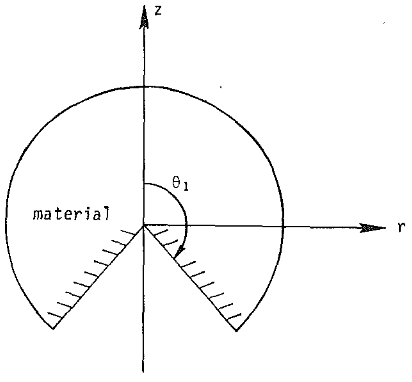


Fig. 1(*a*) One axisymmetric conical boundary

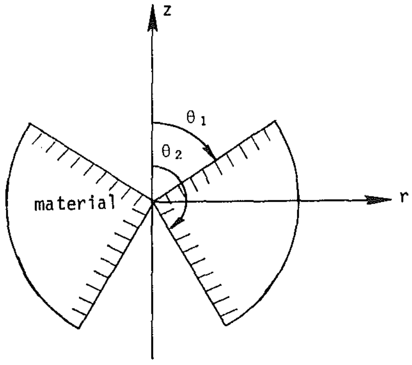


Fig. 1(*b*) Two axisymmetric conical boundaries

or rigidly clamped. The former boundary condition signifies a notch while the latter corresponds to a rigid inclusion. If necessary the method presented here can be used to analyze "mixed" boundary conditions (e.g., a frictionless rigid inclusion). A special case of this problem where the material is assumed to be isotropic and bounded by one conical surface has been analyzed by Bazant and Keer (1974) using results of Thompson and Little (1970). A proper reduction of the more general solution presented here, of course, leads to the recovery of the results reported therein.

The mathematical formulation of the basic equations for a transversely isotropic elastic material under an axisymmetric deformation is presented in Section 2. Separation of variables enables the general solution, in which stresses are proportional to ρ^δ, to be expressed in terms of Legendre functions of the first and second kind. This is described in Section 3 while in Section 4 appropriate boundary conditions are imposed on the general solution to yield an eigenequation for δ. It is also shown that under a certain special class of material properties the general solution degenerates. This degeneracy and the modifications of the solution required to overcome the difficulty arising therefrom are discussed in Section 5. Isotropic materials are only a particular case in this special class of degenerate materials and the results corresponding to isotropic materials are recovered in Section 6. They agree with those reported by Thompson and Little (1970) and Bazant and Keer (1974). Section 7 presents some numerical results obtained by solving the eigenequation.

2 Mathematical Formulation

Let (r, ϕ, z) be a cylindrical coordinate system with the z axis as the axis of material symmetry and let (u_r, u_ϕ, u_z) be the corresponding displacement components. We assume that the deformation is axisymmetric and $u_\phi = 0$ so that u_r and u_z are functions of r and z only. Introducing the displacement potential $\Phi(r, z)$ which gives u_r and u_z (Elliott, 1948, 1949; Green and Zerna, 1975; Ting et al., 1985) by

$$u_r = \frac{\partial \Phi}{\partial r}, \quad u_z = m\frac{\partial \Phi}{\partial z} \tag{1}$$

where m is a constant to be determined, the stresses are obtained as

$$\left.\begin{aligned}
\sigma_r &= c_{11}\frac{\partial^2\Phi}{\partial r^2} + c_{12}\frac{\partial\Phi}{r\partial r} + c_{13}m\frac{\partial^2\Phi}{\partial z^2}\\
\sigma_\phi &= c_{12}\frac{\partial^2\Phi}{\partial r^2} + c_{11}\frac{\partial\Phi}{r\partial r} + c_{13}m\frac{\partial^2\Phi}{\partial z^2}\\
\sigma_z &= c_{13}\frac{\partial^2\Phi}{\partial r^2} + c_{13}\frac{\partial\Phi}{r\partial r} + c_{33}m\frac{\partial^2\Phi}{\partial z^2}\\
\sigma_{rz} &= c_{44}(1+m)\frac{\partial^2\Phi}{\partial r\partial z}
\end{aligned}\right\} \tag{2}$$

in which c_{ij} are the elasticity constants for the tranversely isotropic material. The equations of equilibrium are satisfied if

$$\frac{\partial^2\Phi}{\partial r^2} + \frac{\partial\Phi}{r\partial r} - \frac{1}{p^2}\frac{\partial^2\Phi}{\partial z^2} = 0 \tag{3}$$

where

$$p^2 = \frac{-c_{11}}{mc_{13} + (1+m)c_{44}} = \frac{c_{13} + (1+m)c_{44}}{-mc_{33}} \tag{4a}$$

or, equivalently,

$$-m = \frac{c_{11} + c_{44}p^2}{(c_{13} + c_{44})p^2} = \frac{c_{13} + c_{44}}{c_{44} + c_{33}p^2} \tag{4b}$$

The second equality of (4*a*) and (4*b*), respectively, yields

$$m^2 - 2\left[\frac{c_{11}c_{33} - c_{13}^2}{2c_{44}(c_{13} + c_{44})} - 1\right]m + 1 = 0 \tag{5a}$$

$$p^4 + 2\left[\frac{c_{11}c_{33} - c_{13}^2 - 2c_{13}c_{44}}{2c_{33}c_{44}}\right]p^2 + \frac{c_{11}}{c_{33}} = 0 \tag{5b}$$

Notice that m and p are independent of the elasticity constant c_{12}. Therefore, if the boundary conditions are prescribed in terms of the displacements Φ is independent of c_{12}.

It can be shown (Eshelby et al., 1953) that p cannot be real if the strain energy is positive definite. Therefore we have two pairs of complex conjugates for p and let them be denoted by p_1, $\bar{p}_1$, p_2 and $\bar{p}_2$ where an overbar indicates the complex conjugate. The associated values of m are denoted by m_1, $\bar{m}_1$, m_2 and $\bar{m}_2$ respectively. From (5*a*) we note that

$$m_1 m_2 = 1 \tag{5c}$$

Since (5*b*) is a quadratic equation in p^2 with real coefficients, if p_1 is purely imaginary so is p_2. Then $\bar{p}_1 = -p_1$, $\bar{p}_2 = -p_2$ and m_1, m_2 are real and satisfy $m_1 = \bar{m}_1$, $m_2 = \bar{m}_2$. If p_1 and p_2 are not purely imaginary we can choose

$$p_1 = u + iv = -\bar{p}_2; \quad p_2 = -u + iv = -\bar{p}_1 \tag{6}$$

where u, v are real. In this case m_1 and m_2 are complex and $m_1 = \bar{m}_2$. In view of the fact that the problem is linear, the general solution for displacements and stresses is obtained by superposing the solutions corresponding to Φ's associated with p_1, $\bar{p}_1$, p_2 and $\bar{p}_2$. In the next Section where we present the general solution it is tacitly assumed that $p_1 \neq p_2$. The degenerate case where $p_1 = p_2$ is discussed separately in Section 5.

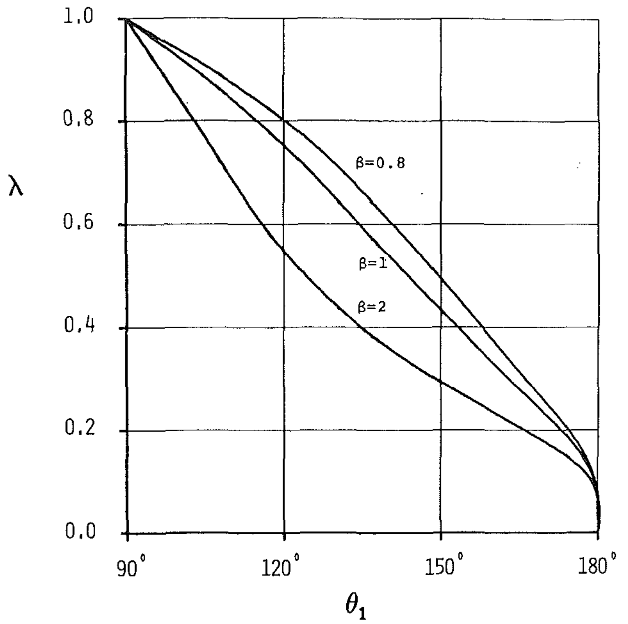


Fig. 2 **Value of λ for a rigid conical inclusion in "degenerate" materials with $\nu = 0.3$ (λ is independent of γ for this case)**

3 Solution

First transform equation (3) into the cylindrical harmonic equation by letting

$$Z = -ipz \tag{7}$$

For isotropic materials $p = i$ is an eigenvalue of the elasticity constants and we have $Z = z$. With (7), (3) becomes

$$\frac{\partial^2 \Phi}{\partial r^2} + \frac{1}{r}\frac{\partial \Phi}{\partial r} + \frac{\partial^2 \Phi}{\partial Z^2} = 0 \tag{8}$$

To obtain a solution of the harmonic equation (8) we let

$$\left.\begin{aligned} R &= (r^2 + Z^2)^{1/2} \\ c &= \cos\psi = Z/R \\ s &= \sin\psi = r/R \end{aligned}\right\} \tag{9}$$

where (R, ψ) are similar to axisymmetric spherical coordinates. It is well known (e.g., Lebedev, 1972) that (8) has two linearly independent solutions given by

$$\Phi = A\, R^{\lambda+1} P_{\lambda+1}(c) + B\, R^{\lambda+1} Q_{\lambda+1}(c) \tag{10}$$

where $P_{\lambda+1}(c)$ and $Q_{\lambda+1}(c)$ are the Legendre functions of the first and second kind respectively and their argument c is defined in (9), while λ is an arbitrary number, real or complex, and A and B are arbitrary multiplicative constants.

In the definition of R in (9) there is a certain ambiguity. R in general is complex and if R is admissible so is $-R$. In fact this would give rise to two more solutions: $\Phi = \bar{A}(-R)^{\lambda+1} P_{\lambda+1}(-c)$ and $\Phi = \bar{B}(-R)^{\lambda+1} Q_{\lambda+1}(-c)$. However $P_{\lambda+1}(-c)$ and $Q_{\lambda+1}(-c)$ are linearly dependent on $P_{\lambda+1}(c)$ and $Q_{\lambda+1}(c)$ (Lebedev, 1972) and the $(-1)^{\lambda+1}$ can be absorbed into $\bar{A}$ and $\bar{B}$. Therefore they do not provide any new independent solutions and can be ignored. By similar considerations it can easily be shown that the use of $-p$ instead of p as the eigenvalue in (7) would not yield any new independent solutions. Therefore one can ignore $\bar{p}_1$ and $\bar{p}_2$ which are related to $-p_1$ and $-p_2$ as discussed earlier and the general solution is a linear combination of the solutions corresponding to Φ's as given in (10) that are associated with p_1 and p_2. Denoting their associations with p_1 and p_2 by subscripts 1 and 2, respectively, we write

$$\left.\begin{aligned} \Phi_1 &= A_1 R_1^{\lambda+1} P_{\lambda+1}(c_1) + B_1 R_1^{\lambda+1} Q_{\lambda+1}(c_1) \\ \Phi_2 &= A_2 R_2^{\lambda+1} P_{\lambda+1}(c_2) + B_2 R_2^{\lambda+1} Q_{\lambda+1}(c_2) \end{aligned}\right\} \tag{11}$$

where A_1, B_1, A_2, B_2 and λ are arbitrary constants. The general solution for the displacements is

$$u_r = \frac{\partial \Phi_1}{\partial r} + \frac{\partial \Phi_2}{\partial r}, \; u_z = m_1 \frac{\partial \Phi_1}{\partial z} + m_2 \frac{\partial \Phi_2}{\partial z}. \tag{12}$$

Similar expressions hold for the stresses.

Equations (1) and (2) which are used to obtain the expressions for the displacements and stresses corresponding to the solutions given by (11) require the evaluation of certain derivatives of Φ. First consider

$$\Phi = R^{\lambda+1} P_{\lambda+1}(c) \tag{13}$$

The necessary derivatives of the above Φ are obtained by using the differentiation rule and the recurrence relations corresponding to Legendre functions (Lebedev, 1972) and the relations given by (9). The results are:

$$\left.\begin{aligned} \frac{\partial \Phi}{\partial r} &= R^{\lambda}\left[\frac{\lambda}{s}(cP_\lambda - P_{\lambda-1})\right] \\ \frac{\partial \Phi}{\partial z} &= R^{\lambda}[(\lambda+1)(-ip)P_\lambda] \\ \frac{\partial^2 \Phi}{\partial r^2} &= R^{\lambda-1}\left[-\frac{\lambda c}{s^2}P_\lambda - \lambda\left\{(\lambda+1) - \frac{1}{s^2}\right\}P_{\lambda-1}\right] \\ \frac{\partial^2 \Phi}{\partial z^2} &= R^{\lambda-1}[-\lambda(\lambda+1)p^2 P_{\lambda-1}] \\ \frac{\partial^2 \Phi}{\partial r \partial z} &= R^{\lambda-1}\left[\frac{\lambda(\lambda+1)(-ip)}{s}(P_\lambda - cP_{\lambda-1})\right] \end{aligned}\right\} \tag{14}$$

In (14) as well as in the following Sections it is understood that the argument of P_λ, $P_{\lambda-1}$, Q_λ and $Q_{\lambda-1}$ is c. The rule for differentiating Q_λ and its recurrence relations are identical to those for P_λ. Hence the derivatives corresponding to $\Phi = R^{\lambda+1} Q_{\lambda+1}(c)$ are obtained simply by replacing P_λ and $P_{\lambda-1}$ in (14) by Q_λ and $Q_{\lambda-1}$, respectively.

When the derivatives of (14) are substituted into (1) and (2) we obtain expressions of the following form

$$\left.\begin{aligned} u_r &= u_r^{(P)} = R^{\lambda}\tilde{u}_r^{(P)} \\ u_z &= u_z^{(P)} = R^{\lambda}\tilde{u}_z^{(P)} \\ \sigma_r &= \sigma_r^{(P)} = R^{\lambda-1}\tilde{\sigma}_r^{(P)} \\ \sigma_\phi &= \sigma_\phi^{(P)} = R^{\lambda-1}\tilde{\sigma}_\phi^{(P)} \\ \sigma_z &= \sigma_z^{(P)} = R^{\lambda-1}\tilde{\sigma}_z^{(P)} \\ \sigma_{rz} &= \sigma_{rz}^{(P)} = R^{\lambda-1}\tilde{\sigma}_{rz}^{(P)} \end{aligned}\right\} \tag{15}$$

where the superscript P refers to the fact that these correspond to the solution $\Phi = R^{\lambda+1} P_{\lambda+1}(c)$. In an identical manner similar expressions corresponding to the solution $\Phi = R^{\lambda+1} Q_{\lambda+1}(c)$ can be obtained. Those relationships also may be expressed by equation (15) itself by merely replacing superscript P by superscript Q. Expressions for the terms $\tilde{u}_r^{(P)}, \ldots$ etc. in (15) do not contain R.

Hereafter a subscript 1 or 2 (following a comma if other subscripts are present) will be used to indicate whether the quantity is associated with the eigenvalue p_1 or p_2 of the material properties. The exceptions are θ_1 and θ_2 which are the values of the coordinate θ at the conical boundaries.

With the above notation, the general solution for displacements and stresses corresponding to Φ's given by (11) is

$$\left.\begin{aligned}
u_r &= A_1 R_1^{\lambda}\tilde{u}_{r,1}^{(P)} + B_1 R_1^{\lambda}\tilde{u}_{r,1}^{(Q)} + A_2 R_2^{\lambda}\tilde{u}_{r,2}^{(P)} + B_2 R_2^{\lambda}\tilde{u}_{r,2}^{(Q)}\\
u_z &= A_1 R_1^{\lambda}\tilde{u}_{z,1}^{(P)} + B_1 R_1^{\lambda}\tilde{u}_{z,1}^{(Q)} + A_2 R_2^{\lambda}\tilde{u}_{z,2}^{(P)} + B_2 R_2^{\lambda}\tilde{u}_{z,2}^{(Q)}\\
\sigma_r &= A_1 R_1^{\lambda-1}\tilde{\sigma}_{r,1}^{(P)} + B_1 R_1^{\lambda-1}\tilde{\sigma}_{r,1}^{(Q)}\\
&\quad + A_2 R_2^{\lambda-1}\tilde{\sigma}_{r,2}^{(P)} + B_2 R_2^{\lambda-1}\tilde{\sigma}_{r,2}^{(Q)}\\
\sigma_\phi &= A_1 R_1^{\lambda-1}\tilde{\sigma}_{\phi,1}^{(P)} + B_1 R_1^{\lambda-1}\tilde{\sigma}_{\phi,1}^{(Q)}\\
&\quad + A_2 R_2^{\lambda-1}\tilde{\sigma}_{\phi,2}^{(P)} + B_2 R_2^{\lambda-1}\tilde{\sigma}_{\phi,2}^{(Q)}\\
\sigma_z &= A_1 R_1^{\lambda-1}\tilde{\sigma}_{z,1}^{(P)} + B_1 R_1^{\lambda-1}\tilde{\sigma}_{z,1}^{(Q)}\\
&\quad + A_2 R_2^{\lambda-1}\tilde{\sigma}_{z,2}^{(P)} + B_2 R_2^{\lambda-1}\tilde{\sigma}_{z,2}^{(Q)}\\
\sigma_{rz} &= A_1 R_1^{\lambda-1}\tilde{\sigma}_{rz,1}^{(P)} + B_1 R_1^{\lambda-1}\tilde{\sigma}_{rz,1}^{(Q)}\\
&\quad + A_2 R_2^{\lambda-1}\tilde{\sigma}_{rz,2}^{(P)} + B_2 R_2^{\lambda-1}\tilde{\sigma}_{rz,2}^{(Q)}
\end{aligned}\right\}\qquad(16)$$

Further, from (7) and (9), R can be expressed as

$$R=\rho\zeta \qquad (17a)$$

$$\zeta=(\sin^2\theta-p^2\cos^2\theta)^{1/2} \qquad (17b)$$

$$r=\rho\sin\theta,\quad z=\rho\cos\theta \qquad (17c)$$

where $(\rho,\ \theta,\ \phi)$ refer to the spherical coordinates. Then it is clear that the stresses are proportional to ρ^{δ} where $\delta=\lambda-1$ is the order of stress singularity. When $\lambda<1$, the stress is singular at $\rho=0$. However, for the strain energy to be bounded at the origin we require that $\lambda>-1/2$. On the other hand if no concentrated force applies at the vertex, then boundedness of the displacements at $\rho=0$ demands that $\lambda>0$. Therefore the possibility of a stress singularity exists if $0<\lambda<1$. In the case of λ being complex this requirement becomes $0<\mathrm{Re}(\lambda)<1$ where Re refers to the real part.

4 Boundary Conditions and the Eigenequation

The boundary conditions applicable at $\theta=\theta_1$ and (in the case of two boundaries) at $\theta=\theta_2$ are as follows: For the rigidly connected boundary

$$u_r=0,\quad u_z=0, \qquad (18a)$$

and for the traction free boundary

$$\left.\begin{aligned}\sigma_r\cos\theta-\sigma_{rz}\sin\theta=0\\ \sigma_{rz}\cos\theta-\sigma_z\sin\theta=0\end{aligned}\right\}\qquad(18b)$$

4.1 Problem (A): One Conical Boundary. In this case the domain of validity of the solution includes the positive z axis where c assumes the value of 1. But $Q_\lambda(c)$ is unbounded at $c=1$ and in order to obtain a finite solution along the z axis we require that $B_1=B_2=0$. Then there are only two arbitrary multiplicative constants in the general solution (16). By substituting (16) into (18a) or (18b) as the case may be and using (17), we obtain two simultaneous linear algebraic equations for A_1 and A_2 in the form

$$\mathbf{K}^{(P)}(\theta_1)\mathbf{q}=\mathbf{0} \qquad (19)$$

In (19),

$$\mathbf{q}=\begin{bmatrix}A_1\\ A_2\end{bmatrix} \qquad (20a)$$

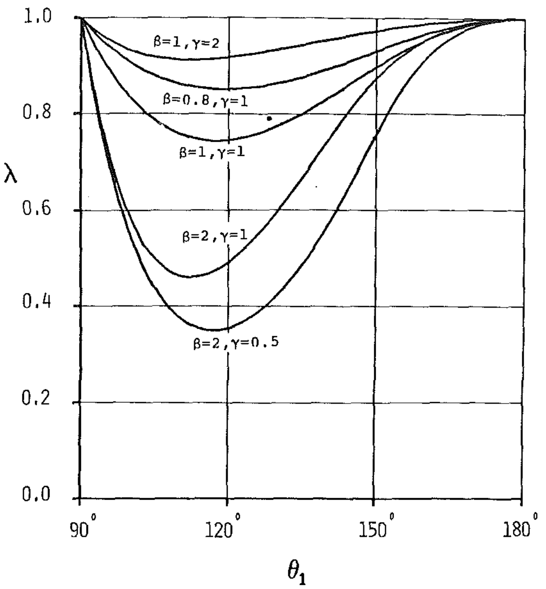


Fig. 3 Value of λ for a notch in "degenerate" materials with ν = 0.3

and the (2×2) matrix $\mathbf{K}^{(P)}(\theta)$ is given by

$$\mathbf{K}^{(P)}(\theta)=\begin{bmatrix}\zeta_1^{\lambda}\tilde{u}_{r,1}^{(P)} & \zeta_2^{\lambda}\tilde{u}_{r,2}^{(P)}\\ \zeta_1^{\lambda}\tilde{u}_{z,1}^{(P)} & \zeta_2^{\lambda}\tilde{u}_{z,2}^{(P)}\end{bmatrix} \qquad (20b)$$

for the rigid inclusion and

$$\mathbf{K}^{(P)}(\theta)=\begin{bmatrix}\zeta_1^{\lambda-1}(\tilde{\sigma}_{r,1}^{(P)}\cos\theta-\tilde{\sigma}_{rz,1}^{(P)}\sin\theta) & \zeta_2^{\lambda-1}(\tilde{\sigma}_{r,2}^{(P)}\cos\theta-\tilde{\sigma}_{rz,2}^{(P)}\sin\theta)\\ \zeta_1^{\lambda-1}(\tilde{\sigma}_{rz,1}^{(P)}\cos\theta-\tilde{\sigma}_{z,1}^{(P)}\sin\theta) & \zeta_2^{\lambda-1}(\tilde{\sigma}_{rz,2}^{(P)}\cos\theta-\tilde{\sigma}_{z,2}^{(P)}\sin\theta)\end{bmatrix} \qquad (20c)$$

for the notch. For (19) to yield a nontrivial solution for $\mathbf{q}$ we require

$$\|\mathbf{K}^{(P)}(\theta_1)\|=0 \qquad (21)$$

and this is the eigenequation for determining λ.

4.2 Problem (B): Two Conical Boundaries. The positive z axis is not included in the domain over which the solution is to be valid in this case. Therefore the complete solution given in (16) must be used. It contains 4 arbitrary constants A_1, B_1, A_2 and B_2. At each of the two boundaries $\theta=\theta_1$ and $\theta=\theta_2$ two conditions (either (18a) or (18b)) are imposed leading to the following system of four simultaneous linear algebraic equations for the four constants:

$$\mathbf{L}\mathbf{w}=\mathbf{0} \qquad (22)$$

where

$$\mathbf{w}=\begin{bmatrix}A_1\\ A_2\\ B_1\\ B_2\end{bmatrix} \qquad (23a)$$

$$\mathbf{L}=\begin{bmatrix}\mathbf{K}^{(P)}(\theta_1) & \mathbf{K}^{(Q)}(\theta_1)\\ \mathbf{K}^{(P)}(\theta_2) & \mathbf{K}^{(Q)}(\theta_2)\end{bmatrix} \qquad (23b)$$

In (23b), $\mathbf{K}^{(P)}(\theta_1)$, is given by (20$b$) or (20$c$) depending on whether $\theta=\theta_1$ is a fixed boundary or a traction free boundary. $\mathbf{K}^{(Q)}(\theta_1)$ is obtained from $\mathbf{K}^{(P)}(\theta_1)$ by replacing the Legendre functions P_λ, $P_{\lambda-1}$ by Q_λ, $Q_{\lambda-1}$. Similar definitions apply to $\mathbf{K}^{(P)}(\theta_2)$ and $\mathbf{K}^{(Q)}(\theta_2)$.

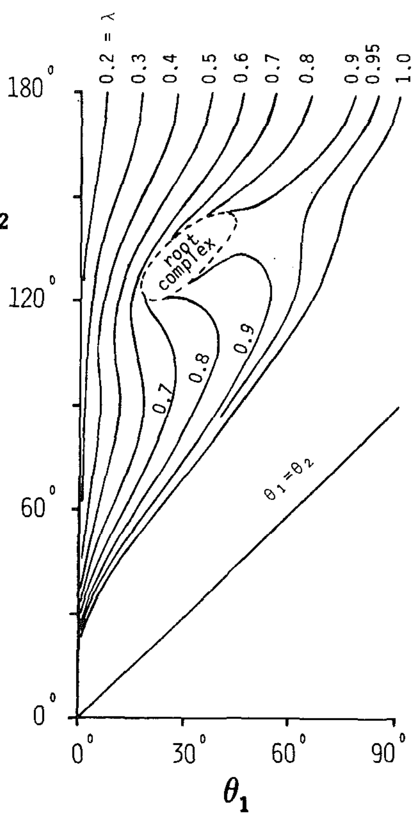


Fig. 4 Value of λ for two conical boundaries in isotropic materials with ν = 0.3 for the fixed-free case

For (22) to yield a nontrivial solution for **w** we require

$$\|\mathbf{L}\| = 0 \tag{24}$$

and this is the eigenequation for determining λ in this case.

5 The Degenerate Case $p_1 = p_2$

When the material properties are such that $p_1 = p_2$ the terms associated with the two independent solutions given in (11) become identical and cause the general solution to degenerate. This situation requires certain modifications of the solution which will be presented in this Section.

The degenerate case is discussed by Ting et al. (1985) and we follow the approach presented by them. From equation (5*b*) and the fact that p cannot be real it is seen that when $p_1 = p_2$, p must be purely imaginary. Also by equations (4*b*) and (5*a*) we have $m_1 = m_2 = \pm 1$. By setting $m = -1$ in the second equality of (4*a*), we obtain $c_{11}c_{33} - c_{13}^2 = 0$. This would imply that the matrix c_{ij} is not positive definite and hence $m \neq -1$. If we set $m = 1$ in the second equality of equation (4*a*) we get

$$(c_{13} + 2c_{44})^2 = c_{11}c_{33} \tag{25}$$

Thus when (25) is satisfied, $p_1 = p_2$ and we have a degenerate case. The five independent material constants are now reduced to four by the relation (25). Following Ting et al. (1985) we introduce the new material constants α, μ, γ and β by letting

$$\left.\begin{aligned} c_{11} &= (\alpha + 2\mu)\beta^2 \\ c_{33} &= (\alpha + 2\mu)/\beta^2 \\ c_{44} &= \mu \\ c_{13} &= \alpha \\ c_{11} - c_{12} &= 2\gamma\mu \end{aligned}\right\} \tag{26a}$$

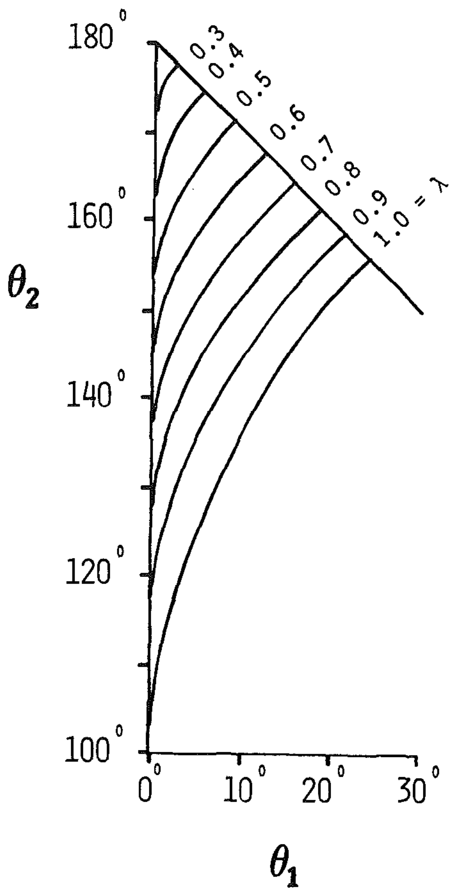


Fig. 5 Value of λ for two conical boundaries in isotropic materials with ν = 0.3 for the fixed-fixed case

and for later convenience we define ν by the relation

$$\alpha = 2\mu\nu/(1 - 2\nu) \tag{26b}$$

Equations (26*a*) satisfy (25) and when substituted into (5) they yield:

$$p = i\beta, \quad m = 1. \tag{27}$$

We also note, by differentiating the second equality of (4*b*) and substituting (26), that when $p_1 = p_2 = i\beta$,

$$\frac{dm}{dp} = \frac{4(1-\nu)}{\beta} i \tag{28}$$

In the particular case of $\gamma = \beta = 1$, the material becomes isotropic and α and μ become the Lame constants while ν is the Poisson's ratio.

In order to deduce the second independent solution we proceed as follows: Let Φ_1 and Φ_2 be the two solutions corresponding to p_1 and p_2. The general solution for the displacements is

$$u_r = A_1 \frac{\partial \Phi_1}{\partial r} + A_2 \frac{\partial \Phi_2}{\partial r} \tag{29}$$

$$u_z = A_1 m_1 \frac{\partial \Phi_1}{\partial z} + A_2 m_2 \frac{\partial \Phi_2}{\partial z} \tag{30}$$

In the degenerate case when $p_1 = p_2$, we have $\Phi_1 = \Phi_2$. To consider the limit as p_1 approaches p_2 let

$$\left.\begin{aligned} A_1 &= \frac{\tilde{A}}{2} + \frac{\tilde{B}}{p_1 - p_2} \\ A_2 &= \frac{\tilde{A}}{2} - \frac{\tilde{B}}{p_1 - p_2} \end{aligned}\right\} \tag{31}$$

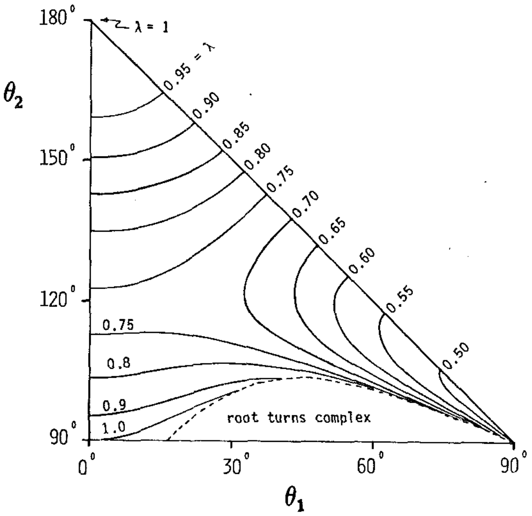


Fig. 6 Value of λ for two conical boundaries in isotropic materials with ν = 0.3 for the free-free case

Substitution of (31) into (29) and (30) and then taking the limit as $p_2 \to p_1$ yields, after omitting the subscript 1,

$$\left.\begin{aligned} u_r &= \tilde{A}\frac{\partial\Phi}{\partial r} + \tilde{B}\frac{\partial}{\partial p}\left(\frac{\partial\Phi}{\partial r}\right) \\ u_z &= \tilde{A}m\frac{\partial\Phi}{\partial r} + \tilde{B}\left[m\frac{\partial}{\partial p}\left(\frac{\partial\Phi}{\partial z}\right) + \frac{dm}{dp}\frac{\partial\Phi}{\partial z}\right] \end{aligned}\right\} \quad (32)$$

The terms associated with $\tilde{B}$ in (32) are easily identified as the second solution and by a similar approach the corresponding stress results may also be derived. Thus while the first solution is given by (1) and (2), the second solution is

$$\left.\begin{aligned} u_r &= \frac{\partial}{\partial p}\left(\frac{\partial\Phi}{\partial r}\right) \\ u_z &= m\frac{\partial}{\partial p}\left(\frac{\partial\Phi}{\partial z}\right) + \frac{dm}{dp}\frac{\partial\Phi}{\partial z} \\ \sigma_r &= c_{11}\frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial r^2}\right) + c_{12}\frac{\partial}{\partial p}\left(\frac{\partial\Phi}{r\partial r}\right) \\ &\quad + c_{13}\left[m\frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial z^2}\right) + \frac{dm}{dp}\frac{\partial^2\Phi}{\partial z^2}\right] \\ \sigma_\phi &= c_{12}\frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial r^2}\right) + c_{11}\frac{\partial}{\partial p}\left(\frac{\partial\Phi}{r\partial r}\right) \\ &\quad + c_{13}\left[m\frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial z^2}\right) + \frac{dm}{dp}\frac{\partial^2\Phi}{\partial z^2}\right] \\ \sigma_z &= c_{13}\frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial r^2}\right) + c_{13}\frac{\partial}{\partial p}\left(\frac{\partial\Phi}{r\partial r}\right) \\ &\quad + c_{33}\left[m\frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial z^2}\right) + \frac{dm}{dp}\frac{\partial^2\Phi}{\partial z^2}\right] \\ \sigma_{rz} &= c_{44}\left[(1+m)\frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial r\partial z}\right) + \frac{dm}{dp}\frac{\partial^2\Phi}{\partial r\partial z}\right] \end{aligned}\right\} \quad (33)$$

where Φ is given in (10). It should be noted that the material properties c_{ij}, p, m and dm/dp in (33) have to be assigned their values for the degenerate case given in (26), (27) and (28).

Equation (33) requires the differentiation of Φ with respect to p also. Consider first

$$\Phi = R^{\lambda+1}P_{\lambda+1}(c) \quad (34)$$

Use of the differentiation rule and the recurrence relations for $P_\lambda(c)$ and equations (7) and (9) leads to

$$\left.\begin{aligned} \frac{\partial\Phi}{\partial p} &= -iR^\lambda[(\lambda+1)s\cot\theta P_\lambda] \\ \frac{\partial}{\partial p}\left(\frac{\partial\Phi}{\partial r}\right) &= -iR^\lambda[\lambda(\lambda+1)\cot\theta(P_\lambda - cP_{\lambda-1})] \\ \frac{\partial}{\partial p}\left(\frac{\partial\Phi}{\partial z}\right) &= -iR^\lambda[(\lambda+1)(P_\lambda + c\lambda P_{\lambda-1})] \\ \frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial r^2}\right) &= -iR^{\lambda-1}\left[\lambda(\lambda+1)s\cot\theta\left\{\left(\lambda-\frac{1}{s^2}\right)P_\lambda\right.\right. \\ &\quad \left.\left. + \left(\frac{1}{s^2} - 2\lambda + 1\right)cP_{\lambda-1}\right\}\right] \\ \frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial z^2}\right) &= -iR^{\lambda-1}[\lambda(\lambda+1)(-ip)(-c\lambda P_\lambda \\ &\quad + \{2 + (2\lambda-1)c^2\}P_{\lambda-1})] \\ \frac{\partial}{\partial p}\left(\frac{\partial^2\Phi}{\partial r\partial z}\right) &= -iR^{\lambda+1}\left[\frac{\lambda(\lambda-1)}{s}\{(1+c^2\lambda)P_\lambda\right. \\ &\quad \left. + \{(\lambda-2) - c^2(2\lambda-1)\}cP_{\lambda-1}\}\right] \end{aligned}\right\} \quad (35)$$

The derivatives corresponding to $\Phi = R^{\lambda+1}Q_{\lambda+1}(c)$ are obtained, as before, by simply replacing P by Q in (35).

When the relations in (35) are substituted into (33), once again we would obtain expressions for displacements and stresses in the form of equation (15). Of course, the detailed expressions for $\tilde{u}_r^{(P)}$, . . . etc. are now different. In a similar fashion the solution corresponding to $\Phi = R^{\lambda+1}Q_{\lambda+1}(c)$ is also obtained and this completes the general solution for $p_1 = p_2$. The imposition of boundary conditions and the derivation of eigenequation can now be carried out as explained in Section 4.

6 Rigid Inclusion in Degenerate Materials and Isotropic Notch

Isotropic materials are only a particular case of the special class of "degenerate" materials discussed in the above Section. Even though there are four material constants for degenerate materials, by non-dimensionalizing with respect to μ one can make λ depend only on three parameters γ, β and ν. For isotropic materials, $\gamma = \beta = 1$ and λ depends on the Poisson's ratio only. Previous investigators have reported explicit eigenequations in the case of isotropic materials for problem (A), i.e. for the case of one conical boundary. In this section we will consider a rigid inclusion in degenerate materials and a notch in isotropic materials. For later use we note here that for $\beta = 1$ (which results in $p = i$) equations (9) and (17*b*) reduce to

$$c = \cos\theta,\ s = \sin\theta,\ \zeta = 1. \quad (36)$$

6.1 Rigid Inclusion in Degenerate Materials. Explicit expansion of (21) using (20*b*) results in the eigenequation

$$\tilde{u}_{r,1}^{(P)}\tilde{u}_{z,2}^{(P)} - \tilde{u}_{r,2}^{(P)}\tilde{u}_{z,1}^{(P)} = 0, \quad (37)$$

where it is tacitly understood that the expression is to be evaluated at $\theta = \theta_1$. Substitution of (14), (26) and (27) into (1) and the use of (15) gives

$$\left.\begin{aligned} \tilde{u}_{r,1}^{(P)} &= \frac{\lambda}{s}(cP_\lambda - P_{\lambda-1}) \\ \tilde{u}_{z,1}^{(P)} &= (1+\lambda)\beta P_\lambda \end{aligned}\right\} \qquad (38a)$$

To obtain the displacement components associated with the second set of solutions relevant expressions from (35), (27) and (28) should be substituted into u_r and u_z in (33). This results in

$$\left.\begin{aligned} \tilde{u}_{r,2}^{(P)} &= -i\lambda(1+\lambda)\cot\theta(P_\lambda - cP_{\lambda-1}) \\ \tilde{u}_{z,2}^{(P)} &= -i(1+\lambda)[\{1-4(1-\nu)\}P_\lambda + c\lambda P_{\lambda-1}] \end{aligned}\right\} \qquad (38b)$$

Substitution of (38) into (37) yields the eigenequation for the rigid inclusion in degenerate materials as

$$\frac{\lambda(1+\lambda)}{s}\Big\{[cP_\lambda - P_{\lambda-1}][\{1-4(1-\nu)\}P_\lambda + c\lambda P_{\lambda-1}] \\ -(1+\lambda)\beta s\cot\theta[P_\lambda - cP_{\lambda-1}]P_\lambda\Big\} = 0 \qquad (39)$$

It is worth noting that λ depends on β and ν but not on γ. The same situation occurs in the case of two conical boundaries if both boundaries are rigidly clamped. Now the eigenequation corresponding to isotropic materials can be easily obtained by letting $\beta = 1$ and using (36) in (39). After simplification it leads to:

$$\frac{\lambda(1+\lambda)}{s}\Big[c(1+\eta\lambda)P_\lambda^2 + c\lambda\eta P_{\lambda-1}^2 \\ -\{\eta c^2(2\lambda+1) + (1-\eta)\}P_\lambda P_{\lambda-1}\Big] = 0 \qquad (40a)$$

where

$$\eta = 1/[4(1-\nu)] \qquad (40b)$$

and c and s stand for $\cos\theta$ and $\sin\theta$ respectively as given by (36). This eigenequation is seen to be the same as that reported by Bazant and Keer (1974) when due attention is paid to the differences in notation.

6.2 Notch in Isotropic Materials. For this case the explicit expansion of (21) using (20c) gives

$$(\tilde{\sigma}_{r,1}^{(P)}\tilde{\sigma}_{rz,2}^{(P)} - \tilde{\sigma}_{rz,1}^{(P)}\tilde{\sigma}_{r,2}^{(P)})\cos^2\theta \\ +(\tilde{\sigma}_{z,1}^{(P)}\tilde{\sigma}_{r,2}^{(P)} - \tilde{\sigma}_{r,1}^{(P)}\tilde{\sigma}_{z,2}^{(P)})\cos\theta\sin\theta \\ +(\tilde{\sigma}_{rz,1}^{(P)}\tilde{\sigma}_{z,2}^{(P)} - \tilde{\sigma}_{z,1}^{(P)}\tilde{\sigma}_{rz,2}^{(P)})\sin^2\theta = 0 \qquad (41)$$

Here too it is tacitly understood that the expression should be evaluated at $\theta = \theta_1$. To reduce the amount of algebraic manipulations involved we straightaway consider the case of isotropic materials by letting $\beta = \gamma = 1$ which also makes $p = i$ and $m = 1$. Use of (26) with these values along with (14) in (2) results in the following expressions for the relevant stress components in the first set of solutions:

$$\left.\begin{aligned} \tilde{\sigma}_{r,1}^{(P)} &= -\frac{2\mu\lambda c}{s^2}P_\lambda + 2\mu\lambda\Big\{\frac{1}{s^2} - (1+\lambda)\Big\}P_{\lambda-1} \\ \tilde{\sigma}_{z,1}^{(P)} &= 2\mu\lambda(1+\lambda)P_{\lambda-1} \\ \tilde{\sigma}_{rz,1}^{(P)} &= \frac{2\mu\lambda(1+\lambda)}{s}(P_\lambda - cP_{\lambda-1}) \end{aligned}\right\} \qquad (42a)$$

Similarly the use of equations (26), (27), (28), (35), and (36) in (33) leads to the expressions for the relevant stress components in the second set of solutions

$$\left.\begin{aligned} \tilde{\sigma}_{r,2}^{(P)} &= -i\lambda(1+\lambda)\Big\{2\mu\Big(\lambda - \frac{1}{s^2}\Big)cP_\lambda \\ &\quad + \Big[2\mu\Big\{\frac{1}{s^2} - (2\lambda-1)\Big\}c^2 + \alpha\Big(2 - \frac{1}{\eta}\Big)\Big]P_{\lambda-1}\Big\} \\ \tilde{\sigma}_{z,2}^{(P)} &= -i\lambda(1+\lambda)\Big\{-2\mu\lambda cP_\lambda \\ &\quad + \Big[2\mu\Big\{(2\lambda-1)c^2 + \Big(2 - \frac{1}{\eta}\Big)\Big\} + \alpha\Big(2 - \frac{1}{\eta}\Big)\Big]P_{\lambda-1}\Big\} \\ \tilde{\sigma}_{rz,2}^{(P)} &= -i\lambda(1+\lambda)\Big\{2\mu\Big[(1+\lambda c^2) - \frac{1}{2\eta}\Big]\frac{1}{s}P_\lambda \\ &\quad + 2\mu\Big[(\lambda-2) - (2\lambda-1)c^2 + \frac{1}{2\eta}\Big]\frac{c}{s}P_{\lambda-1}\Big\} \end{aligned}\right\} \qquad (42b)$$

where η is defined in (40b) and c and s have been reduced to $\cos\theta$ and $\sin\theta$ respectively in accordance with (36). Substitution of (42) into (41) followed by much simplification yields the required eigenequation:

$$4\mu^2\lambda^2(1+\lambda)^2\frac{c}{s^3}\Big\{\Big[(c^2-1)\lambda^2 + (c^2-1)\lambda + \frac{c^2}{2\eta}\Big]P_\lambda^2 \\ + \Big[(c^2-1)\lambda^2 + (c^2-1)\lambda + \frac{1}{2\eta}\Big]P_{\lambda-1}^2 \\ -[2c^2(c^2-1)\lambda^2 + (c^2-1)(3c^2-1)\lambda \\ + \{c^4 + 2(1-2\nu)c^2 + 1\}]P_\lambda P_{\lambda-1}\Big\} = 0 \qquad (43)$$

where $c = \cos\theta$ and $s = \sin\theta$. Once again this is seen to be the same as that reported by Thompson and Little (1970) and Bazant and Keer (1974) except for the slight differences in notation. (There is a typographical error in equation (16) of Bazant and Keer (1974). In the coefficient of $P_\lambda(x)\ P_{\lambda-1}(x)$ the term $2(1-\nu)\lambda^2$ should read $2(1-2\nu)x^2$.)

7 Numerical Results

To illustrate the application of the solution procedure the eigenequation corresponding to several examples was solved for real values of λ that lie between 0 and 1. The search was carried out by evaluating the expression on the left hand side of the eigenequation at the two ends of a specified region on the real axis. If the two values differ in sign a root is present within the region. (The region should not be too large; otherwise this method will fail to detect roots when an even number of them are present in the same region). A promising region is bisected and each subdivision is tested separately to determine which one contains the root. This procedure is repeated until the region size reaches the tolerance level of accuracy. The computations presented here were carried out in IBM double precision and the values of λ were refined to ± 0.00005 so that they would be accurate to 4 decimal digits.

7.1 One Boundary. The material corresponding to degenerate case defined by (μ, β, γ, ν) was used. As we stated in Section 6, the material property matrix can be normalized with respect to μ causing λ to depend only on β, γ and ν. When assigning values to these parameters one must ensure that they result in a positive definite material stiffness matrix. This requires

$$\mu > 0,\ \gamma > 0,\ \nu < ½,\ \beta^2 > \frac{1-\nu}{2}\gamma \qquad (44)$$

The eigenequation for a rigid inclusion as given by (39) is independent of γ. Results obtained for varying β with $\nu = 0.3$ are shown in Fig. 2. Presented in Fig. 3 are similar results obtained for the case of a notch for different values of β and γ

with $\nu=0.3$. In the case of isotropic material (i.e. $\beta=\gamma=1$) the results agree with those presented by Bazant and Keer (1974). Even though both figures are for $\nu=0.3$ the trends they depict are typical of other values of ν also. It is seen that the anisotropy represented by the deviations of β and γ from their "isotropic value" of 1 does exert a substantial influence on the order of the stress singularities.

7.2 Two Boundaries. In this case there are many variable parameters: the material properties, two boundary angles and two boundary conditions. For an example we have picked the case of an isotropic material with $\nu=0.3$. The occurrence of stress singularities for three combinations of boundary conditions at $\theta = \theta_1$ and $\theta = \theta_2$ are considered; namely fixed-fixed, free-free, and fixed-free conditions.

The admissible ranges of values for the boundary angles are: 0 deg $< \theta_1 < \theta_2 <$ 180 deg. In the fixed-free case this entire region on the θ_1/θ_2 plane was scanned for possible stress singularities. Fig. 4 presents the results as contour lines of λ. Only the region that contained singularities on the admissible part of θ_1/θ_2 plane is shown there. This plot can be used to find λ corresponding to free-fixed case also because λ for free θ_1/fixed θ_2 is the same as that for fixed (180 deg $-\theta_2$)/free(180 deg $-\theta_1$).

In the fixed-fixed and free-free cases the λ corresponding to θ_1/θ_2 is the same as that for (180 deg $-\theta_2$)/(180 deg $-\theta_1$) which means that on the θ_1/θ_2 plane the λ values are symmetric about the line $\theta_1 + \theta_2 = 180$ deg. Therefore only the region $0<\theta_1<90$ deg, $\theta_1<\theta_2<$(180 deg $-\theta_1$) had to be considered. The parts of this region that contained possible stress singularities are shown in Figs. 5 and 6 with contours of λ.

The λ's shown in the figures are the smallest admissible real values between 0 and 1. In both the free-free and fixed-free cases we encountered the situation where there are two real λ's which turn into a pair of complex conjugate roots at some point in the θ_1/θ_2 plane. The real part of the complex roots would still be between 0 and 1 and would lead to a stress singularity. In Figs. 4 and 6 we have shown the boundary on which this transition takes place. It should be pointed out that on the boundary where real roots turn into complex roots, λ is a double root and the stress may have the $\rho^{\lambda-1}(\ln \rho)$ singularity in addition to the $\rho^{\lambda-1}$ singularity (Dempsey and Sinclair, 1979).

Acknowledgments

This work was supported by the Army Materials and Mechanics Research Center, Watertown, MA, under contract No. DAAG 46-83-K-0159 with the University of Illinois at Chicago. The support of AMMRC and the encouragement of its technical monitor Dr. S. C. Chou are gratefully acknowledged.

References

Bazant, Z. P., 1974, "Three-Dimensional Harmonic Functions Near Termination or Intersection of Gradient Singularity Lines: A General Numerical Method," *International Journal of Engineering Science*, Vol. 12, pp. 221-243.

Bazant, Z. P., and Estenssoro, L. F., 1977, "General Numerical Method for Three Dimensional Singularities in Cracked or Notched Elastic Solids," *Fracture 1977, Proceedings of the 4th International Conference on Fracture*, Taplin, D. M. R., ed., University of Waterloo, Ontario, Canada, Vol. 3, pp. 371-385.

Bazant, Z. P., and Estenssoro, L. F., 1979, "Surface Singularity and Crack Propagation," *International Journal of Solids and Structures*, Vol. 15, pp. 405-426. See also "Erratum," in Vol. 19, pp. 661 (1983).

Bazant, Z. P. and Keer, L. M., 1974, "Singularities of Elastic Stresses and of Harmonic Functions at Conical Notches and Inclusions," *International Journal of Solids and Structures*, Vol. 10, pp. 957-964.

Benthem, J. P., 1977, "State of Stress at the Vertex of a Quarter-Infinite Crack in a Half Space," *International Journal of Solids and Structures*, Vol. 13, pp. 479-492.

Benthem, J. P., 1980, "The Quarter-Infinite Crack in a Half Space; Alternative and Additional Solutions," *International Journal of Solids and Structures*, Vol. 16, pp. 119-130.

Dempsey, J. P., and Sinclair, G. B., 1979, "On the Stress Singularities in the Plane Elasticity of the Composite Wedge," *Journal of Elasticity*, Vol. 9, pp. 373-391.

Elliott, H. A., 1948, "Three-Dimensional Stress Distributions in Hexagonal Aeolotropic Crystals," *Proceedings of the Cambridge Philosophical Society*, Vol. 44, pp. 522-533.

Elliott, H. A., 1949, "Axial Symmetric Stress Distributions in Aeolotropic Hexagonal Crystals. The Problem of the Plane and Related Problems," *Proceedings of the Cambridge Philosophical Society*, Vol. 45, pp. 621-630.

Erdogan, F., 1983, "Stress Intensity Factors," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 992-1002.

Eshelby, J. D., Read, W. T., and Shockley, W., 1953, "Anisotropic Elasticity with Applications to Dislocation Theory," *Acta Metallurgy*, Vol. 1, pp. 251-259.

Green, A. E., and Zerna, W., 1975, *Theoretical Elasticity*, Oxford University Press, Oxford.

Kassir, M. K., and Sih, G. C., 1975, *Mechanics of Fracture 2—Three Dimensional Crack Problems*, Noordhoff International Publishing, Leyden.

Kawai, T., Fujitani, Y., and Kumagai, K., 1977, "Analysis of Singularity at the Root of the Surface Crack Problem," *Proceedings of an International Conference on Fracture Mechanics and Technology*, Sih, G. C. and Chow, C. L., ed., Vol. 11, pp. 1157-1163.

Lebedev, N. N., 1972, *Special Functions and Their Applications*, Revised edition, Silverman, Richard A., trans., Dover, New York.

Lin, K. Y., and Mar, J. W., 1976, "Finite Element Analysis of Stress Intensity Factors for Cracks at a Bi-Material Interface," *International Journal of Fracture*, Vol. 12, pp. 521-531.

Pian, T. H. H., Tong, P., and Luk, C. H., 1972, "Elastic Crack Analysis by a Finite Element Hybrid Method," *Proceedings of the 3rd Conference on Matrix Methods of Structural Analysis*, Wright Patterson Air Force Base.

Rhee, H. C., and Atluri, S. N., 1982, "Hybrid Stress Finite Element Analysis of Bending of a Plate with a Through Flaw," *International Journal for Numerical Methods in Engineering*, Vol. 18, pp. 259-271.

Sih, G. C., 1971, "A Review of the Three Dimensional Stress Problem for a Cracked Plate," *International Journal of Fracture Mechanics*, Vol. 9, pp. 39-61.

Somaratna, Nihal, and Ting, T. C. T., 1986, "Three-Dimensional Stress Singularities in Anisotropic Materials and Composites," *International Journal of Engineering Science*, to appear.

Thompson, Tommie R., and Little, Robert W., 1970, "End Effects in a Truncated Semi-Infinite Cone," *Quarterly Journal of Mechanics and Applied Mathematics*, Vol. 23, Pt. 2, pp. 185-195.

Ting, T. C. T., Jin, Yijian, and Chou, S. C., 1985, "Eigenfunctions at a Singular Point in Transversely Isotropic Materials Under Axisymmetric Deformations," ASME *Journal of Applied Mechanics*, Vol. 52, No. 3, pp. 565-570.

Tong, P., Pian, T. H. H., and Lasry, S. J., 1973, "A Hybrid-Element Approach to Crack Problems in Plane Elasticity," *International Journal for Numerical Methods in Engineering*, Vol. 7, pp. 297-308.

Nonlinear Analysis of Crazes

Y. Weitsman
Professor,
Mechanics and Materials Center,
Civil Engineering Department,
Texas A&M University,
College Station, Texas 77843

This paper concerns the nonlinear mechanical behavior of a single craze of finite length contained in an extended linear elastic medium. The craze is modeled as a distributed spring with a nonlinear force-displacement relation, which exhibits a hysteresis loop upon unloading. Stresses, displacements, and energy release rates are computed and compared against results for a linear craze. The case of a central crack within the craze is also considered.

Introduction

Crazes are thin, elongated defects that develop in many polymers. Although resembling cracks in their geometric configuration, namely length to thickness ratio of $0(10^3)$, crazes consist of interconnected voids that are transversed by thin fibrils which span the gap between opposite faces of the bulk polymer. These fibrils transmit substantial loads, thereby reducing the "craze opening displacement" to a level that is one order smaller than the opening displacement of a crack of an equal length. A schematic drawing, representing the craze region as an array of parallel fibrils between opposite faces of bulk polymeric material, is shown in Fig. 1.

Crazes grow by a process of fibrillation at their tips and thicken due to both stretching and drawing mechanisms. The lengths of the "process zones" at the tips are about two to three orders of magnitude smaller than the length of the craze, and typical craze opening profiles terminate with "cusps" within those tip regions.

The formation, growth, and shape of crazes have been studied by many investigators [11-12]. The subject is technically important because crazing appears to be a preferred energy absorbing mechanism in many circumstances, and its efficient use can find applications in both adhesives and composite materials. Nevertheless, the modeling of the mechanics of crazing appears to be deficient at the present time. Many investigators associated the craze region with the tip zone ahead of cracks (e.g., a plastic "Dugdale zone" [6, 8, 10, 12]) and attempted to correlate the behavior of crazes with considerations of crack stability. This approach suffers from the severe handicap of being unable to handle the case of crazes that develop in many polymers in the *absence* of cracks. In fact, it is well known that, in most circumstances, crazes are the precursors of cracks. Furthermore, even in the presence of cracks crazes can grow while crack lengths remain stationary.

This paper extends a recent work [13] where the craze was modeled as an elastic foundation with a linear force-displacement relation. The formulation, which was akin to fracture mechanics, led to a singular solution associated with stress intensity factors. These factors can be related to the familiar energy release "rates" that govern the stability and growth behavior of crazes. A subsequent work [14], which incorporated time-dependent response for both craze and bulk polymeric materials, demonstrated that the now time-dependent stress intensity factors could explain various observations regarding craze growth and arrest.

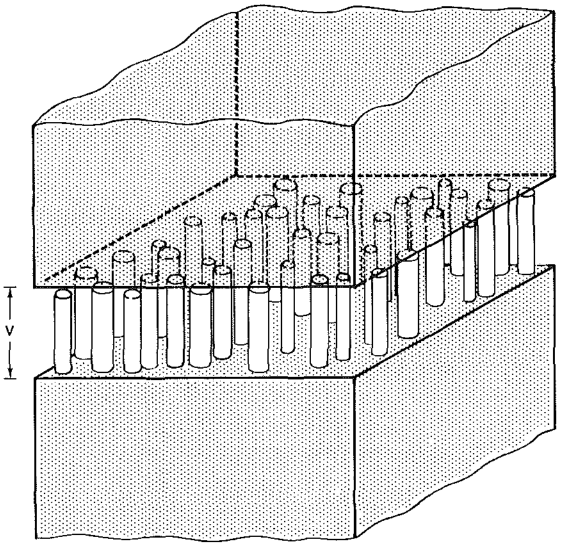


Fig. 1 A craze model showing parallel fibrils connecting opposite faces of the bulk polymer

Several experimental investigators [1], [5] noted nonlinear force-displacements relations for the craze fibrils. Furthermore, it was observed that significant hysteresis loops developed upon unloading. A schematic drawing of the craze response during the first three loading-unloading cycles is shown in Fig. 2. In the present work the nonlinear response of the fibrils, as sketched in Fig. 2, is represented by a nonlinear boundary condition, which is imposed on the linear elastic exterior region.

Formulation

Consider an infinite, isotropic, homogeneous, linear elastic region with G, ν denoting the shear modulus and Poisson's

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami Beach, Fla., November 17-22, 1985, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May, 1984, final revision, February, 1985. Paper No. 85-WA/APM-19.

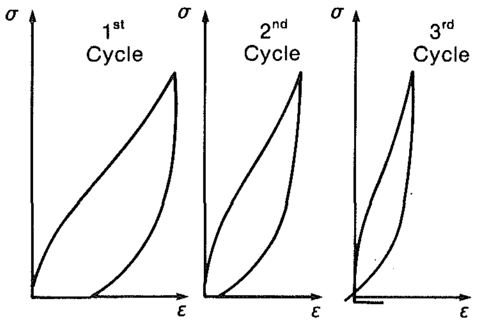


Fig. 2 A sketch of a typical stress-strain response of the craze material

ratio, respectively. Assume a state of plane strain and let σ_X, σ_Y, τ_{XY}, U, and V designate stresses and displacements relative to the Cartesian coordinates X, Y.

Let the region be subjected to $\sigma_Y = T$ as $|Y| \to \infty$ and consider a diminishingly thin craze along the X-axis at $-a < X < a$.

Modeling the craze as a nonlinear elastic foundation with a force-displacement relation given by $\sigma = F(V)$, it suffices to consider the following boundary value problem

$$\begin{array}{ll} \tau_{XY}(X,0)=0 & -\infty<X<\infty \\ V(X,0)=0 & |X|>a \\ \sigma_Y(X,0)=-T+F[V(X,0),X] & |X|<a \\ \sigma_Y(X,Y)\to 0 & |Y|\to\infty \end{array} \tag{1}$$

It can be shown [15] that if $(1)_1$, and $(1)_2$ hold then

$$\sigma_Y(X,0)=\frac{G}{1-\nu}\,\frac{1}{\pi}\int_{-a}^{a}\frac{\partial V(\xi,0)}{\partial \xi}\,\frac{d\xi}{\xi-X} \tag{2}$$

whereby $(1)_3$ gives

$$T=F(V,X)-\frac{G}{1-\nu}\,\frac{1}{\pi}\int_{-a}^{a}\frac{\partial V}{\partial \xi}\,\frac{d\xi}{\xi-X} \tag{3}$$

In view of the properties of Hilbert transform [16], equation (3) can be rewritten as

$$T=F(V,X)-\frac{G}{1-\nu}\,\frac{1}{\pi}\,\frac{d}{dX}\int_{-a}^{a}\frac{V}{\xi-X}\,d\xi \tag{4}$$

In (3) and (4) $V = V(X,0)$, while under the integrals $V = V(\xi,0)$.

The explicit dependence of the nonlinear response function F on the coordinate X along the craze allows the incorporation of the case of a central crack within the craze, as well as the case of unloading over a portion of the craze region.[1]

Assuming that $F(V,X)$ is continuous in V it follows that the solution of the singular integral equation (4) will be continuous over $[-a,a]$ if and only if $V \sim (a^2-X^2)^{1/2}$. More specifically, V takes the form

$$V=\frac{\Delta}{a}\sqrt{a^2-X^2}\,\chi(X/a) \tag{5}$$

In (5) χ is analytic over $[-a, a]$ and Δ is some length parameter.

Similarly to [13], it is advantageous to nondimensionalize (4) and (5). For this purpose let $v = V/\Delta$, $x = X/a$, $F(V,X) = Tf(v,x)$, and $s = \xi/a$. Then, in view of (5), we have $V = \Delta\sqrt{1-x^2}\,\chi(x)$ and (4) can be rewritten as follows

$$1=f(v,x)-\lambda\frac{1}{\pi}\,\frac{d}{dx}\int_{-1}^{1}\frac{v}{s-x}\,ds \tag{6}$$

In (6) $\lambda = G\Delta/[(1-\nu)aT]$

If we select, in particular, $\Delta = (1-\nu)aT/G = \delta$ then $\lambda = 1$. This value of δ corresponds to the opening displacement of a *crack* of length a in the same extended region. We shall adhere to this choice of $\Delta = \delta$, for which $\chi(1)$ provides the nondimensional stress intensity factor (the ratio between the stress intensity factor of the craze and that of a crack of equal length).

Following the same procedure as in [13] it is possible to convert the singular integral equation (6) to a Fredholm integral equation with a logarithmic singularity as follows

$$\chi(x)+\frac{1}{\pi}\int_{0}^{1}R_0(s^2,x^2)f[(1-s^2)^{1/2}\chi(s),s]ds=1 \tag{7}$$

In (7) $R_0(s^2,x^2) = (1-s^2)^{-1/2}R(s^2,x^2)$

where

$$\sqrt{\frac{1-x^2}{1-s^2}}\,R(s^2,x^2)=$$

$$\begin{cases} -\ln(x^2-s^2)+\ln(1-s^2)+2\ln\left[1+\sqrt{\dfrac{1-x^2}{1-s^2}}\right] & 0<s<x \quad (8a) \\ -\ln(s^2-x^2)+\ln(1-x^2)+2\ln\left[1+\sqrt{\dfrac{1-s^2}{1-x^2}}\right] & x<s<1 \quad (8b) \end{cases}$$

Assuming further that, during the first loading cycle, the leading term in $f(v,x)$ is linear in v we can express f as $f(v,x) = vg(v,x)$. In this case equation (7) gives

$$\chi(x)+\frac{1}{\pi}\int_{0}^{1}R(s^2,x^2)\chi(s)g[(1-s^2)^{1/2}\chi(s),s]ds=1 \tag{9}$$

Obviously, if a symmetric crack of length $2C(0<C\le a)$ occupies the central portion of the craze, then $g=0$ for $0<s\le c$, with $c = C/a$.

As noted previously in [13], the form (5) yields a displacement profile that terminates with a vertical tangent at $X = a$(i.e., at $x = 1$). In view of the exceedingly small length of the process zone at the craze tip it is possible to correct this deficiency by means of a Barenblatt-type analysis [17], as detailed in the Appendix.

The Unloading Case

Consider the first unloading path shown in Fig. 2. The force-displacement relation along this curve can be expressed by

$$\sigma=\sigma^{(m)}f_u(v/v^{(m)}) \tag{10}$$

where $\sigma^{(m)}$, $v^{(m)}$ are the maximal values attained by the stress and the displacement during the loading stage. The function f_u is monotonic in its argument and attains its largest value of unity at $v/v(m) = 1$.

Since only one unloading curve is available experimentally, we shall assume that equation (10) expresses the unloading path from *any* stress level.

In view of Fig. 2, there remains some positive residual strain upon complete, first cycle unloading. Consequently, we assume that f_u takes the form

$$f_u(v/v^{(m)})=-k_0+g_u(v/v^{(m)}) \tag{11}$$

where g_u can be expanded in powers of its argument.

Consider the case when the remote load T is reduced to $\alpha T (0 \le \alpha < 1)$. Since, in the present circumstance, the formulation of the boundary value problem results in an expression that is essentially similar to equation (4), we still have a nondimensional displacement $v(x)$ of the form $v(x) = (1-$

---

[1] It is possible to extend F to $F(V,X,a)$, accounting for a redistribution of the response with craze length a. However, this will not be considered here due to the absence of pertinent data.

$x^2)^{1/2}\chi(x)$. Therefore, the singular integral equation results in a form analogous to equation (9) as follows

$$\chi(x)+\frac{1}{\pi}\int_0^1 R(s^2,x^2)g(v^{(m)}(s))g_u[\chi(s)/\chi^{(m)}(s),s]ds$$
$$=\alpha+R(x) \qquad (12)$$

where

$$R(x)=\frac{k_0}{\pi}\int_0^1 R(s^2,x^2)\chi^{(m)}(s)g[(1-s^2)^{1/2}\chi^{(m)}(s),s]ds \qquad (13)$$

In equations (12) and (13) $v^{(m)}(s)$ and thereby $\chi^{(m)}(s)$ are presumed known from a prior solution to the loading case.

Work and Energy-Release Within the Craze

In analogy with the concepts of fracture mechanics, it is reasonable to assume that crazes will grow when the stress intensity factor, hence $\chi(1)$, exceeds a critical value. For this reason $\chi(1)$ is the essential quantity to be calculated in the present work. Nevertheless, it is of some interest to note that in contrast with cracks, where the entire energy release "rate" is associated with the work done at the tip zone due to crack advancement, crazes absorb energy also throughout the fibrillated region as their tips advance. For that reason the two portions of energy release "rate" are calculated below.

The work W in the portion $0\leq X<a$ of the craze is given by

$$W=\int_0^a\left(\int_0^{V(X)}\sigma\bullet 2d\tilde{V}\right)dX=2\int_0^a\left(\int_0^{V(X)}F(\tilde{V},X)d\tilde{V}\right)dX \qquad (14)$$

whereby, the "work relase rate" is

$$\frac{\partial W}{\partial a}=2\int_0^{V(a)}F(\tilde{V},X)d\tilde{V}$$
$$+2\int_0^a\left(\frac{\partial}{\partial a}\int_0^{V(X)}F(\tilde{V},X)d\tilde{V}\right)dX \qquad (15)$$

The first integral in (15) vanishes since, by hypothesis, $V(a)=0$. Employing the chain rule $\partial/\partial a=\partial V/\partial a\partial/\partial V$ we obtain

$$\frac{\partial W}{\partial a}=2\int_0^a\frac{\partial V}{\partial a}F(V,X)dX \qquad (16)$$

In view of (5), and with $\Delta=\delta$, we have

$$\frac{\partial V}{\partial a}=\frac{\delta}{a}\left(\frac{\chi}{\sqrt{1-x^2}}-x\sqrt{1-x^2}\,\frac{\partial\chi}{\partial x}\right)$$

Therefore, with $F(V,X)=Tf(v,x)=Tvg(v,x)$, and $v=\sqrt{1-x^2}\chi(x)$, (16) yields

$$\frac{1}{T\delta}\,\frac{\partial W}{\partial a}=2\int_0^1\left[\chi-x(1-x^2)\,\frac{\partial\chi}{\partial x}\right]\chi g(\sqrt{1-x^2}\chi,x)dx \qquad (17)$$

The nondimensional energy release term (17) should be compared with $\pi[\chi(1)]^2$, which provides the energy release "rate" due to the tip zone mechanism.

For a linear craze $f(v,x)$ reduces to $k(x)v$, whereby (17) reduces to

$$\frac{1}{T\delta}\,\frac{\partial W}{\partial a}=2\int_0^1 k(x)\chi\left[\chi-x(1-x^2)\,\frac{\partial\chi}{\partial x}\right]dx \qquad (18)$$

When a crack extends over the portion $0<x\leq c$, the integrals (17) and (18) should be evaluated between the limits c and 1 instead of 0 and 1.

Numerical Computations and Results

Data on stress-strain behavior of crazes [1, 5] can be converted to the forms $\sigma=f(v)$ and $\sigma=\sigma^{(m)}f_u(v/v^{(m)})$ as required by the present formulation. Unfortunately, those data are not supplemented by information on craze opening displacements and craze lengths under prescribed loads,

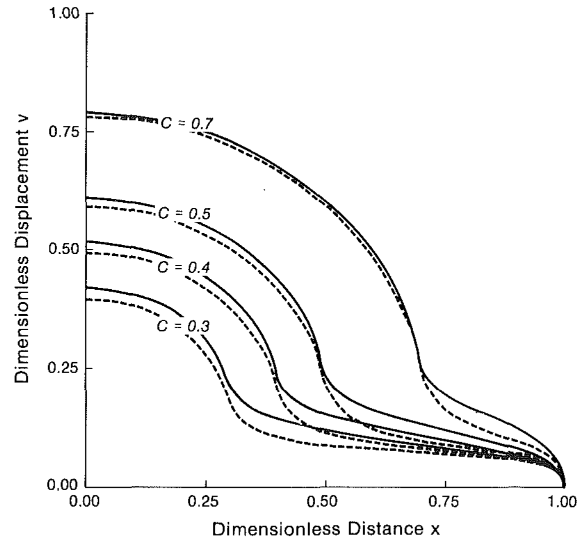


Fig. 3 Comparative opening displacement profiles $v(x)$ versus x for selected values of c. Nonlinear results (solid lines) and linear values (dashed lines).

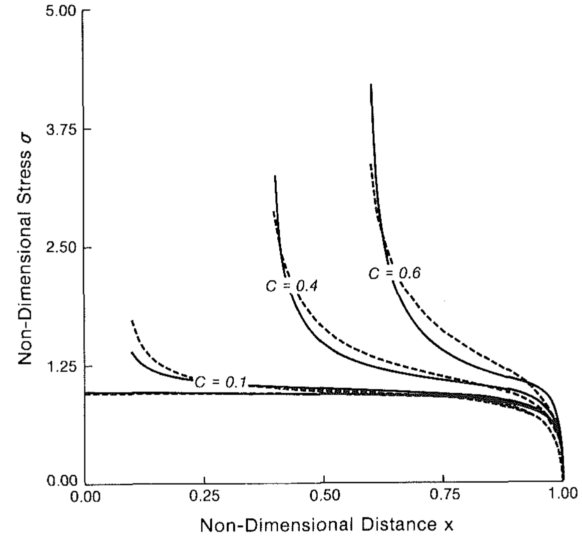


Fig. 4 Nondimensional stresses $\sigma(x)$ versus x within the craze, for various values of c. Nonlinear results (solid lines) and linear values (dashed lines).

which is also required in the present analysis. Consequently, the computations reported in this section are based on ad-hoc combinations of stress-strain data [1, 5] and craze profile information [2, 4].

The stress-strain curve for the initial load was approximated by the cubic equation $\sigma=80\epsilon-301.235\epsilon^2+438.958\epsilon^3$. To convert strain to displacement V it was assumed that $V=\epsilon\delta_o$, where the "premordial craze thickness" was $\delta_o=0.1\mu$m. Considering typical properties of polystyrene, let $G=1000$ MPA and $\nu=0.35$[2]. Finally, assume that under $T=7.2$ MPa the craze length is $a=50\mu$m. In accordance with the above stress-strain relation $\sigma=7.2$ MPa corresponds to $\epsilon=0.18$, thereby yielding a craze opening displacement $V(o)=0.018\mu$m as compared with a crack opening displacement $\delta=(1-\nu)Ta/G=0.234\mu$m.

Converting to nondimensional quantities we have, for the first loading stage, $f(v)=vg(v)$ where $g(v)$ was curve-fitted by

---

[2] These values are somewhat different from those employed in Ref. [13].

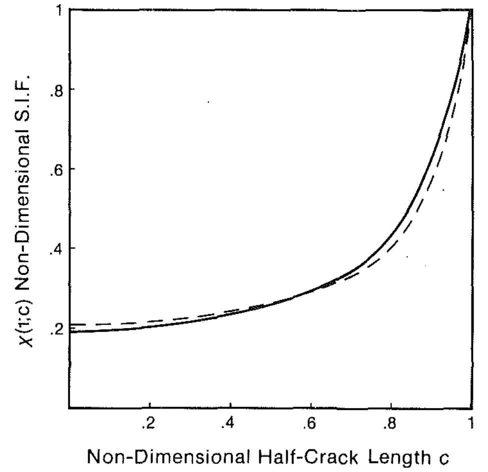


Fig. 5 Nondimensional stress intensity factors $\chi(1)$ versus c. Nonlinear results (solid line) and linear values (dashed line).

$$g(v)=b_o+b_1v+b_2v^2 \tag{19}$$

with $b_o=26$, $b_1=-229.1$, $b_2=781.2$.

The first unloading curve in [5] was approximated according to equation (11) with $k_o=0.47578$ and

$$g_u(v/v^{(m)})=B_o+B_1(v/v^{(m)})+B_2(v/v^{(m)})^2 \tag{20}$$

where $B_o=2.79494$, $B_1=-5.77698$, $B_2=4.45782$.

The solution to equation (9) with $g(v)$ given in (19) was obtained numerically. In view of the logarithmic singularity in $R(s^2,x^2)$ the segment [0,1] was divided into n equal subintervals, and $g(v)$ interpolated linearly over each of these subintervals. The resulting subintegrals were evaluated analytically, in analogy with the linear case detailed in the Appendix to [13]. In the present nonlinear case it is only necessary to redefine the quantity $L(x,s)$ so that instead of the previous relation $L(x,s)=k(s)\sqrt{(1-s^2)/(1-x^2)}$, we consider now the three expressions

$$\begin{aligned} L^{(0)}(x,s) &= b_o(s)\sqrt{(1-s^2)/(1-x^2)} \\ L^{(1)}(x,s) &= b_1(s)\sqrt{(1-s^2)}\sqrt{(1-s^2)/(1-x^2)} \\ L^{(2)}(x,s) &= b_2(s)(1-s^2)\sqrt{(1-s^2)/(1-x^2)} \end{aligned} \tag{21}$$

These $L^{(k)}(x,s)$ $(k=0,1,2)$ are now employed, in manner analogous to [13], to generate matrices $A_{ij}^{(k)}$ $(k=0,1,2)$ which yield the following nonlinear system of algebraic equations for $\chi_i=\chi(x_i)$

$$\sum_{j=0}^{n}(\delta_{ij}\chi_j+A_{ij}^{(0)}\chi_j+A_{ij}^{(1)}\chi_j^2+A_{ij}^{(2)}\chi_j^3)=1 \tag{22}$$

$$i=0,1,\ldots n$$

As noted in [13], $\lim_{x\to 1^-} R(x^2,s^2)=2$; thereby, if numerical integration employs the trapezoidal rule, we have $A_{nj}^{(0)}=(2/n\pi)b_{0j}$ for $1\le j\le n-1$ and $A_{nj}^{(0)}=(1/n\pi)b_{0j}$ for $j=0$ and $j=n$. Also, $A_{nj}^{(1)}=A_{nj}^{(2)}=0$ for all j.

The solution of the system (22) was obtained by means of the Newton-Raphson method, employing the linear solution as an initial guess. For an accuracy of $0(10^{-6})$ convergence was typically achieved after five iterations. The results were somewhat sensitive to the number of subdivisions n, and it was found that an accuracy of $0(10^{-3})$ was attained with $n=80$.

For comparison, the linear solution was also evaluated with the same values of a, G, and ν. To attain a value of $v(0)=0.06712$, which is the opening displacement obtained for the nonlinear case, it was necessary to compute the linear solution with $k=14.19$.

The case of a central crack of nondimensional length c within the craze was handled by letting $k(s)=0$ in the linear case, and $b_0(s)=b_1(s)=b_2(s)=0$ in the nonlinear case, for $0<s<c$.

Various results, with comparisons between linear and nonlinear values, are shown in Figs. 3, 4, and 5. Note that the differences between linear and nonlinear displacement profiles, as shown in Fig. 3, are surprisingly small. Therefore, measurements of those profiles cannot provide conclusive information regarding the response of the craze material.

It should be noted that in the absence of a criterion for craze "disintegration," namely, the condition for crack formation within the craze, the stress profiles shown in Fig. 4 should not be interpreted as "failure stresses." These stress profiles are merely computational results that demonstrate the differences between linear and nonlinear values.

Turning to the unloading case, the "fully loaded" quantities $\chi_j^{(m)}$ and thereby $\sigma_j^{(m)}$ were identified with the solution to equation (22). In view of equation (13), the values of $R_i=R(x_i)$ in (12) are given by

$$R_i=k_0\sum_{j=0}^{n}[A_{ij}^{(0)}\chi_j^{(m)}+A_{ij}^{(1)}(\chi_j^{(m)})^2 + A_{ij}^{(2)}(\chi_j^{(m)})^3] \qquad i=0,1,\ldots n \tag{23}$$

To solve for $\chi(x)$ in equation (12), we employ the numerical scheme of [13] with a modification that is analogous to that prescribed in equations (21), with the following additional replacements

$$\begin{aligned} b_0(s) &\rightarrow B_0(s)P^{(m)}(s) \\ b_1(s) &\rightarrow B_1(s)P^{(m)}(s)/\chi^{(m)}(s) \\ b_2(s) &\rightarrow B_2(s)P^{(m)}(s)/(\chi^{(m)}(s))^2 \end{aligned} \tag{24}$$

where $P^{(m)}(s)=b_0+b_1v^{(m)}(s)+b_2(v^{(m)}(s))^2$.

These replacements generate matrices $A_{ij}^{(ku)}$ $k=0,1,2$ and lead to the nonlinear system for the unloading solution χ

$$\sum_{j=0}^{n}(\delta_{ij}\chi_j+A_{ij}^{(0u)}\chi_j+A_{ij}^{(1u)}\chi_j^2+A_{ij}^{(2u)}\chi_j^3)=\alpha+R_i \tag{25}$$

$$i=0,1,\ldots n$$

The solution for the unloading case, as expressed in equation (25), was again obtained by the Newton-Raphson method. Covergence was attained after about six iterations.

Results for $\alpha=1$ (no unloading), $\alpha=0.5$ (partial unloading to half the maximal load) and $\alpha=0$ (complete unloading) are shown in Fig. 6.[3] For purposes of comparison, the corresponding displacement profiles for the linear case, where the displacements are proportional to the load level, are also exhibited. The substantial differences between the linear and nonlinear values of the displacement profiles indicate that unloading tests should provide valuable information on the response of the fibrillated, craze region.

Finally, the energy-release "rates" at the tip of the craze W_{tip} and the work release "rates" within the craze region W_a are plotted in Fig. 7 versus the nondimensional length c of a central crack. Note that as the crack length approaches unity $W_a\rightarrow 0$ and $W_{\text{tip}}\rightarrow\pi$, as expected.

Concluding Remarks

In this work it was shown that the analysis of the response of crazes can be disengaged from crack growth con-

---

[3] It should be remarked that upon complete unloading ($\alpha=0$) the present formulation results in slightly compressive stresses within the craze region ($\sigma_{\text{compression}}=2$-$5$ percent of the remote load T). If the craze fibrils cannot support compression, then the present analysis requires a suitable modification. Such a modification was not included herein.

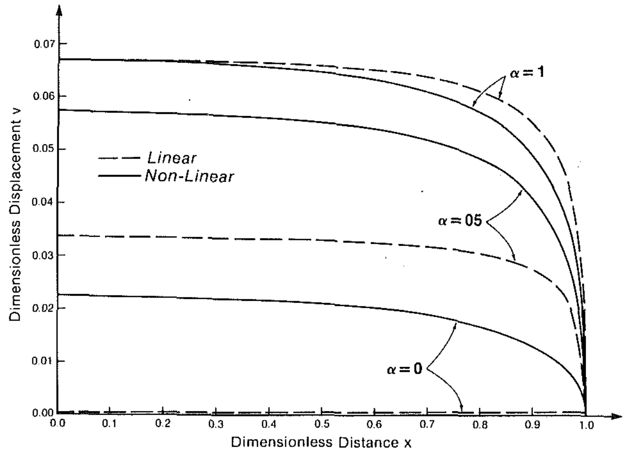


Fig. 6 Nondimensional opening displacement $v(x)$ versus the nondimensional distance x during first unloading cycle. Cases of full load ($\alpha = 1$), partial unloading ($\alpha = 0.5$) and complete unloading ($\alpha = 0$). Nonlinear results (solid lines) and linear values (dashed lines).

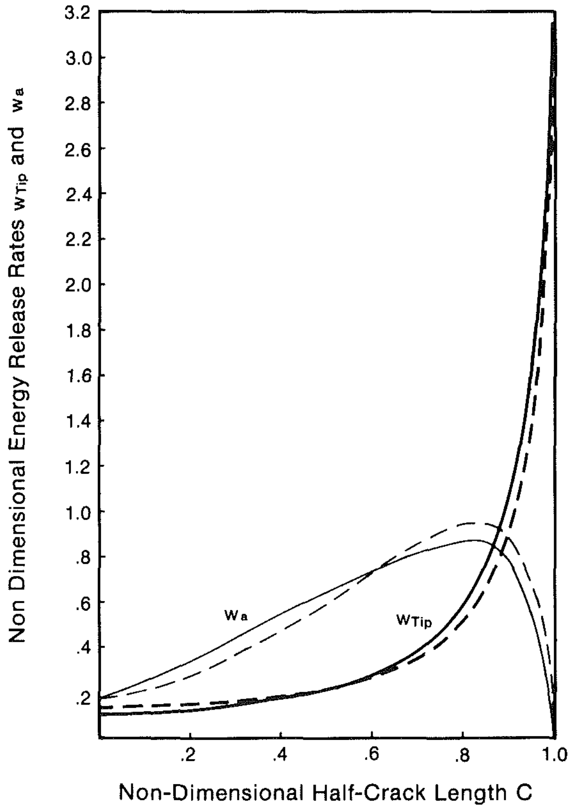


Fig. 7 Energy release rates at the craze tip, $w_{Tip} = \pi[\chi(1)]^2$, and work release rates within the craze region $w_a = (1/T\delta)\ \partial W/\partial a$, versus the nondimensional half-crack length c. Nonlinear results (solid lines) and linear values (dashed lines).

siderations. This feature of the present work provides a basis for predicting the behavior of crazes which develop and grow in the absence of cracks.

The present work indicates that the departure between linear and nonlinear behavior of the craze fibrils is likely to be much more noticeable under unloading than during loading.

It should be pointed out that the opening profiles for $\alpha = 1$ as shown in Fig. 6 do not agree with observed results. The discrepancy is most likely due to the fact that, in the absence of any single set of complete data, the computations presented in this paper were based on disjoint, partial sets of data. Further progress in this subject requires experimental programs which provide complete information on craze response. Such information should include craze lengths and opening profiles under various load levels within both the loading and unloading regimes.

Acknowledgments

The author is greatly indebted to Professor J. R. Walton for several helpful discussions and to Professor R. A. Schapery for a useful suggestion. The assistance of Mrs. J. Wells in obtaining the numerical results presented in this work is gratefully acknowledged.

The work was conducted under Contract N00014-82-K-0562 from the Office of Naval Research (ONR). The author wishes to express his thanks to Dr. Y. Rajapakse of the Mechanics Division, Engineering Sciences Directorate, ONR, for his support and encouragement.

References

1 Kambour, R. P., and Kopp, R. W., "Cyclic Stress-strain Behavior of the Dry Polycarbonate Craze," *J. Polymer Science*, A-2, Vol. 7, 1969, pp. 183-200.

2 Lauterwasser, B. D., and Kramer, E. J., "Microscopic Mechanisms of Craze Growth and Fracture," *Phil. Mag.*, A, Vol. 39, No. 4, 1979, pp. 469-495.

3 Kramer, E. J., "Crazing," *Proc. of the Int. Spring School of Plastic Deformation of Amorphous and Semicrystalline Materials*, D. Forelich, ed., Les Houches, France, Apr. 19-29, 1982.

4 Farrar, N. R., and Kramer, E. J., "Microstructure and Mechanics of Crazes in Oriented Polystyrene," *Polymer*, Vol. 22, 1981. pp. 691-698.

5 Hull, D., "The Microstructure and Properties of Crazes," in *Deformation and Fracture of High Polymers*, Kausch, Hassell and Jaffee, eds., Plenum Press, 1973, pp. 171-189.

6 Marshall, G. P., Culver, L. E., and Williams, J. G., "Craze Growth in Polymethylmethacrylate: A Fracture Mechanics Approach," *Proc. Royal Soc.*, Series A, Vol. 319, 1970, pp. 165-187.

7 Verheuplen-Heymens, N., and Bauwens, J. C., "Effects of Stress and Temperature on Dry Craze Growth Kinetics During Low-Stress Creep of Polycarbonate," *J. Materials Sci.*, Vol. 11, 1976, pp. 1-6.

8 Andrews, E. H., and Bevan, L., "Mechanics and Mechanism of Environmental Crazing in a Polymeric Glass," *Polymer*, Vol. 13, 1972, pp. 337-346.

9 Argon, A. S., and Salama, M. M., "Growth of Crazes in Glassy Polymers," *Phil. Mag*, Vol. 36, No. 5, 1977, pp. 1217-1234.

10 Doll, W., Seidelmann, U., and Konczol, L., "On the Validity of the Dugdale Model for Craze Zones at Crack Tips in PMMA," *J. Materials Sci.*, Vol. 15, 1980, pp. 2389-2394.

11 Doll, W., Konczol, L., and Schinker, M. G., "Zur Zeit-und Temperaturabhangigen Verstreckung von Polymermaterial vor Rissspitzen in PMMA bei Langzeitiger Statischer Belastung," *J. Colloid and Polymer Sci.*, Vol. 259, 1981, pp. 171-181.

12 Trassaert, P., and Schirrer, R., "The Disentangelment Time of the Craze Fibrils in Polymethylmethacrylate," *J. Materials Sci.*, forthcoming.

13 Walton, J. R., and Weitsman, Y., "Deformations and Stress Intensities Due to a Craze in an Extended Elastic Material," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, No. 1, 1984, pp. 84-92.

14 Weitsman, Y., "Viscoelastic Effects on Stresses and Stress Intensities in Crazes," in *Recent Developments in Applied Mathematics*," F. F. Ling and T. G. Tadjbakhsh, eds., Rensselaer Press, 1983, pp. 220-234.

15 Gladwell, G. M. L., *Contact Problems in the Classical Theory of Elasticity*, Sitjhoff & Noordhoff, 1980.

16 Sneddon, L. H., *The Use of Integral Transforms*, McGraw-Hill, New York, 1972, p. 235.

17 Barenblatt, G. I., "The Mathematical Theory of Equilibrium Cracks in Brittle Fracture," in *Advances in Applied Mechanics*, Vol. VII, Academic Press, New York, 1962, pp. 59-129.

APPENDIX

Tip Zone Correction

It is known that the fibrillated, craze region terminates with process zones at its tips. These process zones resemble a fingerlike array of polymeric matter interspersed with voids, as shown in Figs. 3, 4, and 5 of [13]. It has been suggested that plastic regions develop within the bulk polymer ahead of the process zones, in which case the tip regions would encompass both plastic and process zones.

Since the lengths of the tip regions are about two orders of magnitude smaller than the length of the craze, it is possible to consider their effects in eliminating the mathematical singularity associated with the present analysis in a manner that is analogous to Barenblatt's approach, and which resembles the linear case of [13].

Denote by $-\sigma_\epsilon(X)$ the compressive stresses acting within the tip regions $|a-\alpha|<|X|<|a|$ and introduce the non-dimensional quantities $-p_\epsilon(x)=-\sigma_\epsilon(X)/T$ and $\epsilon=\alpha/a$.

In analogy with equation (9) we now obtain

$$\chi_\epsilon(x)+\frac{1}{\pi}\int_0^1 R(s^2,x^2)\chi_\epsilon(s)g[(1-s^2)^{1/2}\chi_\epsilon(s),s]ds$$

$$=-\frac{1}{\pi}\int_{1-\epsilon}^1 R(s^2,x^2)(1-s^2)^{-1/2}p_\epsilon(s)ds \qquad \text{(A1)}$$

In $(A1)$ $\chi_\epsilon(x)=(1-x^2)^{-1/2}v_\epsilon(x)$, where $v_\epsilon(x)$ is the non-dimensional displacement due to $-p_\epsilon(x)$.

Taking the limit $x\to 1^-$ in (A1) we obtain, in view of the form of $R(s^2,x^2)$ given in equation (8)

$$\chi_\epsilon(1)+\frac{2}{\pi}\int_0^1 \chi_\epsilon(s)g[(1-s^2)^{1/2}\chi_\epsilon(s)]ds$$

$$=-\frac{2}{\pi}\int_{1-\epsilon}^1 \frac{p_\epsilon(s)}{\sqrt{1-s^2}}ds \qquad \text{(A2)}$$

Consider now the asymptotic limit of $\chi_\epsilon(x)$ when $x<1-\epsilon$ (fixed) while $\epsilon\to 0$, denoted by $\chi_0(x)$.

In this case the right-hand side of (A1) is given by

$$-\frac{1}{\pi}\lim_{\epsilon\to 0}\int_{1-\epsilon}^1 R(s^2,x^2)(1-s^2)^{-1/2}p_\epsilon(s)ds \qquad \text{(A3)}$$

Obviously, in the asymptotic limit under consideration, the value of x in (A3) is fixed, while $s\to 1^-$. In view equation (8b), we have that $\lim_{s\to 1^-} R(s^2,x^2)\,(1-s^2)^{-1/2}=0$ and therefore (A3) vanishes. Consequently, (A1) yields that $\chi_0(x)$ vanishes.

We thereby recover the well-known result that, for vanishingly small tip regions, the tip zone correction does not affect the solution outside those regions.

The tip zone correction is thus confined only near $x=1$, and it follows from (A2) that

$$\chi_\epsilon(1)=-\frac{2}{\pi}\int_{1-\epsilon}^1 \frac{p_\epsilon(s)}{\sqrt{1-s^2}}ds=-C_\epsilon \quad \text{(say)} \qquad \text{(A4)}$$

The combined solution, denoted by $\chi_c(x)$, due to the remote load T and tip region stress $-\sigma_\epsilon(X)$, is given by the singular solution for $0<x<1-\epsilon$ namely

$$\chi_c(x)=\chi(x) \text{ for } 0<x<1-\epsilon \qquad \text{(A5)}$$

while $\chi_c(1)$ is determined from

$$\chi_c(1)+\frac{2}{\pi}\int_0^1 \chi(s)g[(1-s^2)^{1/2}\chi(s)]ds=1-C_\epsilon \qquad \text{(A6)}$$

The amplitude and distribution of $p_\epsilon(x)$ in (A4) must be such so as to yield a value of C_ϵ which leads to $\chi_c(1)=0$, eliminating the mathematical singularity associated with the solution $\chi(x)$.

E. Tsuchida

T. Mura

J. Dundurs

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

The Elastic Field of an Elliptic Inclusion With a Slipping Interface

The paper analyzes the elastic fields caused by an elliptic inclusion which undergoes a uniform expansion. The interface between the inclusion and the matrix cannot sustain shear tractions and is free to slip. Papkovich-Neuber displacement potentials are used to solve the problem. In contrast to the perfectly bonded interface, the solution cannot be expressed in closed form and involves infinite series. The results are illustrated by numerical examples.

Introduction

The elliptic inclusion is defined in this paper as an elliptic subdomain, in an infinite two-dimensional space, which undergoes a uniform eigenstrain (stress-free strain, transformation strain) ϵ_{ij}^* and which also has elastic moduli different from those of the matrix.

Most of the inclusion problems solved by Eshelby (1957) and others assume continuity of displacements at the interface of the inclusion, i.e., perfect bonding. Recently Mura and Furuhashi (1984) found that the stress field of an ellipsoidal inclusion with a slipping interface vanishes when the eigenstrain has only shear components with respect to the principal axes of the ellipsoid.

In this paper the complementary part of the theory is developed. The eigenstrain in this paper is not of the shear type. For mathematical simplicity, however, we consider a two-dimensional inclusion. As expected, the solution cannot be expressed in a closed form, and is in the form of infinite series. Numerical results are given to illustrate the nature of the elastic fields.

Relations for Elliptic Coordinates

Elliptic coordinates are obtained by the coordinate transformation

$$x = c\cosh\alpha\cos\beta \qquad y = c\sinh\alpha\sin\beta \tag{1}$$

Eliminating first β and then α from these equations,

$$\frac{x^2}{c^2\cosh^2\alpha} + \frac{y^2}{c^2\sinh^2\alpha} = 1 \qquad \frac{x^2}{c^2\cos^2\beta} - \frac{y^2}{c^2\sin^2\beta} = 1 \tag{2}$$

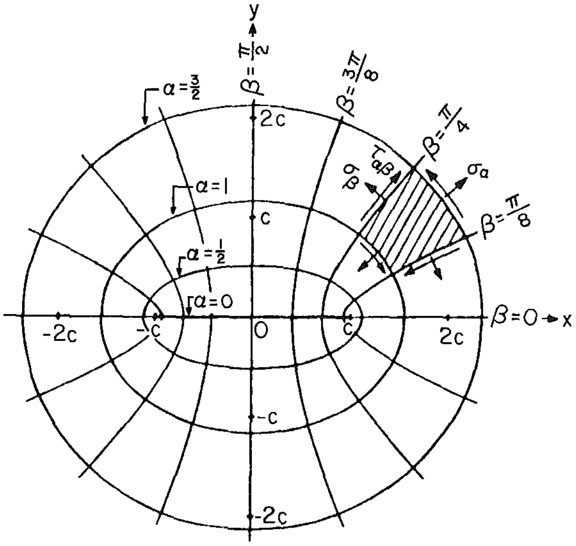


Fig. 1 Elliptic coordinate system used

and the coordinate lines are seen to be ellipses and hyperbolas (see Fig. 1).

The relations between the displacements and strains in the α and β directions are (Love, 1927, p. 54)

$$\epsilon_\alpha = h_1\frac{\partial U_\alpha}{\partial\alpha} + h_1h_2U_\beta\frac{\partial}{\partial\beta}\left(\frac{1}{h_1}\right)$$
$$\epsilon_\beta = h_2\frac{\partial U_\beta}{\partial\beta} + h_1h_2U_\alpha\frac{\partial}{\partial\alpha}\left(\frac{1}{h_2}\right) \tag{3}$$
$$\gamma_{\alpha\beta} = \frac{h_1}{h_2}\frac{\partial}{\partial\alpha}(h_2U_\beta) + \frac{h_2}{h_1}\frac{\partial}{\partial\beta}(h_1U_\alpha)$$

where

$$h_1^2 = \left(\frac{\partial\alpha}{\partial x}\right)^2 + \left(\frac{\partial\alpha}{\partial y}\right)^2, \quad h_2^2 = \left(\frac{\partial\beta}{\partial x}\right)^2 + \left(\frac{\partial\beta}{\partial y}\right)^2 \tag{4}$$

From (1), we have

$$h_1 = h_2 = h = \{2/c^2(\cosh 2\alpha - \cos 2\beta)\}^{1/2} \tag{5}$$

The relations between the strains and the stresses are given by Hooke's law

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami, Fla., November 17-21, 1985 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publicaiton of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 16, 1984; final revision, May 31, 1985. Paper No. 85-WA/APM-28.

$$\sigma_\alpha = \frac{G}{\kappa - 1} \{ (\kappa + 1)\epsilon_\alpha + (3 - \kappa)\epsilon_\beta \}$$

$$\sigma_\beta = \frac{G}{\kappa - 1} \{ (\kappa + 1)\epsilon_\beta + (3 - \kappa)\epsilon_\alpha \} \tag{6}$$

$$\tau_{\alpha\beta} = G\gamma_{\alpha\beta}$$

where G is the shear modulus, $\kappa = 3 - 4\nu$ for plain strain and $\kappa = (3 - \nu)/(1 + \nu)$ for the plane stress, and ν is Poisson's ratio.

The displacement components in the elliptic coordinates are given in terms of those in the Cartesian coordinates by

$$U_\alpha = h_1 \frac{\partial x}{\partial \alpha} U_x + h_1 \frac{\partial y}{\partial \alpha} U_y$$

$$U_\beta = h_2 \frac{\partial x}{\partial \beta} U_x + h_2 \frac{\partial y}{\partial \beta} U_y \tag{7}$$

or

$$U_\alpha = ch \sinh\alpha \cos\beta \, U_x + ch \cosh\alpha \sin\beta \, U_y$$

$$U_\beta = -ch \cosh\alpha \sin\beta \, U_x + ch \sinh\alpha \cos\beta \, U_y \tag{8}$$

where h is defined by (5).

The Papkovich–Neuber displacement functions ϕ_0, ϕ_1 and ϕ_2 are used to represent U_x and U_y as

$$2GU_x = \frac{\partial}{\partial x} (\phi_0 + x\phi_1 + y\phi_2) - (\kappa + 1)\phi_1$$

$$2GU_y = \frac{\partial}{\partial y} (\phi_0 + x\phi_1 + y\phi_2) - (\kappa + 1)\phi_2 \tag{9}$$

where

$$\nabla^2\phi_0 = 0, \quad \nabla^2\phi_1 = 0, \quad \nabla^2\phi_2 = 0 \tag{10}$$

Boundary Conditions

Consider an elliptic inclusion that undergoes *uniform* eigenstrains ϵ_x^* and ϵ_y^*. The present problem is to find the stress field when the elliptic inclusion is free to slip at the interface that cannot sustain shear tractions.

The boundary conditions at the interface $\alpha = \alpha_0$ are

$$U_\alpha = \bar{U}_\alpha + \bar{U}_\alpha^*, \quad \sigma_\alpha = \bar{\sigma}_\alpha, \quad \tau_{\alpha\beta} = 0, \quad \bar{\tau}_{\alpha\beta} = 0 \tag{11}$$

where the quantities with the bar pertain to the inclusion. The meaning of $\bar{U}_\alpha^*$ is as follows:

The elliptic inclusion with ϵ_x^*, ϵ_y^* will deform corresponding to the displacements $\bar{U}_x^* = \epsilon_x^* x$, $\bar{U}_y^* = \epsilon_y^* y$ if it has no constraint from the matrix. Because of the constraint, however, the displacement for the inclusion is the sum of $\bar{U}_i^*$ and the elastic displacement $\bar{U}_i$ caused by stresses and given by (9). When $\bar{U}_i^*$ is transformed into the elliptic coordinates by (8),

$$\bar{U}_\alpha^* = (c^2 h/4)\sinh 2\alpha[(1 + \cos 2\beta)\epsilon_x^* + (1 - \cos 2\beta)\epsilon_y^*],$$

$$\bar{U}_\beta^* = -(c^2 h/4)\sin 2\beta[(1 + \cosh 2\alpha)\epsilon_x^* + (1 - \cosh 2\alpha)\epsilon_y^*]. \tag{12}$$

It may be noted that, if the eigenstrain also has a ϵ_z^* component as in case of thermal expansion, the terms ϵ_x^* and ϵ_y^* in (12) must be replaced for plane strain with $\epsilon_x^* + \nu\epsilon_z^*$ and $\epsilon_y^* + \nu\epsilon_z^*$, respectively. The boundary conditions (11) state the continuity of the normal component of displacement and the continuity of normal traction and vanishing shear stress.

It has been found by Mura and Furuhashi (1984) that a shear eigenstrain ϵ_{xy}^* does not contribute to the stress field for the slipping inclusion. Therefore, ϵ_{xy}^* is not considered here. Consequently, the present problem has symmetry about the x and y axes.

The Papkovich–Neuber Functions

For an elliptic inclusion ($a_1 > a_2$), the Papkovich–Neuber displacement functions are chosen as

$$\phi_0 = c_0 \left[F_0\alpha + \sum_{n=1}^{\infty} A_n e^{-n\alpha} \cos n\beta \right]$$

$$\phi_1 = c_0 \sum_{n=1}^{\infty} B_n e^{-n\alpha} \cos n\beta \tag{13}$$

for the matrix ($\alpha > \alpha_0$) and

$$\phi_0 = c_0 \sum_{n=1}^{\infty} \bar{A}_n \cosh n\alpha \cos n\beta$$

$$\phi_1 = c_0 \sum_{n=1}^{\infty} \bar{B}_n \cosh n\alpha \cos n\beta \tag{14}$$

for the inclusion ($\alpha < \alpha_0$). Here $\phi_2 = 0$ and $c_0 = 2G\epsilon_x^*$ or $2G\epsilon_y^*$. In the expressions above, F_0, A_n, B_n, $\bar{A}_n$, $\bar{B}_n$ are unknown constants which are determined from the boundary conditions (11).

Next we evaluate the displacement and stress components corresponding to the terms in the series (13) and (14). The quantities defined in the inclusion carry bars.

For $\phi_0 = c_0 F_0 \alpha$,

$$2GU_\alpha/c_0 = F_0 h, \; U_\beta = 0,$$

$$\sigma_\alpha/c_0 = -(1/2)F_0 c^2 h^4 \sinh 2\alpha,$$

$$\sigma_\beta/c_0 = (1/2)F_0 c^2 h^4 \sinh 2\alpha, \tag{15}$$

$$\tau_{\alpha\beta}/c_0 = -(1/2)F_0 c^2 h^4 \sin 2\beta.$$

For $\phi_0 = c_0 \sum_{n=1}^{\infty} A_n e^{-n\alpha} \cos n\beta$,

$$\frac{2GU_\alpha}{c_0} = -h \sum_{n=1}^{\infty} A_n n e^{-n\alpha} \cos n\beta$$

$$\frac{2GU_\beta}{c_0} = -h \sum_{n=1}^{\infty} A_n n e^{-n\alpha} \sin n\beta$$

$$\frac{\sigma_\alpha}{c_0} = (1/4)c^2 h^4 \sum_{n=1}^{\infty} A_n n[-(n-1)e^{-n\alpha}\cos(n+2)\beta + \{ (n+1)e^{-(n-2)\alpha} + (n-1)e^{-(n+2)\alpha} \}\cos n\beta - (n+1)e^{-n\alpha}\cos(n-2)\beta]$$

$$\frac{\sigma_\beta}{c_0} = (1/4)c^2 h^4 \sum_{n=1}^{\infty} A_n n[(n-1)e^{-n\alpha}\cos(n+2)\beta - \{ (n+1)e^{-(n-2)\alpha} + (n-1)e^{-(n+2)\alpha} \}\cos n\beta + (n+1)e^{-n\alpha}\cos(n-2)\beta] \tag{16}$$

$$\frac{\tau_{\alpha\beta}}{c_0} = -(1/4)c^2 h^4 \sum_{n=1}^{\infty} A_n n[(n-1)e^{-n\alpha}\sin(n+2)\beta - \{ (n+1)e^{-(n-2)\alpha} + (n-1)e^{-(n+2)\alpha} \}\sin n\beta + (n+1)e^{-n\alpha}\sin(n-2)\beta]$$

For $\phi_1 = c_0 \sum_{n=1}^{\infty} B_n e^{-n\alpha} \cos n\beta$,

$$\frac{2GU_\alpha}{c_0} = -(1/4)ch \sum_{n=1}^{\infty} B_n[(n+\kappa)e^{-(n-1)\alpha} + (n-\kappa)e^{-(n+1)\alpha}] \times \{\cos(n+1)\beta + \cos(n-1)\beta\},$$

$$\frac{2GU_\beta}{c_0} = -(1/4)ch \sum_{n=1}^{\infty} B_n[e^{-(n-1)\alpha}$$

$$
\begin{aligned}
&+e^{-(n+1)\alpha}]\{(n-\kappa)\sin(n+1)\beta \\
&+(n+\kappa)\sin(n-1)\beta\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\sigma_\alpha}{c_0} &= (1/16)c^3h^4\sum_{n=1}^{\infty} B_n n[-\{(n+1)e^{-(n-1)\alpha} \\
&+(n-\kappa)e^{-(n+1)\alpha}\}\cos(n+3)\beta \\
&+\{(n+3)e^{-(n-3)\alpha}-2(-2+\kappa)e^{-(n-1)\alpha}+4e^{-(n+1)\alpha} \\
&+(n-\kappa)e^{-(n+3)\alpha}\}\cos(n+1)\beta+\{(n+\kappa)e^{-(n-3)\alpha} \\
&-4e^{-(n-1)\alpha}+2(-2+\kappa)e^{-(n+1)\alpha} \\
&+(n-3)e^{-(n+3)\alpha}\}\cos(n-1)\beta \\
&-\{(n+\kappa)e^{-(n-1)\alpha}+(n-1)e^{-(n+1)\alpha}\}\cos(n-3)\beta],
\end{aligned}
\tag{17}
$$

$$
\begin{aligned}
\frac{\sigma_\beta}{c_0} &= (1/16)c^3h^4\sum_{n=1}^{\infty} B_n n[\{(n-3)e^{-(n-1)\alpha} \\
&+(n-\kappa)e^{-(n+1)\alpha}\}\cos(n+3)\beta \\
&+\{-(n-1)e^{-(n-3)\alpha}+2(-2+\kappa)e^{-(n-1)\alpha}+4e^{-(n+1)\alpha} \\
&-(n-\kappa)e^{-(n+3)\alpha}\}\cos(n+1)\beta-\{(n+\kappa)e^{-(n-3)\alpha} \\
&+4e^{-(n-1)\alpha}+2(-2+\kappa)e^{-(n+1)\alpha} \\
&+(n+1)e^{-(n+3)\alpha}\}\cos(n-1)\beta \\
&+\{(n+\kappa)e^{-(n-1)\alpha}+(n+3)e^{-(n+1)\alpha}\}\cos(n-3)\beta],
\end{aligned}
$$

$$
\begin{aligned}
\frac{\tau_{\alpha\beta}}{c_0} &= (1/16)c^3h^4\sum_{n=1}^{\infty} B_n n[-\{(n-1)e^{-(n-1)\alpha} \\
&+(n-\kappa)e^{-(n+1)\alpha}\}\sin(n+3)\beta \\
&+\{(n+1)e^{-(n-3)\alpha}-2(-2+\kappa)e^{-(n-1)\alpha} \\
&+(n-\kappa)e^{-(n+3)\alpha}\}\sin(n+1)\beta+\{(n+\kappa)e^{-(n-3)\alpha} \\
&+2(-2+\kappa)e^{-(n+1)\alpha}+(n-1)e^{-(n+3)\alpha}\}\sin(n-1)\beta \\
&-\{(n+\kappa)e^{-(n-1)\alpha}+(n+1)e^{-(n+1)\alpha}\}\sin(n-3)\beta].
\end{aligned}
$$

For $\phi_0 = c_0\sum_{n=1}^{\infty}\bar{A}_n \cosh n\alpha\cos n\beta$,

$$
\frac{2\bar{G}\bar{U}_\alpha}{c_0} = h\sum_{n=1}^{\infty}\bar{A}_n n\sinh n\alpha\cos n\beta,
$$

$$
\frac{2\bar{G}\bar{U}_\beta}{c_0} = -h\sum_{n=1}^{\infty}\bar{A}_n n\cosh n\alpha\sin n\beta,
$$

$$
\begin{aligned}
\frac{\bar{\sigma}_\alpha}{c_0} &= -(1/4)c^2h^4\sum_{n=1}^{\infty}\bar{A}_n n[(n-1)\cosh n\alpha\cos(n+2)\beta \\
&-\{(n+1)\cosh(n-2)\alpha+(n-1)\cosh(n \\
&+2)\alpha\}\cos n\beta \\
&+(n+1)\cosh n\alpha\cos(n-2)\beta],
\end{aligned}
\tag{18}
$$

$$
\begin{aligned}
\frac{\bar{\sigma}_\beta}{c_0} &= (1/4)c^2h^4\sum_{n=1}^{\infty}\bar{A}_n n[(n-1)\cosh n\alpha\cos(n+2)\beta \\
&-\{(n+1)\cosh(n-2)\alpha+(n-1)\cosh(n \\
&+2)\alpha\}\cos n\beta \\
&+(n+1)\cosh n\alpha\cos(n-2)\beta],
\end{aligned}
$$

$$
\begin{aligned}
\frac{\bar{\tau}_{\alpha\beta}}{c_0} &= (1/4)c^2h^4\sum_{n=1}^{\infty}\bar{A}_n n[(n-1)\sinh n\alpha\sin(n+2)\beta \\
&-\{(n+1)\sinh(n-2)\alpha+(n-1)\sinh(n+2)\alpha\}\sin n\beta \\
&+(n+1)\sinh n\alpha\sin(n-2)\beta].
\end{aligned}
$$

For $\phi_1 = c_0\sum_{n=1}^{\infty}\bar{B}_n \cosh n\alpha\cos n\beta$,

$$
\begin{aligned}
\frac{2\bar{G}\bar{U}_\alpha}{c_0} &= (1/4)ch\sum_{n=1}^{\infty}\bar{B}_n\{(n+\bar{\kappa})\sinh(n-1)\alpha \\
&+(n-\bar{\kappa})\sinh(n+1)\alpha\} \\
&\times\{\cos(n+1)\beta+\cos(n-1)\beta\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{2\bar{G}\bar{U}_\beta}{c_0} &= -(1/4)ch\sum_{n=1}^{\infty}\bar{B}_n\{\cosh(n-1)\alpha+\cosh(n+1)\alpha\} \\
&\{(n-\bar{\kappa})\sin(n+1)\beta \\
&+(n+\bar{\kappa})\sin(n-1)\beta\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\bar{\sigma}_\alpha}{c_0} &= (1/16)c^3h^4\sum_{n=1}^{\infty}\bar{B}_n n[-\{(n+1)\cosh(n-1)\alpha \\
&+(n-\bar{\kappa})\cosh(n+1)\alpha\}\cos(n+3)\beta \\
&+\{(n+3)\cosh(n-3)\alpha-2(-2+\bar{\kappa})\cosh(n-1)\alpha \\
&+4\cosh(n+1)\alpha+(n-\bar{\kappa})\cosh(n+3)\alpha\}\cos(n+1)\beta \\
&+\{(n+\bar{\kappa})\cosh(n-3)\alpha-4\cosh(n-1)\alpha \\
&+2(-2+\bar{\kappa})\cosh(n+1)\alpha \\
&+(n-3)\cosh(n+3)\alpha\}\cos(n-1)\beta \\
&-\{(n+\bar{\kappa})\cosh(n-1)\alpha+(n-1)\cosh(n+1)\alpha\} \\
&\cosh(n-3)\beta],
\end{aligned}
$$

$$
\begin{aligned}
\frac{\bar{\sigma}_\beta}{c_0} &= (1/16)c^3h^4\sum_{n=1}^{\infty}\bar{B}_n n[\{(n-3)\cosh(n-1)\alpha \\
&+(n-\bar{\kappa})\cosh(n+1)\alpha\}\cos(n+3)\beta \\
&+\{-(n-1)\cosh(n-3)\alpha+2(-2+\bar{\kappa})\cosh(n-1)\alpha \\
&+4\cosh(n+1)\alpha-(n-\bar{\kappa})\cosh(n+3)\alpha\}\cos(n+1)\beta \\
&-\{(n+\bar{\kappa})\cosh(n-3)\alpha+4\cosh(n-1)\alpha \\
&+2(-2+\bar{\kappa})\cosh(n+1)\alpha \\
&+(n+1)\cosh(n+3)\alpha\}\cos(n-1)\beta \\
&+\{(n+\bar{\kappa})\cosh(n-1)\alpha \\
&+(n+3)\cosh(n+1)\alpha\}\cos(n-3)\beta],
\end{aligned}
\tag{19}
$$

$$
\begin{aligned}
\frac{\bar{\tau}_{\alpha\beta}}{c_0} &= (1/16)c^3h^4\sum_{n=1}^{\infty}\bar{B}_n n[\{(n-1)\sinh(n-1)\alpha \\
&+(n-\bar{\kappa})\sinh(n+1)\alpha\}\sin(n+3)\beta \\
&+\{-(n+1)\sinh(n-3)\alpha+2(-2+\bar{\kappa})\sinh(n-1)\alpha \\
&-(n-\bar{\kappa})\sinh(n+3)\alpha\}\sin(n+1)\beta \\
&-\{(n+\bar{\kappa})\sinh(n-3)\alpha+2(-2+\bar{\kappa})\sinh(n+1)\alpha \\
&+(n-1)\sinh(n+3)\alpha\}\sin(n-1)\beta \\
&+\{(n+\bar{\kappa})\sinh(n-1)\alpha+(n+1)\sinh(n+1)\alpha\} \\
&\sin(n-3)\beta].
\end{aligned}
$$

Using the results above and considering the symmetry of the problem about the y axis it is seen that A_n ($n=1, 3, 5, \ldots$), $\bar{A}_n$ ($n=1, 3, 5, \ldots$), B_n ($n=0, 2, 4, \ldots$), and $\bar{B}_n$ ($n=0, 2, 4, \ldots$) vanish.

Determination of the Unknown Constants

The remaining unknown constants A_i, $\bar{A}_i$, B_i, $\bar{B}_i$ and F_0 are determined from the boundary conditions (11).

The condition $U_\alpha = \bar{U}_\alpha + \bar{U}_\alpha^*$ at $\alpha = \alpha_0$ yields

$$
F_0 + \sum_{n=0}^{\infty}[(k_{A1}A_n + k_{B1}B_{n-1} + k_{B2}B_{n+1})
$$

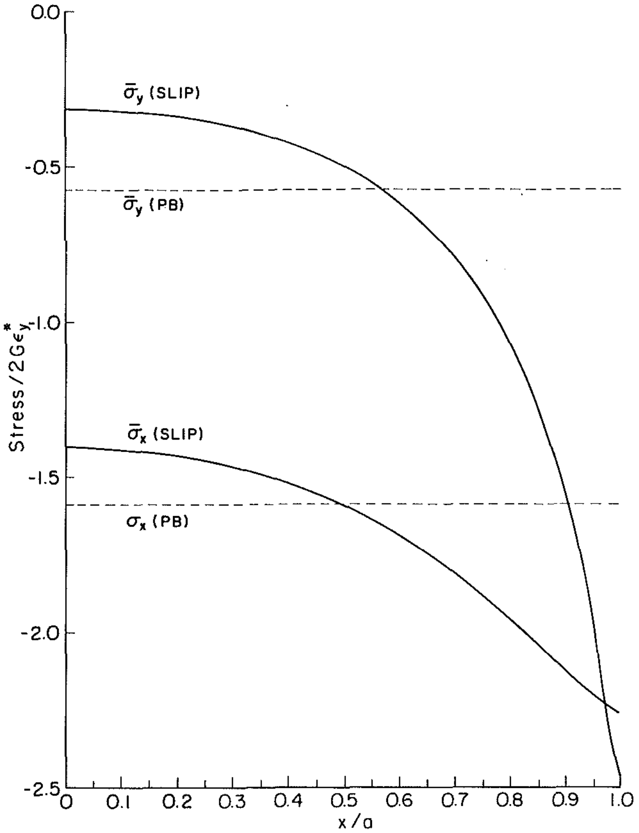


Fig. 2 The stress distributions along the *x* axis

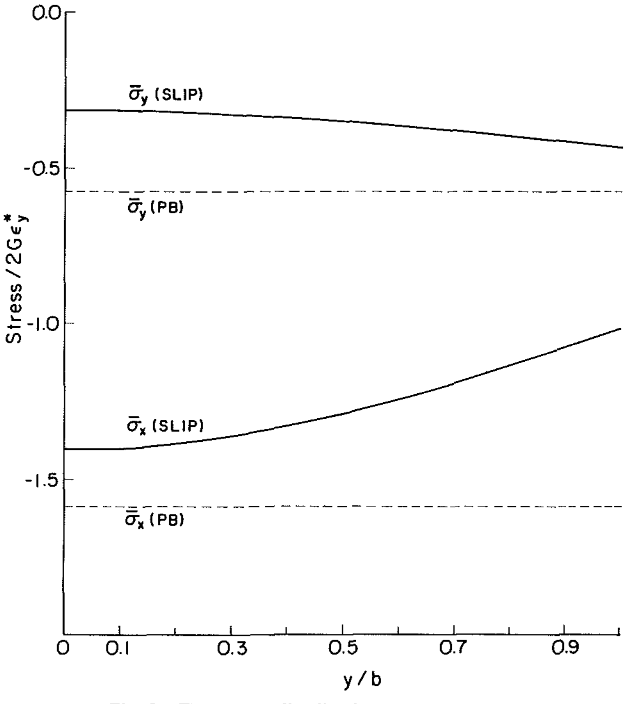


Fig. 3 The stress distribution along the *y* axis

$$-(1/\Gamma)\{\bar{k}_{A1}\bar{A}_n+\bar{k}_{B1}\bar{B}_{n-1}+\bar{k}_{B2}\bar{B}_{n+1}\}]\cos n\beta \tag{20}$$
$$=(G/2c_0)c^2\sinh 2\alpha_0\{(\epsilon_x^*+\epsilon_y^*)+(\epsilon_x^*-\epsilon_y^*)\cos 2\beta\}$$

The condition $\sigma_\alpha = \bar{\sigma}_\alpha$ at $\alpha = \alpha_0$ gives

$$S_{F0}F_0+\sum_{n=0}^{\infty}[(S_{A1}A_{n-2}+S_{A2}A_n+S_{A3}A_{n+2}$$

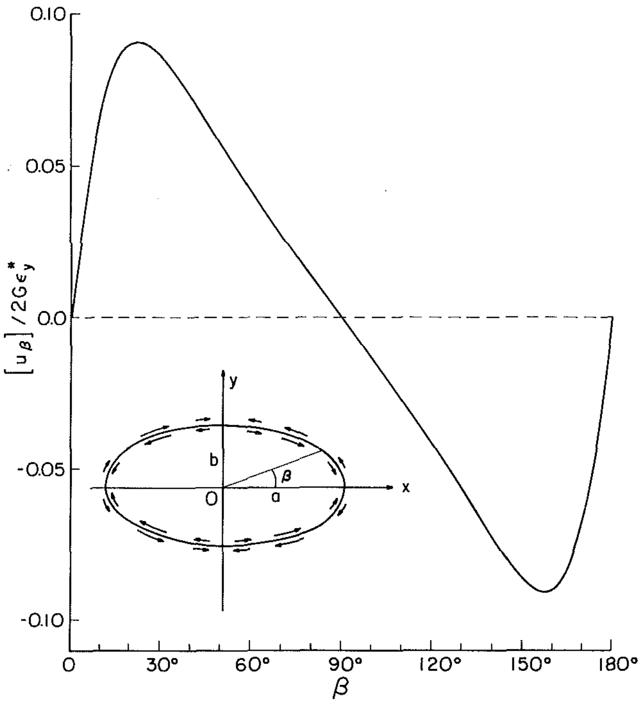


Fig. 4 Slip magnitude along the interface of inclusion

$$+S_{B1}B_{n-3}+S_{B2}B_{n-1}+S_{B3}B_{n+1}+S_{B4}B_{n+3})$$
$$-(\bar{S}_{A1}\bar{A}_{n-2}+\bar{S}_{A2}\bar{A}_n+\bar{S}_{A3}A_{n+2}+\bar{S}_{B1}\bar{B}_{n-3}$$
$$+\bar{S}_{B2}\bar{B}_{n-1}+\bar{S}_{B3}\bar{B}_{n+1}+\bar{S}_{B4}\bar{B}_{n+3})]\cos n\beta=0$$

The condition $\tau_{\alpha\beta} = 0$ at $\alpha = \alpha_0$ corresponds to

$$t_{F0}F_0\sin 2\beta+\sum_{n=1}^{\infty}(t_{A1}A_{n-2}+t_{A2}A_n+t_{A3}A_{n+2}$$
$$+t_{B1}B_{n-3}+t_{B2}B_{n-1}+t_{B3}B_{n+1}+t_{B4}B_{n+3})\sin n\beta=0. \tag{22}$$

The condition $\bar{\tau}_{\alpha\beta} = 0$ at $\alpha = \alpha_0$ yields

$$\sum_{n=1}^{\infty}(\bar{t}_{A1}\bar{A}_{n-2}+\bar{t}_{A2}\bar{A}_n+\bar{t}_{A3}\bar{A}_{n+2}+\bar{t}_{B1}\bar{B}_{n-3}$$
$$+\bar{t}_{B2}\bar{B}_{n-1}+\bar{t}_{B3}\bar{B}_{n+1}+\bar{t}_{B4}\bar{B}_{n+3})\sin n\beta=0 \tag{23}$$

where $\Gamma = \bar{G}/G$. The constants $k_{Ai}, k_{Bi}, \bar{k}_{Ai}, \bar{k}_{Bi}$, S_{F0}, S_{Ai}, $S_{Bi}, \bar{S}_{Ai}, \bar{S}_{Bi}, t_{F0}, t_{Ai}, t_{Bi}, \bar{t}_{Ai}$, and $\bar{t}_{Bi}$ are known functions of n, α_0, κ and $\bar{\kappa}$. Equating the coefficients of $\cos n\beta$ and $\sin n\beta$ ($n = 1, 2, 3, \ldots$) in (20) or (23), we obtain an infinite system of algebraic equations for F_0, A_n, B_n, $\bar{A}_n$, and $\bar{B}_n$.

Numerical Calculations

For a numerical example we calculate the stress field and slip magnitude when an elliptic inclusion of the same material as the matrix and with an axis ratio $b/a = 1/2$ undergoes the eigenstrain $\epsilon_x^* = 2\epsilon_y^*$. The stress components along the x and y axes inside the inclusion are shown in Figs. 2 and 3. The solid curves are for the slipping inclusion and the dotted curves for the perfectly bonded inclusion. The stress field for the slipping inclusion is not uniform as it is for the perfectly bonded inclusion. The stress field generally relaxes but may increase locally due to slip at the interface. The corresponding slip magnitude along the interface is shown in Fig. 4. The point $x = a, y = 0$ corresponds to $\beta = 0$ and the point $x = 0, y = b$ corresponds to $\beta = 90$ deg. The slip direction is indicated by the arrows in Fig. 4.

Since the problem contains a large number of independent parameters, an exhaustive representation of the results numerically is not feasible. It may be noted, however, that similarly to problems with specified surface tractions, (Dundurs, 1969) eigenstrain problems also show a reduced dependence on the elastic constants. Thus, instead of a general dependence on Γ, κ_1 and κ_2, the stress components in eigenstrain problems are of the form $\bar{G}/[\Gamma(\kappa_1+1)+\kappa_2+1]$ times a function which contains only two combinations of the elastic constants. The derivation of this result is given in the Appendix, and it was confirmed for some numerical runs which are not quoted here.

Acknowledgments

This research was supported by the U.S. Army Research grant No. DAAG29-81-K-0090.

References

Dundurs, J., 1969, Discussion of a paper by D. B. Bogy, ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 650–652.

Eshelby, J. D., 1957, "The Determination of the Elastic Fields of an Ellipsoidal Inclusion and Related Problems," *Proceedings of the Royal Society of London*, Ser. A, Vol. 241, pp. 376–396.

Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., Cambridge University Press.

Mura, T., and Furuhashi, R., 1984, "The Elastic Inclusion with a Sliding Interface," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 308–310.

Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff.

APPENDIX

To show the reduced dependence of the stress components on the elastic constants we use the complex-variable formulation of plane elasticity (Muskhelishvili, 1953) in which the displacement and stress components are related to the complex potentials $\phi(z)$ and $\psi(z)$ through

$$2G(U_x+iU_y)=\kappa\phi-z\overline{\phi'}-\overline{\psi} \tag{A1}$$

$$\sigma_{xx}+\sigma_{yy}=2(\phi'+\overline{\phi'}) \tag{A2}$$

$$\sigma_{yy}-\sigma_{xx}+2i\sigma_{xy}=2(\bar{z}\phi''+\psi') \tag{A3}$$

It is seen from (A2) and (A3) that the stress components depend on the elastic constants in the same way as the complex potentials. When the eigenstrains are *compatible*, the elastic constants can enter the complex potentials only through the boundary conditions at the interface. The requirements that the normal tractions be continuous and that the shear tractions vanish at the interface are homogeneous and contain no elastic constants. Hence, only the condition on the normal displacements need be examined in detail.

The normal component of displacements in the complex formulation is given by the expression

$$4GU_n=e^{-i\theta}(\kappa\phi-z\overline{\phi'}-\overline{\psi})+e^{i\theta}(\kappa\overline{\phi}-\bar{z}\phi'-\psi) \tag{A4}$$

where θ is the angle from the x-axis to the outer normal. The differential relation

$$d(\phi+z\overline{\phi'}+\overline{\psi})=i(t_x+it_y)\,ds \tag{A5}$$

is also needed in the derivation. In (A5), t_x and t_y denote traction components, and the arc-coordinate must be chosen so that the material is on the left when moving in the direction of increasing s along the boundary.

Whenever the eigenstrains in the inclusion are compatible and can thus be integrated for displacements, the required continuity of normal displacements at the interface can be written as the first equation in (A1) or

$$U_n^{(1)}+U_n^{(2)}=f(s) \tag{A6}$$

For (A6), the superscripts 1 and 2 refer respectively to the matrix and the inclusion, U_n is normal displacement caused by stresses and reckoned positive in the direction of the outer normal. Moreover, $f(s)$ is a known function corresponding to the normal displacement of the integrable eigenstrain in the inclusion. It is important to note that (A6) is the *nonhomogeneous* condition in the problem. Substituting (A4), (A6) becomes

$$\begin{aligned}\Gamma[&\exp(-i\theta_1)(\kappa_1\phi_1-z\overline{\phi_1'}-\overline{\psi_1})+\exp(i\theta_1)(\kappa_1\overline{\phi_1}\\ &-\bar{z}\phi_1'-\psi_1)]+\exp(-i\theta_2)(\kappa_2\phi_2-z\overline{\phi_2'}-\overline{\psi_2})\\ &+\exp(i\theta_2)(\kappa_2\overline{\phi_2}-\bar{z}\phi_2'-\psi_2)=4G_2f(s)\end{aligned} \tag{A7}$$

where the subscripts 1 and 2 are used to refer to the matrix and the inclusion, and $\Gamma=G_2/G_1$. At this point, the left side of (A7) contains the three elastic constants Γ, κ_1, and κ_2.

Newton's third law requires that at the interface

$$d(\phi_1+z\overline{\phi_1'}+\overline{\psi_1})=d(\phi_2+z\overline{\phi_2'}+\overline{\psi_2}) \tag{A8}$$

Due to the fact that the interface does not transmit a net force and, moreover, that an arbitrary complex constant γ can be added to $\phi(z)$ and $\kappa\bar{\gamma}$ to $\psi(z)$ without affecting the displacements, it is easy to show that (A8) yields

$$\phi_1+z\overline{\phi_1'}+\overline{\psi_1}=\phi_2+z\overline{\phi_2'}+\overline{\psi_2} \tag{A9}$$

at the interface. Noting that $\exp(i\theta_2)=-\exp(i\theta_1)$, it follows from (A9) that

$$\begin{aligned}&\exp(-i\theta_1)(\phi_1+z\overline{\phi_1'}+\overline{\psi_1})\\ &\qquad+\exp(-i\theta_2)(\phi_2+z\overline{\phi_2'}+\overline{\psi_2})=0\end{aligned} \tag{A10}$$

$$\begin{aligned}&\exp(i\theta_1)(\overline{\phi_1}+\bar{z}\phi_1'+\psi_1)\\ &\qquad+\exp(i\theta_2)(\overline{\phi_2}+\bar{z}\phi_2'+\psi_2)=0\end{aligned} \tag{A11}$$

Finally forming the linear combination (A7) + ½ $(\Gamma+1)$ [(A10) + (A11)] of the preceding equations and dividing by $\Gamma(\kappa_1+1)+\kappa_2+1$, the result is

$$\begin{aligned}&(2+\hat{\alpha}+\hat{\beta})[\exp(-i\theta_1)\phi_1+\exp(i\theta_1)\overline{\phi_1}]\\ &\quad-(\hat{\alpha}-\hat{\beta})[\exp(-i\theta_1)(z\overline{\phi_1'}+\overline{\psi_1})+\exp(i\theta_1)(\bar{z}\phi_1'\\ &\quad+\psi_1)]+(2-\hat{\alpha}-\hat{\beta})[\exp(-i\theta_2)\phi_2+\exp(i\theta_2)\overline{\phi_2}]\\ &+(\hat{\alpha}-\hat{\beta})[\exp(-i\theta_2)(z\overline{\phi_2'}+\overline{\psi_2})+\exp(i\theta_2)(\bar{z}\phi_2'+\psi_2)]\\ &\qquad=\frac{4G_2}{\Gamma(\kappa_1+1)+\kappa_2+1}f(s)\end{aligned} \tag{A12}$$

where

$$\hat{\alpha}=\frac{\Gamma(\kappa_1+1)-(\kappa_2+1)}{\Gamma(\kappa_1+1)+\kappa_2+1},\quad \hat{\beta}=\frac{\Gamma(\kappa_1-1)-(\kappa_2-1)}{\Gamma(\kappa_1+1)+\kappa_2+1} \tag{A13}$$

The constants $\hat{\alpha}$ and $\hat{\beta}$ (Dundurs, 1969) are measures for the mismatches in the uniaxial and voluminal compliances of the two elastic materials. It may be noted that $-1\le\hat{\alpha}\le+1$, $-½\le\hat{\beta}\le½$, that $\hat{\alpha}=\hat{\beta}=0$ for identical materials, and that $\hat{\alpha}$ and $\hat{\beta}$ simply assume opposite signs upon interchange of the two materials. The last fact is clear by reflected in the structure of the left side of (A12).

Since (A12) is the only nonhomogeneous condition among those that determine $\phi(z)$ and $\psi(z)$, it is clear that the complex potentials are of the form $\{G_2/[\Gamma(\kappa_1+1)+\kappa_2+1]\}\,F(z;\hat{\alpha},\hat{\beta})$. Consequently, on basis of (A2) and (A3), the stress components have the same dependence on the elastic constants. It may be noted that the shape of the inclusion does not enter this derivation.

Z. P. Duan[1]

J. W. Eischen
Graduate Student.
Assoc. Mem. ASME

G. Herrmann
Professor.
Fellow ASME

Division of Applied Mechanics,
Stanford University,
Stanford, CA 94305

Harmonic Wave Propagation in Nonhomogeneous Layered Composites

A new method for analyzing plane wave propagation in a periodically layered, elastic, nonhomogeneous composite body is proposed. The nonhomogeneity considered is a variation of the material properties within each composite layer. Results from probability theory are used to arrive at the two fundamental solutions of the governing second order ordinary differential equations. Floquet's wave theory is combined with a Wronskian formula to yield the dispersion relationship for this nonhomogeneous composite. Numerical results show that the presence of material nonhomogeneity affects the range of frequencies which can pass through the composite unattenuated.

1 Introduction

Problems of wave propagation in layered elastic composites have attracted a great deal of attention from researchers during recent years [1-6]. Several studies [7-9] have used Floquet's theory for one-dimensional wave propagation or Bloch's theory for three-dimensional wave propagation. These investigations have shown that when the wavelength of a harmonic wave is comparable to the characteristic length of the composite layers, successive reflection and refraction of the waves from the interfaces between layers leads to a significant dispersive effect. Such phenomena cannot be predicted by so-called "effective modulus" theories. For anti-plane or plane strain waves, the dispersion relationship can be interpreted geometrically as a surface in the wave number-frequency space. The important feature that was discovered is the presence of pass bands and stop bands, i.e., regions in the frequency spectrum where harmonic waves are either propagated freely or attenuated, respectively. The curves on the surface which define the boundary between the pass bands and stop bands divide the surface into so-called Brillouin zones.

The analyses made by Delph, Herrmann, and Kau [7-9] and by other researchers were performed with the assumption that the material properties within each layer of the composite were homogeneous. However, considering realistic manufacturing processes and/or naturally occurring variations it may not be reasonable to expect a uniform distribution of the elastic constants and mass density throughout each composite layer. It is the purpose of this paper to present a general method to analyze the sitution in which the cells in the periodically layered composite structure are nonhomogeneous, i.e., the elastic constants and mass density depend on the spatial coordinates within each layer.

This method is based on a procedure of representing the solution of the governing second order ordinary differential equations by means of a technique taken from probability theory. Combining Floquet's wave theory with properties of a special Wronskian formula, the dispersion relationship for wave propagation in certain nonhomogeneous composites is derived. Numerical calculations pertaining to the dispersion relationship for nonhomogeneous composites have shown that the presence of a material nonhomogeneity within each layer of the composite alters the width of the stop band and affects the dissipative characteristics of the medium.

2 Derivation of the Dispersion Relationship

The system under consideration consists of an infinite sequence of two alternating layers, each of which are taken to be nonhomogeneous and elastic. Perfect bonding is assumed between the adjoining layers. A unit cell is defined as the union of any two adjacent layers. As shown in Fig. 1, the two lamellae of the *N-th* unit cell have variable Lamé moduli $\{\lambda_m(x),\ \mu_m(x)\}$, $\{\lambda_f(x),\ \mu_f(x)\}$, variable mass densities $\{\rho_m(x),\ \rho_f(x)\}$, and thicknesses $\{2h_m,\ 2h_f\}$, where the subscripts m and f refer to "matrix" and "fiber" layers, respectively.

Let u, v, and w be the three Cartesian components of the displacement vector in the x, y and z directions, respectively. The layers lie in the y-z plane. Consideration will be given only to waves propagating in a direction normal to the layers. For a one-dimensional longitudinal strain wave propagtaing in the x-direction, only the u component of displacement is nonzero. Therefore, we take

$$u = u(x,t) \quad v = w = 0 \tag{2.1}$$

where the function $u(x, t)$ satisfies the equation of motion

---

[1] Visiting scholar, on leave from Institute of Mechanics, Chinese Academy of Sciences, Beijing, China.

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

Discussion of this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 27, 1984; final revision, August 13, 1985.

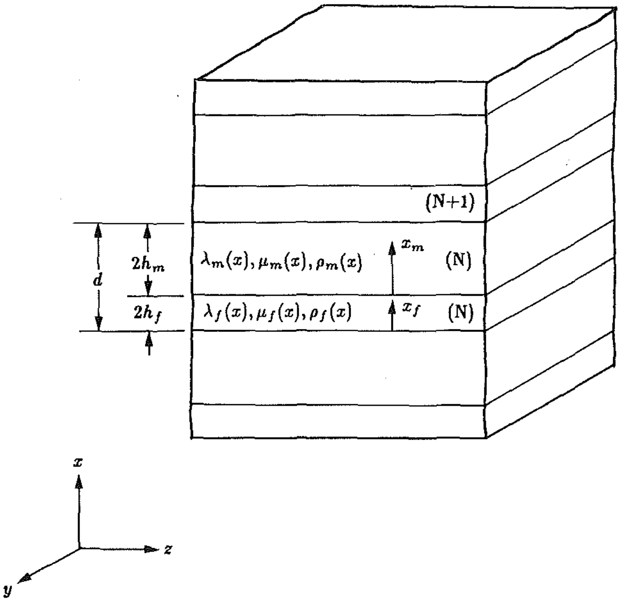


Fig. 1 Geometry of nonhomogeneous layered composite

$$\rho(x)\,\frac{\partial^2 u}{\partial t^2}=\frac{\partial}{\partial x}\left(D(x)\,\frac{\partial u}{\partial x}\right) \tag{2.2}$$

where $D(x) = \lambda(x) + 2\mu(x)$. The material properties $D(x)$ and $\rho(x)$ satisfy periodicity conditions

$$D(x+d)=D(x) \qquad \rho(x+d)=\rho(x) \qquad (2.3)_{1\text{-}2}$$

where $d = 2(h_m + h_f)$ represents the thickness of one unit cell, while

$$D(x)=\begin{cases} D_m(x), & \text{for } matrix \\ D_f(x), & \text{for } fiber \end{cases} \tag{2.4}$$

and

$$\rho(x)=\begin{cases} \rho_m(x), & \text{for } matrix \\ \rho_f(x), & \text{for } fiber \end{cases} \tag{2.5}$$

The global coordinate x is replaced with two local coordinates, x_m and x_f, in the *N-th* unit cell, as shown in Fig. 1. Thus, (2.2) holds in each layer

$$\rho_m(x_m)\,\frac{\partial^2 u_m}{\partial t^2}=\frac{\partial}{\partial x_m}\left(D_m(x_m)\,\frac{\partial u_m}{\partial x_m}\right) \quad (0\le x_m\le 2h_m) \qquad (2.6)_1$$

$$\rho_f(x_f)\,\frac{\partial^2 u_f}{\partial t^2}=\frac{\partial}{\partial x_f}\left(D_f(x_f)\,\frac{\partial u_f}{\partial x_f}\right) \quad (0\le x_f\le 2h_f) \qquad (2.6)_2$$

The solutions of $(2.6)_{1\text{-}2}$ are taken in the form of harmonic plane waves as

$$u_m=\left[\frac{D_m(0)}{D_m(x_m)}\right]^{1/2} U_m(x_m)e^{-i\omega t} \qquad (2.7)_1$$

$$u_f=\left[\frac{D_f(0)}{D_f(x_f)}\right]^{1/2} U_f(x_f)e^{-i\omega t} \qquad (2.7)_2$$

where ω is the circular frequency measured in radians per second, $U_m(x_m)$ and $U_f(x_f)$ are two unknown functions representing the amplitude of vibration. The longitudinal wave speeds will be denoted by $c_m = \sqrt{D_m/\rho_m}$ and $c_f = \sqrt{D_f/\rho_f}$.

We next introduce nondimensional parameters and nondimensional dependent and independent variables according to

$$\bar{x}_f=\frac{x_f}{2h_f} \quad \bar{x}_m=\frac{x_m}{2h_m} \quad \bar{U}_f=\frac{U_f}{2h_f} \quad \bar{U}_m=\frac{U_m}{2h_m}$$

$$\Omega=\frac{2\omega h_f}{\pi c_{fo}} \quad \tau=\frac{\pi c_{fo}}{2h_f}t \quad \bar{\sigma}_f=\frac{\sigma_f}{D_f(0)} \quad \bar{\sigma}_m=\frac{\sigma_m}{D_m(0)} \tag{2.8}$$

$$\bar{c}_f=\frac{c_f}{c_{fo}} \quad \bar{c}_m=\frac{c_m}{c_{mo}} \quad \bar{D}_f(\bar{x}_f)=\frac{D_f(x_f)}{D_f(0)}$$

$$\bar{D}_m(\bar{x}_m)=\frac{D_m(x_m)}{D_m(0)}$$

where $c_{mo} = \sqrt{D_m(0)/\rho_m(0)}$ and $c_{fo} = \sqrt{D_f(0)/\rho_f(0)}$ represent the longitudinal wave speeds at the interface between the matrix and fiber, respectively. The stresses in the matrix and fiber layers are indicated by σ_m and σ_f, respectively. Using these nondimensional quantities and substituting $(2.7)_{1\text{-}2}$ into $(2.6)_{1\text{-}2}$, the equations of motion are reduced to a system of second order Sturm-Liouville ordinary differential equations

$$\frac{d^2\bar{U}_m}{d\bar{x}_m^2}+V_m(\bar{x}_m,\Omega)\bar{U}_m=0 \qquad (2.9)_1$$

$$\frac{d^2\bar{U}_f}{d\bar{x}_f^2}+V_f(\bar{x}_f,\Omega)\bar{U}_f=0 \qquad (2.9)_2$$

with variable coefficients given by

$$V_m(\bar{x}_m,\Omega)=\frac{\Omega^2\pi^{*2}}{\bar{c}_m^2(\bar{x}_m)}-\frac{1}{2}\frac{d}{d\bar{x}_m}\left(\frac{1}{\bar{D}_m}\frac{d\bar{D}_m}{d\bar{x}_m}\right) -\frac{1}{4}\left(\frac{1}{\bar{D}_m}\frac{d\bar{D}_m}{d\bar{x}_m}\right)^2 \qquad (2.10)_1$$

$$V_f(\bar{x}_f,\Omega)=\frac{\Omega^2\pi^2}{\bar{c}_f^2(\bar{x}_f)}-\frac{1}{2}\frac{d}{d\bar{x}_f}\left(\frac{1}{\bar{D}_f}\frac{d\bar{D}_f}{d\bar{x}_f}\right) -\frac{1}{4}\left(\frac{1}{\bar{D}_f}\frac{d\bar{D}_f}{d\bar{x}_f}\right)^2 \qquad (2.10)_2$$

where $\pi^* = \pi\,\dfrac{h_m c_{fo}}{h_f c_{mo}}$.

For convenience, we now revert to mathematical notations introduced initially to indicate the corresponding nondimensional quantities, thereby dropping the barred notation. It is hoped that this will not confuse the readers. Furthermore, since both independent variables $\bar{x}_m$ and $\bar{x}_f$ vary between (0, 1), there is no need to distinguish between them. In the following derivation, we will let x stand for both $\bar{x}_m$ and $\bar{x}_f$.

Let U_{mi} (x, Ω) and $U_{fi}(x, \Omega)$ $(i = 1, 2)$ be the two fundamental solutions of $(2.9)_{1\text{-}2}$, respectively, satisfying the boundary conditions

$$U_{m1}(0,\Omega)=U_{m2}(1,\Omega)=U_{f1}(0,\Omega)=U_{f2}(1,\Omega)=1 \qquad (2.11)_1$$

$$U_{m1}(1,\Omega)=U_{m2}(0,\Omega)=U_{f1}(1,\Omega)=U_{f2}(0,\Omega)=0 \qquad (2.11)_2$$

Since the coefficients V_m (x, Ω) and V_f (x, Ω) vary with x, determination of analytical expressions for the fundamental solutions is not a simple task. However, when Ω is not an eigenvalue of $(2.9)_{1\text{-}2}$, a method recently developed by Chung [10] which uses probability theory allows the solutions U_{mi} (x, Ω) and U_{fi} (x, Ω) to be expressed in closed-form as

$$U_{m1}(x,\Omega)=\int_{[B_\tau=0]}\exp\left(\int_0^\tau V_m(B_t,\Omega)dt\right)dp^x$$

$$U_{m2}(x,\Omega)=\int_{[B_\tau=1]}\exp\left(\int_0^\tau V_m(B_t,\Omega)dt\right)dp^x \qquad (2.12)_1$$

and

$$U_{f1}(x,\Omega)=\int_{[B_\tau=0]}\exp\left(\int_0^\tau V_f(B_t,\Omega)dt\right)dp^x$$

$$U_{f2}(x,\Omega)=\int_{[B_\tau=1]}\exp\left(\int_0^\tau V_f(B_t,\Omega)dt\right)dp^x \qquad (2.12)_2$$

where $[B_\tau: \tau>0]$ stands for a Brownian motion process, τ means the first exit time from the interval (0, 1), and p^x is the probability distribution of the process $[B_t: t>0]$ starting at x. In fact, using a random walk instead of Brownian motion, a second order finite difference scheme can be formulated to calculate the fundamental solutions $(2.12)_{1\text{-}2}$ very accurately. This procedure will be illustrated in a separate paper.

Complex forms of the fundamental solutions are given in terms of $U_{mi}(x, \Omega)$ and $U_{fi}(x, \Omega)$ by

$$\mathcal{U}_{m1}(x,\Omega) = U_{m1}(x,\Omega) + iU_{m2}(x,\Omega) \qquad (2.13)_1$$

$$\mathcal{U}_{m2}(x,\Omega) = U_{m1}(x,\Omega) - iU_{m2}(x,\Omega) \qquad (2.13)_2$$

$$\mathcal{U}_{f1}(x,\Omega) = U_{f1}(x,\Omega) + iU_{f2}(x,\Omega) \qquad (2.14)_1$$

$$\mathcal{U}_{f2}(x,\Omega) = U_{f1}(x,\Omega) - iU_{f2}(x,\Omega) \qquad (2.14)_2$$

They must satisfy the boundary conditions

$$\mathcal{U}_{m1}(0,\Omega) = \mathcal{U}_{m2}(0,\Omega) = \mathcal{U}_{f1}(0,\Omega) = \mathcal{U}_{f2}(0,\Omega) = 1 \qquad (2.15)$$

and

$$\mathcal{U}_{m1}(1,\Omega) = \mathcal{U}_{f1}(1,\Omega) = i, \quad \mathcal{U}_{m2}(1,\Omega) = \mathcal{U}_{f2}(1,\Omega) = -i \qquad (2.16)_{1\text{-}2}$$

The general solution for the displacement in the matrix layer of the *N-th* unit cell can be written in the form

$$u_m(x,\Omega) = \frac{1}{\sqrt{D_m(x)}}\,[A_m\mathcal{U}_{m1}(x,\Omega) + B_m\mathcal{U}_{m2}(x,\Omega)]e^{-i\Omega\tau} \qquad (2.17)$$

where A_m and B_m are nontrivial complex constants yet to be determined. Using the stress-strain relations, the stress in the matrix layer of the *N-th* unit cell is given by

$$\sigma_m = D_m\frac{\partial u_m}{\partial x} = \frac{1}{\sqrt{D_m(x)}}\,[A_m\sigma_{m1} + B_m\sigma_{m2}]e^{-i\Omega\tau} \qquad (2.18)$$

where

$$\sigma_{m1} = D_m\mathcal{U}'_{m1}(x,\Omega) - \frac{1}{2}\,D'_m\mathcal{U}_{m1}(x,\Omega) \qquad (2.19)_1$$

$$\sigma_{m2} = D_m\mathcal{U}'_{m2}(x,\Omega) - \frac{1}{2}\,D'_m\mathcal{U}_{m2}(x,\Omega) \qquad (2.19)_2$$

The prime represents a derivative of the associated quantity with respect to x. Similarly, in the fiber layer of the *N-th* unit cell, the displacement and stress take the forms

$$u_f(x,\Omega) = \frac{1}{\sqrt{D_f(x)}}\,[A_f\mathcal{U}_{f1}(x,\Omega) + B_f\mathcal{U}_{f2}(x,\Omega)]e^{-i\Omega\tau} \qquad (2.20)$$

$$\sigma_f = D_f\frac{\partial u_f}{\partial x} = \frac{1}{\sqrt{D_f(x)}}\,(A_f\sigma_{f1} + B_f\sigma_{f2})e^{-i\Omega\tau} \qquad (2.21)$$

where

$$\sigma_{f1} = D_f\mathcal{U}'_{f1}(x,\Omega) - \frac{1}{2}\,D'_f\mathcal{U}_{f1}(x,\Omega) \qquad (2.22)_1$$

$$\sigma_{f2} = D_f\mathcal{U}'_{f2}(x,\Omega) - \frac{1}{2}\,D'_f\mathcal{U}_{f2}(x,\Omega) \qquad (2.22)_2$$

To complete specification of the problem, continuity of displacement and traction must be enforced at the interface between matrix and fiber layers, which leads to

$$u_f(1,\Omega) = \xi u_m(0,\Omega) \quad \sigma_f(1,\Omega) = \eta\sigma_m(0,\Omega) \qquad (2.23)_{1\text{-}2}$$

and

$$u^*_f(0,\Omega) = \xi u_m(1,\Omega) \quad \sigma^*_f(0,\Omega) = \eta\sigma_m(1,\Omega) \qquad (2.24)_{1\text{-}2}$$

where $\xi = h_m/h_f$ and $\eta = \rho_{mo}c^2_{mo}/\rho_{fo}c^2_{fo}$. Moreover, u^*_f and σ^*_f represent the displacement and stress in the fiber layer of the $(N + 1)$*–th* unit cell, respectively.

According to Floquet's one-dimensional wave theory, $(2.7)_2$ with its periodic variable coefficients admits quasi-periodic recurrence relations for the displacement and stress between two adjacent cell units as follows

$$u^*_f(0,\Omega) = u_f(0,\Omega)e^{ikd} \quad \sigma^*_f(0,\Omega) = \sigma_f(0,\Omega)e^{ikd} \qquad (2.25)_{1\text{-}2}$$

where $k = k_1 + ik_2$ is the complex Floquet wave number. Combining continuity conditions $(2.24)_{1\text{-}2}$ with $(2.25)_{1\text{-}2}$, we obtain

$$u_f(0,\Omega)e^{ikd} = \xi u_m(1,\Omega) \quad \sigma_f(0,\Omega)e^{ikd} = \eta\sigma_m(1,\Omega) \qquad (2.26)_{1\text{-}2}$$

Substitution of expressions (2.18) and (2.21) for the stresses σ_m and σ_f into $(2.23)_2$ and $(2.26)_2$, respectively, yields a set of four homogeneous algebraic equations from which the unknown constants A_m, B_m, A_f, and B_f can be determined. A nontrivial solution for these constants exists if the corresponding determinant of the matrix of coefficients vanishes. Setting the determinant equal to zero leads to the following dispersion equation

$$\begin{vmatrix} i\epsilon & -i\epsilon & -1 & -1 \\ \epsilon e^{ikd} & \epsilon e^{ikd} & -i & i \\ \mathcal{U}'_{m1}(1,\Omega) + i\beta_m & \mathcal{U}'_{m2}(1,\Omega) - i\beta_m & -\mathcal{U}'_{f1}(0,\Omega) + \beta_f & -\mathcal{U}'_{f2}(0,\Omega) + \beta_f \\ (\mathcal{U}'_{m1}(0,\Omega) - \beta_m)e^{ikd} & (\mathcal{U}'_{m2}(0,\Omega) - \beta_m)e^{ikd} & -\mathcal{U}'_{f1}(1,\Omega) - i\beta_f & -\mathcal{U}'_{f2}(1,\Omega) + i\beta_f \end{vmatrix} = 0 \qquad (2.27)$$

where three new parameters have been introduced as

$$\epsilon = \frac{\eta}{\xi} \quad \beta_f = \frac{1}{2}\,D'_f(0) \quad \beta_m = \frac{1}{2}\,D'_m(0) \qquad (2.28)_{1\text{-}3}$$

We see from (2.27) that the dispersion relationship depends on derivatives of the complex fundamental solutions $\mathcal{U}_{mi}(x, \Omega)$ and $\mathcal{U}_{fi}(x, \Omega)$ evaluated at the end points $x=0$ and $x=1$. These derivatives are related to the corresponding derivatives of the real valued fundamental solutions as follows

$$\begin{aligned} \mathcal{U}'_{m1}(0,\Omega) &= U'_{m1}(0,\Omega) + iU'_{m2}(0,\Omega) & \mathcal{U}'_{m2}(0,\Omega) &= U'_{m1}(0,\Omega) - iU'_{m2}(0,\Omega) \\ \mathcal{U}'_{m1}(1,\Omega) &= U'_{m1}(1,\Omega) + iU'_{m2}(1,\Omega) & \mathcal{U}'_{m2}(1,\Omega) &= U'_{m1}(1,\Omega) - iU'_{m2}(1,\Omega) \\ \mathcal{U}'_{f1}(0,\Omega) &= U'_{f1}(0,\Omega) + iU'_{f2}(0,\Omega) & \mathcal{U}'_{f2}(0,\Omega) &= U'_{f1}(0,\Omega) - iU'_{f2}(0,\Omega) \\ \mathcal{U}'_{f1}(1,\Omega) &= U'_{f1}(1,\Omega) + iU'_{f2}(1,\Omega) & \mathcal{U}'_{f2}(1,\Omega) &= U'_{f1}(1,\Omega) - iU'_{f2}(1,\Omega) \end{aligned} \qquad (2.29)$$

Before expanding the determinant (2.27), we will examine some important features of the fundamental solutions $U_{mi}(x, \Omega)$ and $U_{fi}(x, \Omega)$. Henceforth, we assume that the material properties in each layer of any unit cell are symmetric with respect to the midplane of that layer.

Proposition 1. If the coefficient $V_f(x, \Omega)$ in $(2.9)_2$ is symmetric with respect to the midplane $x = 0.5$ in the interval (0, 1), such that

$$V_f(x,\Omega) = V_f(1-x,\Omega) \tag{2.30}$$

then the derivatives of the fundamental solutions $U'_{f1}(x, \Omega)$ and $U'_{f2}(x, \Omega)$ at both ends $x = 0$ and $x = 1$ satisfy

$$U'_{f1}(1,\Omega) + U'_{f2}(0,\Omega) = 0 \quad U'_{f1}(0,\Omega) + U'_{f2}(1,\Omega) = 0 \qquad (2.31)_{1\text{-}2}$$

Proof: The Wronskian $W_f[U_{f1}, U_{f2}]$ of the fundamental solutions $U_{f1}(x, \Omega)$ and $U_{f2}(x, \Omega)$ is defined as

$$W_f[U_{f1},U_{f2}] \equiv \begin{vmatrix} U_{f1}(x,\Omega) & U_{f2}(x,\Omega) \\ U'_{f1}(x,\Omega) & U'_{f2}(x,\Omega) \end{vmatrix} \tag{2.32}$$

Furthermore, due to the symmetry of $V_f(x, \Omega)$ and based on the basic behavior of the Wronskian, we have

$$\frac{dW_f}{dx} = 0 \qquad (0 \le x \le 1) \tag{2.33}$$

Thus

$$W_f(0,\Omega) = W_f(1,\Omega) \tag{2.34}$$

or written in expanded form

$$\begin{vmatrix} U_{f1}(0,\Omega) & U_{f2}(0,\Omega) \\ U'_{f1}(0,\Omega) & U'_{f2}(0,\Omega) \end{vmatrix} = \begin{vmatrix} U_{f1}(1,\Omega) & U_{f2}(1,\Omega) \\ U'_{f1}(1,\Omega) & U'_{f2}(1,\Omega) \end{vmatrix} \tag{2.35}$$

Therefore, according to the boundary conditions $(2.11)_{1\text{-}2}$, (2.35) is equivalent to $(2.31)_1$. Using the symmetry condition (2.30), $(2.31)_2$ is easily confirmed. The same argument follows for the behavior in the matrix layer.

Making use of (2.29) and (2.31), the dispersion relationship (2.27) reduces to the simple form[1-2]

$$e^{2ikd} + F(\Omega)e^{ikd} + 1 = 0 \tag{2.36}$$

where $F(\Omega)$ is called the "spectrum function" and is given by

$$F(\Omega) = \frac{\left(U'_{f1}(1,\Omega)\right)^2 + \epsilon^2\left(U'_{m1}(1,\Omega)\right)^2 - \left(U'_{f1}(0,\Omega) + \epsilon U'_{m1}(0,\Omega) - \beta_f - \epsilon\beta_m\right)^2}{\epsilon U'_{f1}(1,\Omega)\,U'_{m1}(1,\Omega)} \tag{2.37}$$

The derivatives of the first fundamental solutions $U_{m1}(x, \Omega)$ and $U_{f1}(x, \Omega)$ must be known at the ends $x = 0$ and $x = 1$ in order to determine the function $F(\Omega)$. For this reason, we present two basic properties of these derivatives.

Proposition 2. If Ω is not an eigenvalue of $(2.9)_{1\text{-}2}$, then the derivatives $U'_{m1}(1, \Omega)$ and $U'_{f1}(1, \Omega)$ must not vanish. This will insure that $F(\Omega)$ remains bounded.

Proof: If the independent and dependent variables are transformed according to

$$x^* = 1 - x \quad U^*_{f1}(x^*,\Omega) = U_{f1}(1-x,\Omega) \tag{2.38}$$

and the symmetry condition (2.30) is used, $(2.9)_2$ takes the form

$$\frac{d^2U^*_f}{dx^{*2}_f} + V_f(x^*,\Omega)U^*_f = 0 \tag{2.39}$$

On the other hand, according to $(2.11)_{1\text{-}2}$ and (2.38), we find that

$$U_{f1}(1,\Omega) = 0 \Rightarrow U^*_{f1}(0,\Omega) = 0 \tag{2.40}$$

Therefore, if $U'_{f1}(1,\Omega)$ is to vanish, according to (2.38), $U'^*_{f1}(0,\Omega)$ must vanish also, i.e.,

$$U'^*_{f1}(0,\Omega) = 0 \tag{2.41}$$

The existence and uniqueness theorem states that if Ω is not an eigenvalue of the ordinary differential equation (2.39), and homogeneous boundary conditions (2.40) and (2.41) are posed, then (2.39) has only a trivial solution

$$U^*_{f1}(x^*,\Omega) = 0 \tag{2.42}$$

Obviously

$$U^*_{f1}(1,\Omega) = 0 \Rightarrow U_{f1}(0,\Omega) = 0 \tag{2.43}$$

However, this conclusion is in contradiction to the original assumption $(2.11)_1$, and therefore $U'_{f1}(1, \Omega)$ must not vanish. The same argument can be made concerning $U'_{m1}(1, \Omega)$.

Proposition 3. If Ω is not an eigenvalue of $(2.9)_1$ the derivatives $U'_{f1}(0, \Omega)$ and $U'_{f1}(1, \Omega)$ can be expressed in the form

$$U'_{f1}(0,\Omega) = \frac{b_2u_1(1,\Omega) - b_1u_2(1,\Omega)}{a_2u_1(1,\Omega) - a_1u_2(1,\Omega)} \qquad (2.44)_1$$

$$U'_{f1}(1,\Omega) = \frac{a_2b_1 - a_1b_2}{a_2u_1(1,\Omega) - a_1u_2(1,\Omega)} \qquad (2.44)_2$$

where the functions $u_i(x, \Omega)$ $(i = 1, 2)$ are solutions of the following initial value problems

$$\frac{d^2u_i}{dx^2} + V_f(x,\Omega)u_i = 0 \quad (i = 1,2) \qquad (2.45)_1$$

$$u_i(0,\Omega) = a_i \quad u'_i(0,\Omega) = b_i \qquad (2.45)_2$$

where (a_1, b_1) and (a_2, b_2) are two pairs of arbitrary constants satisfying the condition

$$a_1b_2 - a_2b_1 \neq 0 \tag{2.46}$$

Proof: We form the Wronskians $W_f[u_1, U_{f1}]$ and $W_f[u_2, U_{f1}]$, and then use (2.34), which leads to

$$\begin{aligned} W_f[u_1,U_{f1}] &= u_1(0,\Omega)\,U'_{f1}(0,\Omega) - u'_1(0,\Omega)\,U_{f1}(0,\Omega) \\ &= u_1(1,\Omega)\,U'_{f1}(1,\Omega) - u'_1(1,\Omega)\,U_{f1}(1,\Omega) \end{aligned} \qquad (2.47)_1$$

$$\begin{aligned} W_f[u_2,U_{f1}] &= u_2(0,\Omega)\,U'_{f1}(0,\Omega) - u'_2(0,\Omega)\,U_{f1}(0,\Omega) \\ &= u_2(1,\Omega)\,U'_{f1}(1,\Omega) - u'_2(1,\Omega)\,U_{f1}(1,\Omega) \end{aligned} \qquad (2.47)_2$$

Invoking the boundary conditions $(2.11)_{1\text{-}2}$ and initial conditions $(2.45)_2$, $(2.47)_{1\text{-}2}$ reduces to

$$a_1U'_{f1}(0,\Omega) - u_1(1,\Omega)\,U'_{f1}(1,\Omega) = b_1 \qquad (2.48)_1$$

$$a_2U'_{f1}(0,\Omega) - u_2(1,\Omega)\,U'_{f1}(1,\Omega) = b_2 \qquad (2.48)_2$$

If Ω is not an eigenvalue of (2.9), we can select the constants a_1 and a_2 such that

$$a_1u_2(1,\Omega) - a_2u_1(1,\Omega) \neq 0 \tag{2.49}$$

Therefore, the algebraic equations $(2.48)_{1\text{-}2}$ have the unique solution given by $(2.44)_{1\text{-}2}$. From $(2.44)_{1\text{-}2}$ we conclude that the problem of finding the derivatives $U'_{f1}(0, \Omega)$ and $U'_{f1}(1, \Omega)$, which are needed to specify the dispersion relationship (2.27), reduces to solving the initial value problem $(2.45)_{1\text{-}2}$ to obtain the values of u_i at the end $x = 1$. An identical procedure is followed to find the derivatives $U'_{m1}(0, \Omega)$ and $U'_{m1}(1, \Omega)$.

In order to calculate stresses in the entire interval (0, 1), we need to find the derivatives of the fundamental solutions $U_{mi}(x, \Omega)$ and $U_{fi}(x, \Omega)$. They can be calculated by the following procedure.

Proposition 4. If Ω is not an eigenvalue of $(2.9)_{1\text{-}2}$ we can use Chung's method to express the derivatives of the first and second fundamental solutions $U_{f1}(x, \Omega)$ and $U_{f2}(x, \Omega)$ as

$$U'_{f1}(x,\Omega) = \sqrt{\frac{V_f(x,\Omega)}{V_f(0,\Omega)}}\,(U'_{f1}(0,\Omega)L_{f0}(x,\Omega)$$

$$+\,U'_{f1}(1,\Omega)L_{f1}(x,\Omega)) \tag{2.50$_1$}$$

$$U'_{f2}(x,\Omega)=-\sqrt{\frac{V_f(x,\Omega)}{V_f(0,\Omega)}}\,(U'_{f1}(1,\Omega)L_{f0}(x,\Omega)$$

$$+\,U'_{f1}(0,\Omega)L_{f1}(x,\Omega)) \tag{2.50$_2$}$$

and likewise for $U_{m1}(x, \Omega)$ and $U_{m2}(x, \Omega)$

$$U'_{m1}(x,\Omega)=\sqrt{\frac{V_m(x,\Omega)}{V_m(0,\Omega)}}\,(U'_{m1}(0,\Omega)L_{m0}(x,\Omega)$$

$$+\,U'_{m1}(1,\Omega)L_{m1}(x,\Omega)) \tag{2.51$_1$}$$

$$U'_{m2}(x,\Omega)=-\sqrt{\frac{V_m(x,\Omega)}{V_m(0,\Omega)}}\,(U'_{m1}(1,\Omega)L_{m0}(x,\Omega)$$

$$+\,U'_{m1}(0,\Omega)L_{m1}(x,\Omega)) \tag{2.51$_2$}$$

where

$$L_{m0}(x,\Omega)=\int_{[B_\tau=0]}\exp\Big(\int_0^\tau H_m(B_t,\Omega)dt\Big)dp^x \tag{2.52$_1$}$$

$$L_{m1}(x,\Omega)=\int_{[B_\tau=1]}\exp\Big(\int_0^\tau H_m(B_t,\Omega)dt\Big)dp^x \tag{2.52$_2$}$$

$$L_{f0}(x,\Omega)=\int_{[B_\tau=0]}\exp\Big(\int_0^\tau H_f(B_t,\Omega)dt\Big)dp^x \tag{2.53$_1$}$$

$$L_{f1}(x,\Omega)=\int_{[B_\tau=1]}\exp\Big(\int_0^\tau H_f(B_t,\Omega)dt\Big)dp^x \tag{2.53$_2$}$$

The functions H_m and H_f are given by

$$H_m(x,\Omega)=V_m(x,\Omega)+\frac{1}{2}\frac{d^2}{dx^2}\Big[\log V_m(x,\Omega)\Big]$$

$$-\frac{1}{4}\frac{d}{dx}\Big[\log V_m(x,\Omega)\Big] \tag{2.54$_1$}$$

$$H_f(x,\Omega)=V_f(x,\Omega)+\frac{1}{2}\frac{d^2}{dx^2}\Big[\log V_f(x,\Omega)\Big]$$

$$-\frac{1}{4}\frac{d}{dx}\Big[\log V_f(x,\Omega)\Big] \tag{2.54$_2$}$$

$U'_{f1}(1, \Omega)$, $U'_{f1}(1, \Omega)$, $U'_{f2}(0, \Omega)$, and $U'_{f2}(1, \Omega)$ are known after considering (2.31) and (2.44).

Proof: We introduce a new function $Y_1(x, \Omega)$ by

$$Y_1(x,\Omega)=\sqrt{\frac{V_f(0,\Omega)}{V_f(x,\Omega)}}\,U'_{f1}(x,\Omega) \tag{2.55}$$

Substituting (2.55) into $(2.9)_2$ and using the symmetry condition $V_f(0, \Omega) = V_f(1, \Omega)$, we obtain a boundary value problem for the unknown function $Y(x, \Omega)$ as follows

$$\frac{d^2Y_f}{dx^2}+H_f(x,\Omega)\,Y_1(x,\Omega)=0 \tag{2.56$_1$}$$

$$Y_1(0,\Omega)=U'_{f1}(0,\Omega),\quad Y_1(1,\Omega)=U'_{f1}(1,\Omega) \tag{2.56$_2$}$$

Thus, following the same procedure that lead to $(2.12)_{1\text{-}2}$, we obtain the first and second fundamental solutions of the above boundary value problem as expressed in $(2.50)_{1\text{-}2}$ and $(2.51)_{1\text{-}2}$.

3 Pass and Stop Bands in Nonhomogeneous Composites

The most important feature regarding wave propagation in a periodically layered, elastic, homogeneous medium is the presence of stop band characteristics. Next we investigate how this characteristic is affected by specific material nonhomogeneities.

When the basic dispersion relationship (2.36) is expanded, the following two equations emerge

$$e^{-2k_2d}\cos 2k_2d+F(\Omega)e^{-k_2d}\cos k_1d+1=0 \tag{3.1$_1$}$$

$$e^{-k_2d}\sin k_1d\Big[e^{-k_2d}\cos k_1d+\frac{1}{2}F(\Omega)\Big]=0 \tag{3.1$_2$}$$

where the complex Floquet wave number k has been decomposed into a real part k_1 and an imaginary part k_2, called the dispersion coefficient and dissipation coefficient, respectively. In order to find the specific dependence of k_1 and k_2 on Ω, $(3.1)_{1\text{-}2}$ possesses must be solved simultaneously. The solution depends on the magnitude of the function $F(\Omega)$ as follows

(1) When

$$\Big|\frac{1}{2}F(\Omega)\Big|\le 1 \tag{3.2}$$

$(3.1)_{1\text{-}2}$ possesses the unique solution

$$\cos k_1d=-\frac{1}{2}F(\Omega) \tag{3.3$_1$}$$

$$k_2d=0 \tag{3.3$_2$}$$

Thus, the dissipation coefficient k_2 vanishes, and the pass band in the dispersion spectrum consists of all nondimensional frequencies Ω which satisfy (3.2). In other words, harmonic waves are propagated without attentuation for values of Ω which satisfy (3.2).

(2) When

$$\Big|\frac{1}{2}F(\Omega)\Big|>1 \tag{3.4}$$

then $(3.1)_{1\text{-}2}$ possess the solution

$$k_1d=n\pi \qquad (n=0,1,2,\ .\ .\ .) \tag{3.5$_1$}$$

$$k_2d=-\log\Bigg(\Big|\frac{F(\Omega)}{2}\Big|-\sqrt{\Big(\frac{F(\Omega)}{2}\Big)^2-1}\Bigg) \tag{3.5$_2$}$$

Here the dissipation coefficient k_2 does not vanish. Therefore, when the frequency Ω results in (3.4) being satisfied, harmonic waves are attenuated as they pass through the medium. This is the proof of the presence of stop bands, and $(3.5)_2$ predicts how the dissipation coefficient depends on the frequency.

We have assumed that the nondimensional frequency Ω is not an eigenvalue of $(2.9)_{1\text{-}2}$, so that the function $F(\Omega)$ must be finite. We now examine the case when Ω becomes an eigenvalue of $(2.9)_1$ or $(2.9)_2$. In the following, we refer to the eigenvalue as Ω^*.

Proposition 5. $U'_{f1}(1, \Omega)$ and $U'_{f2}(1, \Omega)$ tend to infinity if and only if Ω tends to Ω^*.

Proof: Let $u^* = u^*(x, \Omega^*)$ be one of the eigenfunctions corresponding to Ω^*. According to the definition of the eigenvalue, we must have

$$u^*(0,\Omega^*)=u^*(1,\Omega)=0 \tag{3.6}$$

If we further assume that

$$u_1=u^*(x,\Omega^*)\quad a_1=0\quad b_1=u^{*\prime}(0,\Omega^*)\ne 0 \tag{3.7}$$

From $(2.44)_{1\text{-}2}$, we find that when $a_1 = 0$ and $u_1(1, \Omega) = 0$

$$U'_{f1}(0,\Omega)\to\infty\quad U'_{f1}(1,\Omega)\to\infty \tag{3.8}$$

On the other hand, for $|u'_1(0, \Omega^*)|$ to become unbounded, it is seen from $(2.44)_2$ that for any two nontrivial constants $u_2(0, \Omega)$ and $u_2(1, \Omega)$, the expression

$$u_1^*(1,\Omega^*)u_2(0,\Omega)-u_1^*(0,\Omega^*)u'_2(0,\Omega)\to 0 \tag{3.9}$$

Obviously, this is the case only if both $u_1^*(0, \Omega^*)$ and $u_1^*(1, \Omega^*)$ tend to zero simultaneously. This also means that Ω must be one of the eigenvalues of $(2.9)_{1\text{-}2}$.

The eigenfrequency Ω^* could reside in either the pass band or the stop band. This depends on the behavior of function $F(\Omega)$ as Ω tends to Ω^*. We take

$$F^* = \lim_{\Omega \to \Omega^*} F(\Omega) \tag{3.10}$$

Therefore, as stated in (3.2), if $|F^*| \leq 2$, Ω^* is in the pass band, otherwise, it falls in a stop band. For nonhomogeneous composites, there may exist a special eigenfrequency where $|F^*|$ becomes infinite. We would call such an eigenfrequency a *pole* in the frequency spectrum. If a pole would occur, the dissipation coefficient $k_2(\Omega^*)$ would become infinite. The amplitude of any harmonic waves would be immediately attenuated at this frequency. In the next section, we show by example that at this special frequency all eigenfrequencies for homogeneous layered composites lie entirely in the pass band. Therefore, a pole cannot exist for homogeneous composites. This raises a very natural question: For nonhomogeneous composites, can a pole actually occur? It has been proved that the derivatives of the first and second fundamental solutions $U_{f1}(x, \Omega)$ and $U_{f2}(x, \Omega)$ become infinite at both ends $x = 0$ and $x = 1$ when Ω is one of the eigenfrequencies. In this case, the spectrum function $F(\Omega)$ is approximated by

$$F(\Omega) \approx \frac{(U'_{f1}(1,\Omega))^2 - (U'_{f1}(0,\Omega))^2}{U'_{f1}(1,\Omega)} + \text{const}\ \frac{U'_{f1}(0,\Omega)}{U'_{f1}(1,\Omega)} \tag{3.11}$$

Thus, if $|U'_{f1}(0, \Omega)| \neq |U'_{f1}(1, \Omega)|$ as Ω tends to Ω^*, then $F^*(\Omega^*)$ tends to infinity. This must indicate the presence of a "pole." A more detailed discussion on this interesting problem is actually needed and will be addressed in another paper. In what follows, we give some examples to illustrate the differences in the behavior of the dispersion relationship between homogeneous and nonhomogeneous composites.

4 Examples and Discussion of Numerical Results

Example 1. As a special case of the general theory presented above, we will calculate the spectrum function and associated dispersion characteristics for an elastic composite with homogeneous layers. In this case, the mass density and elastic moduli are constant within each matrix and fiber layer. We then have

$$V_m = \pi^{*2}\Omega^2, \quad V_f = \pi^2\Omega^2, \tag*{(4.1)$_{1\text{-}2}$}$$

The first and second fundamental solutions of $(2.7)_{1\text{-}2}$ are given by

$$U_{m1}(x,\Omega) = \frac{\sin \pi^*\Omega(1-x)}{\sin \pi^*\Omega} \quad U_{m2}(x,\Omega) = \frac{\sin \pi^*\Omega x}{\sin \pi^*\Omega} \tag*{(4.2)$_{1\text{-}2}$}$$

$$U_{f1}(x,\Omega) = \frac{\sin \pi\Omega(1-x)}{\sin \pi\Omega} \quad U_{f2}(x,\Omega) = \frac{\sin \pi\Omega x}{\sin \pi\Omega} \tag*{(4.3)$_{1\text{-}2}$}$$

Therefore, the derivatives of the first two fundamental solutions $U_{m1}(x, \Omega)$ and $U_{f1}(x, \Omega)$ are given by

$$U'_{m1}(x,\Omega) = -\pi^*\Omega\, \frac{\cos \pi^*\Omega(1-x)}{\sin \pi^*\Omega} \tag*{(4.4)$_1$}$$

$$U'_{f1}(x,\Omega) = -\pi\Omega\, \frac{\cos \pi\Omega(1-x)}{\sin \pi\Omega} \tag*{(4.4)$_2$}$$

From $(4.2)_{1\text{-}2}$, $(4.3)_{1\text{-}2}$, and $(4.4)_{1\text{-}2}$,

$$U'_{m1}(0,\Omega) = -\pi^*\Omega \cot \pi^*\Omega,$$

$$U'_{m1}(1,\Omega) = -\pi^*\Omega \csc \pi^*\Omega \tag*{(4.5)$_{1\text{-}2}$}$$

$$U'_{f1}(0,\Omega) = -\pi\Omega \cot \pi\Omega \quad U'_{f1}(1,\Omega) = -\pi\Omega \csc \pi\Omega \tag*{(4.6)$_{1\text{-}2}$}$$

Substituting $(4.5)_{1\text{-}2}$ and $(4.6)_{1\text{-}2}$ into (2.37), we find the spectrum function has the form

$$G(\Omega) = -\frac{F(\Omega)}{2} = \cos \pi\Omega \cos \pi^*\Omega - \lambda \sin \pi\Omega \sin \pi^*\Omega \tag{4.7}$$

where $\lambda = \frac{1}{2}\left[\frac{\rho_{mo}c_{mo}}{\rho_{fo}c_{fo}} + \frac{\rho_{fo}c_{fo}}{\rho_{mo}c_{mo}}\right]$.

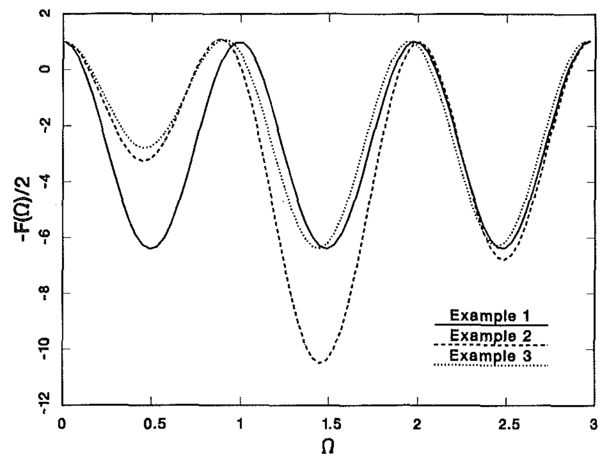


Fig. 2 Spectrum function for three example problems

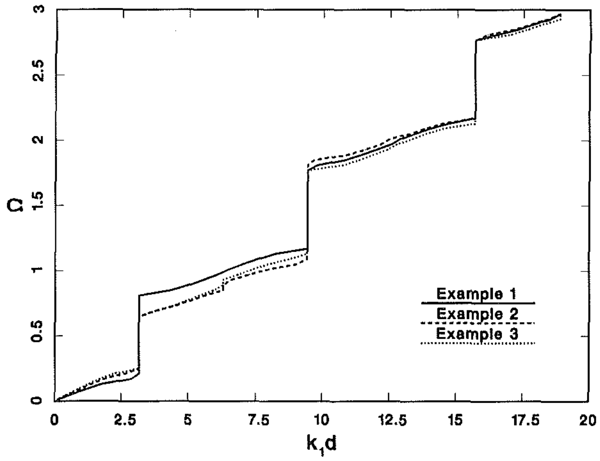


Fig. 3 Dispersion coefficient for three example problems

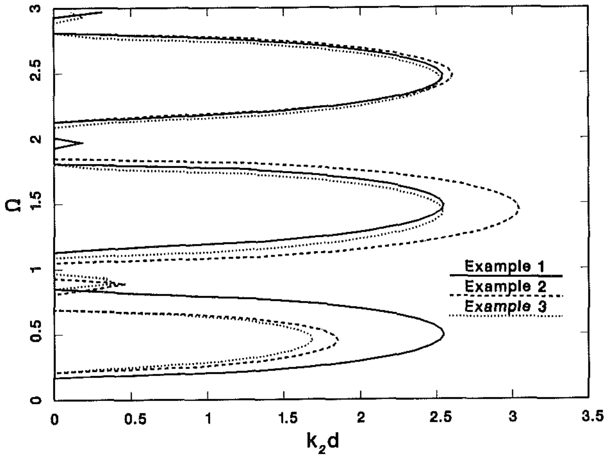


Fig. 4 Dissipation coefficient for three example problems

Thus, based on previous arguments, we obtain the spectral behavior as follows

$$\cos k_1 d = \begin{cases} G(\Omega), & \text{for } |G(\Omega)| \leq 1 \\ (-1)^n, & \text{for } |G(\Omega)| > 1 \end{cases} \tag{4.8}$$

where n is a positive integer and

$$k_2 d = \begin{cases} 0, & \text{for } |G(\Omega)| \leq 1 \\ \log\left(|G(\Omega)| - \sqrt{(G(\Omega))^2 - 1}\right), & \text{for } |G(\Omega)| > 1 \end{cases} \tag{4.9}$$

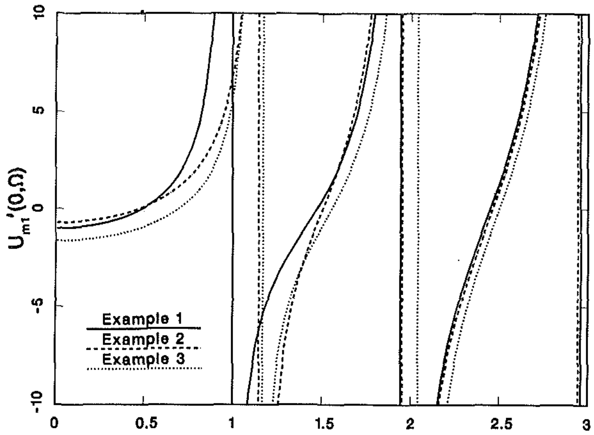


Fig. 5 Derivative of fundamental solution U_{m1} at $x = 0$ for three example problems

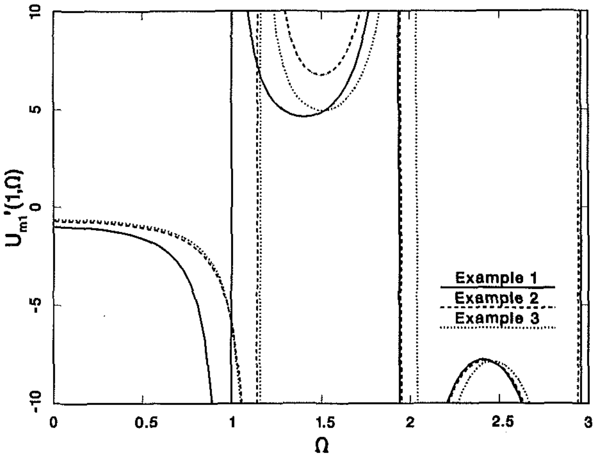


Fig. 6 Derivative of fundamental solution U_{m1} at $x = 1$ for three example problems

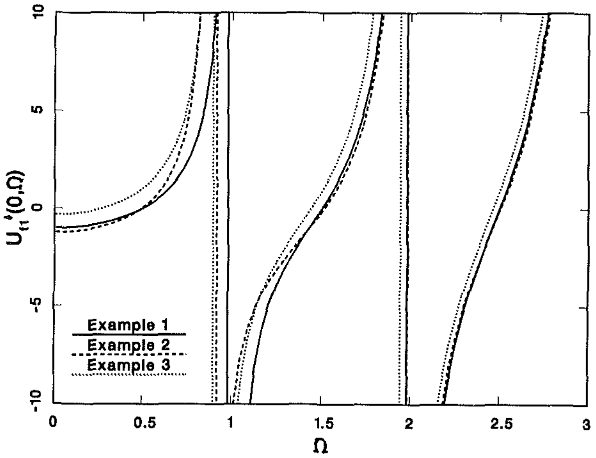


Fig. 7 Derivative of fundamental solution U_{f1} at $x = 0$ for three example problems

Thus, when the relation $|G(\Omega)| \leq 1$ is satisfied, the corresponding frequency is in the pass band. Otherwise, the frequencies are in the stop band. These results are the same as given by Lee and Yang [5] and Delph et al [7-8]. From these solutions, we see that there are two sets of eigenfrequencies

$$\Omega_m^* = \frac{\pi}{\pi^*} n, \quad \Omega_f^* = n \qquad (4.10)_{1\text{-}2}$$

at which both derivatives $U_{f1}'(x, \Omega)$, and $U_{m1}'(x, \Omega)$ tend to infinity at the boundaries $x = 0$ and $x = 1$. After substituting $(4.10)_{1\text{-}2}$ into (4.7), we obtain

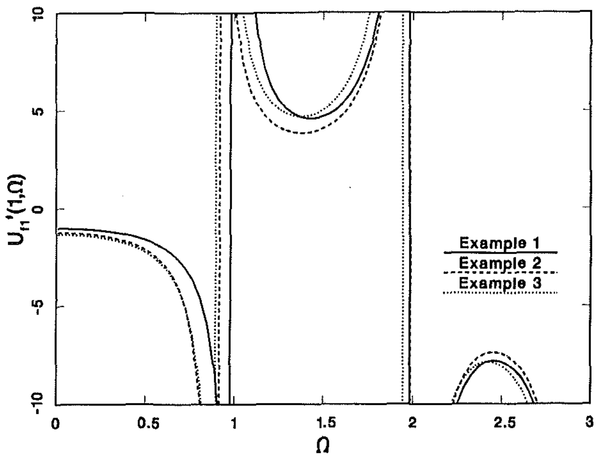


Fig. 8 Derivative of fundamental solution U_{f1} at $x = 1$ for three example problems

$$G(\Omega_f^*) = (-1)^n \cos n\pi^* \qquad (4.11)_1$$

$$G(\Omega_m^*) = (-1)^n \cos \frac{\pi^2}{\pi^*} n \qquad (4.11)_2$$

For these eigenfrequencies, we always have

$$|G(\Omega_f^*)| \leq 1, \quad |G(\Omega_m^*)| \leq 1 \qquad (4.12)_{1\text{-}2}$$

which means that for homogeneous, layered materials, all eigenfrequencies are in the pass band.

Figure 2 shows the spectrum function $G(\Omega) = -\frac{1}{2}F(\Omega)$ versus Ω for this composite. In this example we have assumed that $h_m/h_f = 0.25$, $D_m(0)/D_f(0) = 0.02$, and $\rho_m(0)/\rho_f(0) = 0.33$. These values are the same as those used in the paper by Delph et al. [7]. Figures 3 and 4 show the dispersion relations Ω versus $k_1 d$ and Ω versus $k_2 d$, respectively. Figures 5–8 show the behavior of the derivatives of fundamental solutions U_{m1}' $(0, \Omega)$, U_{m1}' $(1, \Omega)$, U_{f1}' $(0, \Omega)$, and U_{f1}' $(1, \Omega)$, respectively.

Example 2. As an example of a nonhomogeneous elastic composite, we will consider the following variation of the material constants in each of the layers

$$\rho_m(x) = 1 - \frac{\rho_o}{2}[1 + \cos \pi(2x-1)] \qquad (4.13)_1$$

$$D_m(x) = 1 - \frac{D_o}{2}[1 + \cos \pi(2x-1)] \qquad (4.13)_2$$

and

$$\rho_f(x) = 1 + \frac{\rho_o}{2}[1 + \cos \pi(2x-1)] \qquad (4.14)_1$$

$$D_f(x) = 1 + \frac{D_o}{2}[1 + \cos \pi(2x-1)] \qquad (4.14)_2$$

where ρ_o and D_o are two positive parameters each less than unity.

The spectrum function $-F(\Omega)/2$ is shown in Fig. 2 for this nonhomogeneous composite. We have assumed the same values for the parameters h_m/h_f, $D_m(0)/D_f(0)$, and $\rho_m(0)/\rho_f(0)$ as in Example 1. The parameters D_o and ρ_o were both chosen to be 0.5. Figures 3 and 4 show the dispersion relations, while Figs. 5–8 show the derivatives of the fundamental solutions for this example.

Example 3. As another example of a nonhomogeneous elastic composite, we will assume the following quadratic variation of the material constants

$$\rho_m(x) = 1 + \rho_o x(1-x) \qquad (4.15)_1$$

$$D_m(x) = 1 + D_o x(1-x) \qquad (4.15)_2$$

and

$$\rho_f(x) = 1 + \rho_o x(1-x) \quad (4.16)_1$$

$$D_f(x) = 1 + D_o x(1-x) \quad (4.16)_2$$

where ρ_o and D_o are two positive parameters each less than 4.0. For this case, the values of all parameters were taken to be the same as in Example 2, except D_o and ρ_o were both equal to 2.0. This meant that at $x = 0.5$, the magnitude of the material properties was the same as in Example 2. The spectrum function, dispersion relations and the derivatives of the fundamental solutions are shown in Fig. 2, Figs. 3 and 4, and Figs. 5–8, respectively.

Figure 2 exemplifies the effect of material nonhomogeneity on the behavior of the spectrum function. The nonhomogeneity is seen to affect both the amplitude and phase of the spectrum function. Figures 3 and 4 show how the material nonhomogeneity changes the basic dispersion relationship. The vertical line segments in Fig. 3 are the stop bands and indicate the range of frequency where attenuation of the wave amplitude will occur. The nonhomogeneity is seen to affect the width of the stop bands, particularly at low frequencies. A sharp decrease of the attenuation coefficient k_2d at low frequencies can be seen from Fig. 4. The material nonhomogeneities which have been considered are seen not to affect the high frequency behavior of the spectrum function and the corresponding dispersion relationship when compared to composites constructed of homogeneous layers.

Figures 5 and 6 show the derivatives of the fundamental solutions U'_{m1} $(0, \Omega)$ and U'_{m1} $(1, \Omega)$. Eigenfrequencies are indicated where the derivatives become unbounded. Figures 7 and 8 show the derivatives U'_{f1} $(0, \Omega)$ and U'_{f1} $(1, \Omega)$. Here eigenfrequencies also exist. On comparing these results with Figs. 2 and 3, it is seen that the eigenfrequencies lie within the stop bands for the nonhomogeneous composites and within the pass bands for the homogeneous composites. In addition, a "pole" was not discovered during the calculation. Further studies to either discover or rule out the presence of such a feature is necessary.

Our attention has focused only on the investigation of the dispersion relations for nonhomogeneous composites by combining Floquet's wave theory with Wronskian properties of the fundamental solutions of the associated differential equations. We will defer the calculation of the vibrational mode shapes and discussion of further details of Chung's probability theory to a later paper.

Acknowledgments

This research was supported in part by DOE Research Contract No. DE-AT03-82ER12040 to Stanford University. One of the authors, Dr. Z. P. Duan acknowledges many helpful discussions with Dr. Z. X. Chao of the Institute of Mathematics, Chinese Academy of Sciences, during his visit to Stanford University in 1983.

References

1 Sun, C. T., Achenbach, J. D., and Herrmann, G., "Continuum Theory for a Laminated Medium," ASME Journal of Applied Mechanics, Vol. 5, 1968, pp. 467–473.

2 Drumheller, D. S., and Bedford, A., "Wave Propagation in Elastic Laminates Using a Second Order Microstructure," *International Journal of Solids and Structures*, Vol. 10, 1974, pp. 61–76.

3 Hegemier, G. A., and Nayfeh, A. H., "A Continuum Theory for Wave Propagation in Laminated Composites. Case I: Propagation Normal to the Laminates," ASME Journal of Applied Mechanics, Vol. 40, 1973, pp. 503–510.

4 Nemat-Nasser, S., and Fu, F. C. L., "On Harmonic Waves in Layered Composites: Bounds on Frequencies," ASME Journal of Applied Mechanics, Vol. 41, 1974, pp. 288–290.

5 Lee, E. H., and Yang, W. H., "On Waves in Composite Materials with Periodic Structure," *SIAM Journal of Applied Mechanics*, Vol. 25, 1973, pp. 492–499.

6 Bedford, A., Drumheller, D. S., and Sutherland, H. J., "On Modeling the Dynamics of Composite Materials," Nemat-Nasser, S., ed., *Mechanics Today*, Vol. 3, 1976, pp. 1–54.

7 Delph, T. J., Herrmann, G., and Kaul, R. K., "Harmonic Wave Propagation in a Periodically Layered, Infinite Elastic Body: Antiplane Strain," ASME Journal of Applied Mechanics, Vol. 45, 1978, pp. 343–349.

8 Delph, T. J., Herrmann, G., and Kaul, R. K., "Harmonic Wave Propagation in a Periodically Layered, Infinite Elastic Body: Plane Strain, Analytical Results," ASME Journal of Applied Mechanics, Vol. 46, 1978, pp. 113–119.

9 Delph, T. J., Herrmann, G., and Kaul, R. K., "Harmonic Wave Propagation in a Periodically Layered, Infinite Elastic Body: Plane Strain, Numerical Results," ASME Journal of Applied Mechanics, Vol. 47, 1980, pp. 531–537.

10 Chung, K. L., "Probabilistic Approach to Boundary Value Problems for Schrödinger's Equation," *Expositiones Mathematicae*, Vol. 3, 1985, pp. 175–178.

T. Fett
Kernforschungszentrum Karlsruhe,
Institut für Material- und
Festkörperforschung IV,
Karlsruhe, West Germany D-7500

Temperature Distributions and Thermal Stresses in Asymmetrically Heat Radiated Tubes

The transient and stationary temperature distributions in a tube wall caused by an asymmetrical heat flux distribution are evaluated. The results are represented for the case of a heat radiating half-space. In addition, the accompanying stress distributions are computed.

Introduction

The mechanical behavior of asymmetrically heated tubes is of interest in solar energy and fusion energy generation. In both cases the tubes are cyclically heated. For a fatigue evaluation the cyclic temperature and stress distribution of these tubes have to be known. For fusion reactors the temperature distribution must also be known because creep and swelling are temperature dependent.

While stationary temperature distributions in asymmetrically heated tubes are well-known from analysis [1], nonstationary distributions are often treated through the numerical solution of the equation describing transient heat conduction [2, 3] and in the last 10 years by finite-element-analysis [4, 5]. In this paper an analytical solution of this problem is communicated.

1 Temperature Distribution in Tubes

1.1 Basic Equations and Boundary Conditions. The basis of the following calculations is the equation for transient heat conduction

$$\frac{\partial T}{\partial t} = x \Delta T \quad ; \quad x = (\Lambda c d)^{-1} \tag{1}$$

where T is the temperature, t the time, $1/\Lambda$ the thermal conductivity, c the specific heat, and d the density. The Laplacian is written in cylindric coordinates as

$$\Delta = \frac{1}{\rho}\frac{\partial}{\partial \rho} + \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho^2}\frac{\partial^2}{\partial \varphi^2} \tag{2}$$

where the radius ρ and the angle φ are describing a point in the tube wall. The boundary conditions are given by a constant inner surface temperature (arbitrary chosen: $T = 0$) and the heat flux $\dot{Q}$ at the outer surface.

$$T(\rho = r) = 0 \tag{3}$$

$$\frac{\partial T}{\partial \rho}(\rho = R) = -\Lambda \dot{Q}(\varphi) \tag{4}$$

1.2 Fourier Components of the Time Dependent Temperature Distributions. To solve equation (1) under boundary conditions given by equations (3) and (4) it is usual to carry out a Laplace transformation to eliminate the time dependence. If $\bar{T}$ denotes the Laplace transformation of the temperature T

$$\bar{T} = \int_0^\infty T(t)\exp(-pt)\,dt \tag{5}$$

one obtains from equation (1), considering the initial condition $T(t=0) = 0$,

$$\frac{\partial^2 \bar{T}}{\partial \rho^2} + \frac{\partial \bar{T}}{\rho \partial \rho} + \frac{\partial^2 \bar{T}}{\rho^2 \partial \varphi^2} - q^2 \bar{T} = 0 \tag{6}$$

with $q^2 = p/\kappa$.

Since equation (6) is dependent only on ρ and φ, insertion of the usual set-up

$$\bar{T}(\rho,\varphi) = f(\rho)\cdot g(\varphi) \tag{7}$$

into equation (6) gives after separation of both variables

$$\frac{\rho^2 \partial^2 f}{f \partial \rho^2} + \frac{\rho \partial f}{f \partial \rho} - \rho^2 q^2 = \frac{-\partial^2 g}{g \partial \varphi^2} \tag{8}$$

The Laplace-transformed boundary conditions are

$$\bar{T}(\rho = r) = 0 \tag{9}$$

$$\frac{\partial \bar{T}}{\partial \rho}(\rho = R) = \frac{D}{p} = \chi \frac{D}{q^2} \qquad D = -\Lambda \dot{Q} \tag{10}$$

If D_n denotes the nth Fourier component of D, the nth Fourier component of the related temperature becomes after a lengthy analysis [6] for heat flux distributions symmetrical to $\varphi = 0$ (only cos-terms)

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October 11, 1984; final revision, August 30, 1985.

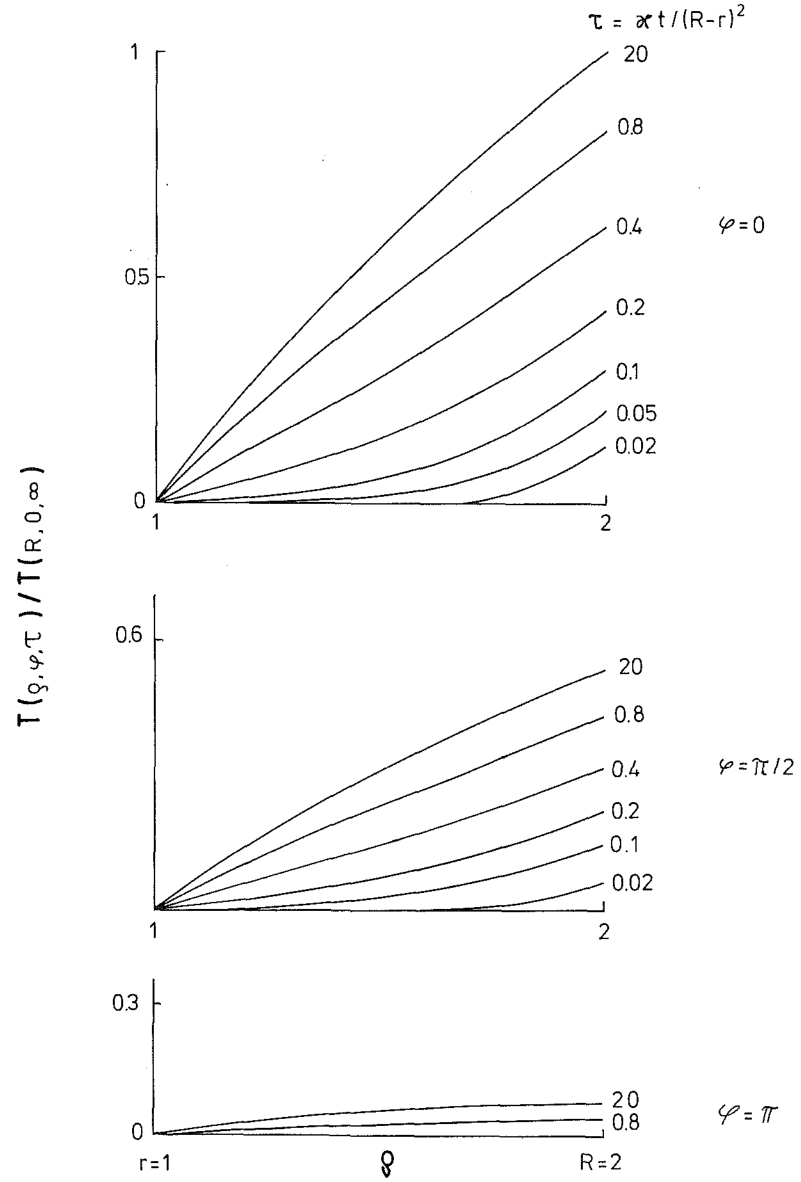


Fig. 1 **Time dependent temperature distribution in a tube in front of a heat radiating half-space**

$$T_{n\rho} = D_n \frac{R}{n} \frac{(\rho/r)^n - (r/\rho)^n}{(R/r)^n + (r/R)^n} - \pi D_n \sum_{k=1}^{\infty} e^{-\chi\alpha_k^2 t}$$

$$\frac{J_n'(\alpha_k R) J_n(\alpha_k r)[J_n(\alpha_k\rho) Y_n(\alpha_k r) - J_n(\alpha_k r) Y_n(\alpha_k\rho)]}{\alpha_k[(1 - n^2/\alpha_k^2 R^2) J_n^2(\alpha_k r) - J_n'^2(\alpha_k R)]}$$

$$\text{for } n > 0 \tag{11}$$

and

$$T_{o\rho} = D_o R \ln(\rho/r) + \pi D_o \sum_{k=1}^{\infty} e^{-\chi\alpha_k^2 t}$$

$$\frac{J_1(\alpha_k R) J_0(\alpha_k r)[J_0(\alpha_k\rho) Y_0(\alpha_k r) - J_0(\alpha_k r) Y_0(\alpha_k\rho)]}{\alpha_k[J_0^2(\alpha_k r) - J_1^2(\alpha_k R)]}$$

$$\text{for } n = 0 \tag{12}$$

where α_k are the positive roots of the equation

$$\frac{n}{R}[Y_n(\alpha_k r) J_n(\alpha_k R) - J_n(\alpha_k r) Y_n(\alpha_k R)]$$
$$-\alpha_k[Y_n(\alpha_k r) J_{n+1}(\alpha_k R) - J_n(\alpha_k r) Y_{n+1}(\alpha_k R)] = 0 \tag{13}$$

J_n and Y_n are the Bessel functions of order n.

The case $n = 0$ is also treated in [1]. The results are in agreement. The general solution is given by

$$T(\rho,\varphi,t) = \sum_{n=0}^{\infty} T_{n\rho}(\rho,t) \cos n\varphi \tag{14}$$

1.3 Temperature Distribution. Since the Fourier components of the temperature are known, any possible heat flux distribution which is symmetrical with respect to $\varphi = 0$ can be evaluated. Here only results for a tube in front of a heat radiating half-space will be communicated. In this special case the heat flux is given by

$$\dot{Q}(\varphi) = \frac{1}{2} \dot{Q}(0)[1 + \cos \varphi] \tag{15}$$

The Fourier series has only two terms and evaluation becomes very simple. Figure 1 shows the time dependent temperature distribution for a thick-walled tube with $R/r = 2$ for $\varphi = 0$, $\pi/2$ and π. The temperature values are normalized to the temperature of the hottest point on the outer surface ($\varphi = 0$, $\rho = R$).

The stationary temperature distribution follows from equations (11) and (12) for $t \to \infty$

$$T_\infty(\rho,\varphi) = \frac{\dot{Q}(0)}{2}\,\Lambda R\left[\ln\rho/r + \frac{\rho/r - r/\rho}{R/r + r/R}\cos\varphi\right] \tag{16}$$

Figure 2 gives a representation of this relation in isotherms.

2 Thermal Stress in the Tube Wall

Thermal stress calculations can be performed by application of the temperature results reported in Chapter 1. This will be shown also for the case of a heat radiating semi-infinite body. To simplify the notations the geometrical data are normalized to the inner radius r by setting $r = 1$.

Axial strains and deflections caused by bending moments should be prevented, i.e. $\epsilon_z = 0$. The elasticity equations are

$$\sigma_\rho = \frac{E}{(1+\mu)(1-2\mu)}[(1-\mu)\epsilon_\rho + \mu\epsilon_\varphi - (1+\mu)\alpha T]$$

$$\sigma_\varphi = \frac{E}{(1+\mu)(1-2\mu)}[(1-\mu)\epsilon_\varphi + \mu\epsilon_\rho - (1+\mu)\alpha T]$$

$$\sigma_{\rho\varphi} = \frac{E}{1+\mu}\,\epsilon_{\rho\varphi} \tag{17}$$

where $\sigma_{\rho\varphi}$ and $\epsilon_{\rho\varphi}$ are the shear stress and strain, μ is Poisson's ratio, α the coefficient of expansion, and E the Young's modulus.

Two equilibrium conditions are given by

$$\frac{\partial\sigma_\rho}{\partial\rho} + \frac{1}{\rho}\frac{\partial\sigma}{\partial\varphi} + \frac{\sigma_\rho - \sigma_\varphi}{\rho} = 0$$

$$\frac{\partial\sigma_{\rho\varphi}}{\partial\rho} + \frac{1}{\rho}\frac{\partial\sigma_\varphi}{\partial\varphi} + \frac{2\sigma_{\rho\varphi}}{\rho} = 0 \tag{18}$$

The strain components can be expressed by the displacements (u, v) in cylindric coordinates

$$\epsilon_\rho = \frac{\partial u}{\partial\rho};\quad \epsilon_\varphi = \frac{1}{\rho}\frac{\partial v}{\partial\varphi} + \frac{u}{\rho};$$

$$\epsilon_{\rho\varphi} = \frac{1}{2}\left(\frac{1}{\rho}\frac{\partial u}{\partial\varphi} + \frac{\partial v}{\partial\rho} - \frac{v}{\rho}\right) \tag{19}$$

Since the temperature distribution is given by a Fourier series it can be concluded that the displacements are

$$u = \sum_n u_n(\rho)\cos n\varphi;\quad v = \sum_n v_n(\rho)\sin n\varphi \tag{20}$$

By introducing equation (20) into equations (17–19), one obtains a system of two coupled differential equations

$$v_n''\rho + v_n'\rho - \left[1 + \frac{2(1-\mu)n^2}{1-2\mu}\right]v_n - \frac{n}{1-2\mu}\,\rho u_n' - n\,\frac{3-4\mu}{1-2\mu}\,u_n$$

$$= -\frac{1+\mu}{1-2\mu}\,2\alpha\rho n T_n$$

$$u_n''\rho + u_n'\rho - \left[1 + \frac{(1-2\mu)n^2}{2(1-\mu)}\right]u_n + \frac{n}{2(1-\mu)}\,\rho v_n'$$

$$-n\,\frac{3-4\mu}{2(1\text{-}\mu)}\,v_n = \frac{1+\mu}{1-\mu}\,\alpha\rho^2 T_n' \tag{21}$$

The solution of this system will yield the displacements and using equations (17) and (19), the related stresses in the tube cross-section. Application of the general set-up.

$$v = A\rho^k \quad u = B\rho^k \tag{22}$$

gives the solution of the homogeneous form of equation (21)

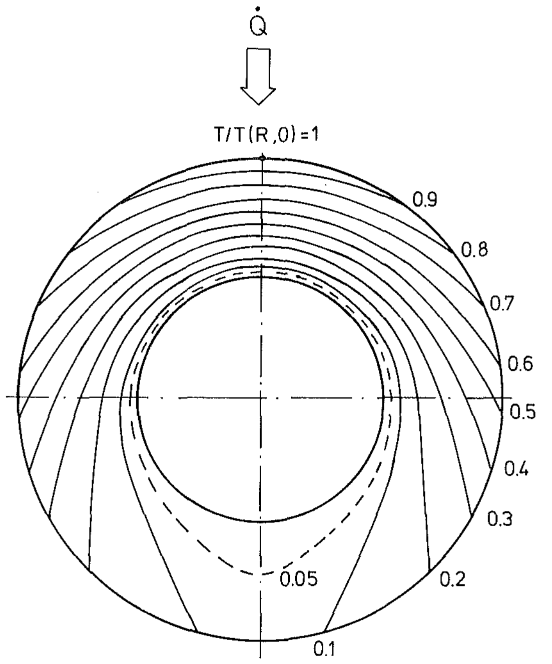


Fig. 2 Isotherms in case of a heat radiating half-space

$$u = A_1\rho + A_2\rho^{-1} \qquad \text{for } n = 0 \tag{23}$$
$$v = 0$$

and

$$u = A_1\frac{1-4\mu}{5-4\mu}\,\rho^2 + A_2\rho^{-2} + \left[\frac{A_4}{3-4\mu} - A_3\right) - A_4\ln\rho$$

$$\text{for } n = 1 \tag{24}$$

$$v = A_1\rho^2 + A_2\rho^{-2} + A_3 + A_4 ln\rho$$

A particular solution of the non-homogeneous system can be evaluated by the method of "variation of parameters".

For $n = 0$ only a single differential equation follows. Its solution is

$$u_i = \alpha\frac{1+\mu}{1-\mu}\,\frac{1}{\rho}\int_1^\rho T\xi d\xi \tag{25}$$

For $n = 1$ the respective displacements are given by

$$u_i = \frac{\alpha}{2}\,\frac{1+\mu}{1-\mu}\left[\frac{1}{\rho^2}\int_1^\rho x^2 T\,dx + \int_1^\rho T\,dx\right.$$

$$\left. -\frac{1}{8}T'(1)\left(\frac{1}{\rho^2} + \frac{4}{3-4\mu} - 4ln\rho - \frac{1-4\mu}{5-4\mu}\rho^2\right)\right]$$

$$v_i = \frac{\alpha}{2}\,\frac{1+\mu}{1-\mu}\left[\frac{1}{\rho^2}\int_1^\rho x^2 T\,dx - \int_1^\rho T\,dx\right.$$

$$\left. +\frac{1}{8}T'(1)\left(\rho^2 - \frac{1}{\rho^2} - 4\,ln\rho\right)\right] \tag{26}$$

The unknown constants A_1, A_2, A_3, A_4 have to satisfy the boundary conditions at the outer and inner surfaces

$$\sigma_\rho(R) = \sigma_\rho(1) = \sigma_{\rho\varphi}(R) = \sigma_{\rho\varphi}(1) = 0 \tag{27}$$

For $n = 0$ one can find

$$\sigma_{\rho 0} = \frac{\alpha E}{1-\mu}\left[\frac{\rho^2 - 1}{R^2 - 1}\,\frac{1}{\rho^2}\int_1^R T\xi d\xi - \frac{1}{\rho^2}\int_1^\rho T\xi d\xi\right];$$

$$\sigma_{\rho\varphi_0} = 0$$

$$\sigma_{\varphi 0} = \frac{\alpha E}{1-\mu}\left[\frac{\rho^2 + 1}{R^2 - 1}\,\frac{1}{\rho^2}\int_1^R T\xi d\xi + \frac{1}{\rho^2}\int_1^\rho T\xi d\xi - T\right] \tag{28}$$

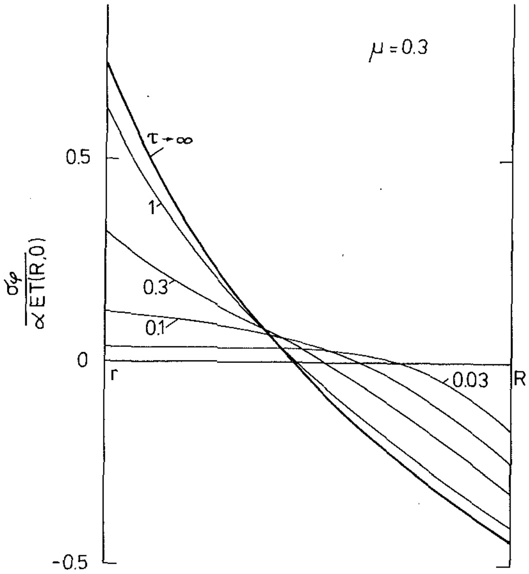


Fig. 3 Instationary circumferential stresses in a tube in front of a heat radiating half-space

and for $n = 1$ the following expressions hold

$$\sigma_{\rho 1} = \frac{\alpha E}{1+\mu}\left(\frac{2}{5-4\mu} A_1\rho - 2A_2\rho^{-3} - \frac{3-2\mu}{3-4\mu} A_4\rho^{-1}\right)$$
$$+ \frac{\alpha E}{1-\mu}\left[\frac{1}{8}\frac{\partial T}{\partial \rho}(1)\left(\rho^{-3} + \frac{6-4\mu}{3-4\mu}\rho^{-1} + \frac{1}{5-4\mu}\rho\right)\right.$$
$$\left. - \frac{1}{\rho^3}\int_1^\rho \xi^2 T\, d\xi\right]$$

$$\sigma_{\rho\varphi 1} = \frac{\alpha E}{1+\mu}\left(\frac{2}{5-4\mu} A_1\rho - 2A_2\rho^{-3} - \frac{1-2\mu}{3-4\mu} A_4\rho^{-1}\right)$$
$$+ \frac{\alpha E}{1-\mu}\left[\frac{1}{8}\frac{\partial T}{\partial \rho}(1)\left(\rho^{-3} + \frac{6-4\mu}{3-4\mu}\rho^{-1} + \frac{1}{5-4\mu}\rho\right)\right.$$
$$\left. - \frac{1}{\rho^3}\int_1^\rho \xi^2 T\, d\xi\right] \tag{29}$$

$$\sigma_{\varphi 1} = \frac{\alpha E}{1+\mu}\left(\frac{6}{5-4\mu} A_1\rho + 2A_2\rho^{-3} + \frac{1-2\mu}{3-4\mu} A_4\rho^{-1}\right)$$
$$+ \frac{\alpha E}{1-\mu}\left[\frac{1}{8}\frac{\partial T}{\partial \rho}(1)\left(\frac{3}{5-4\mu}\rho - \rho^{-3} - \frac{2-4\mu}{3-4\mu}\rho^{-1}\right)\right.$$
$$\left. + \frac{1}{\rho^3}\int_1^\rho \xi^2 T d\xi\right] - \alpha\frac{ET}{1-\mu}$$

with

$$A_1 = -\frac{1}{2}\frac{1+\mu}{1-\mu}\left[\frac{1}{8}\frac{\partial T}{\partial \rho}(1) - (5-4\mu)\frac{\int_1^R \xi^2 T d\xi}{R^4 - 1}\right]$$

$$A_2 = \frac{1}{2}\frac{1+\mu}{1-\mu}\left[\frac{1}{8}\frac{\partial T}{\partial \rho}(1) + \frac{\int_1^R \xi^2 T d\xi}{R^4 - 1}\right]$$

$$A_4 = \frac{1}{4}\frac{1+\mu}{1-\mu}\frac{\partial T}{\partial \rho}(1)$$

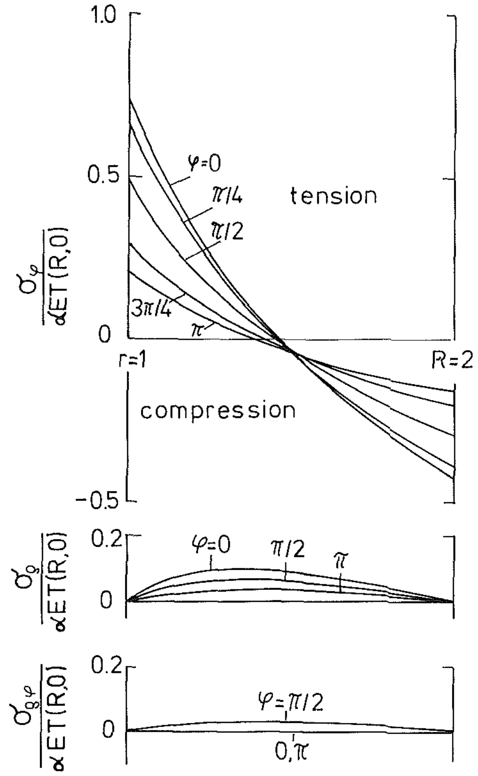


Fig 4 Stationary stress distributions for different values of φ

Taking into account the first terms of equations (11) and (12) the stationary stress components can be written after dropping the restriction $r = 1$ (i.e., $R \to R/r$ and $\rho \to \rho/r$) as

$$\sigma_{\rho 0} = \frac{1}{4}\frac{\alpha E}{1-\mu}\Lambda \dot{Q}_0 R\left[\frac{\rho^2 - r^2}{R^2 - r^2}\frac{1}{\rho^2}\left(2R^2 ln\frac{R}{r}\right.\right.$$
$$\left.\left. - R^2 + r^2\right) - 2\, ln\,\frac{\rho}{r} - \frac{r^2}{\rho^2} + 1\right]$$

$$\sigma_{\varphi 0} = \frac{1}{4}\frac{\alpha E}{1-\mu}\Lambda \dot{Q}_0 R\left[\frac{\rho^2 + r^2}{R^2 - r^2}\frac{1}{\rho^2}\left(2R^2 ln\frac{R}{r}\right.\right.$$
$$\left.\left. - R^2 + r^2\right) - 2\ln\frac{\rho}{r} + \frac{r^2}{\rho^2} - 1\right] \tag{30}$$

$$\sigma_{\rho\varphi_0} = 0; \quad \sigma_{z_0} = \mu(\sigma_{\rho 0} - \sigma_{\varphi 0}) - \alpha E \Lambda \dot{Q}_0 R\; ln\,\frac{\rho}{r}$$

and

$$\sigma_{\rho 1} = \frac{1}{2}\frac{\alpha E}{1-\mu}\dot{Q}_1 \Lambda R\left[\frac{r}{\rho} - \frac{\rho r}{r^2 + R^2} - \frac{R^2}{R^2 + r^2}(r/\rho)^3\right]$$

$$\sigma_{\varphi 1} = \frac{1}{2}\frac{\alpha E}{1-\mu}\dot{Q}_1 \Lambda R\left[\frac{r}{\rho} - 3\frac{\rho r}{r^2 + R^2} + \frac{R^2}{R^2 + r^2}(r/\rho)^3\right]$$

$$\sigma_{\rho\varphi_1} = \sigma_{\rho_1} \tag{31}$$

$$\sigma_{z_1} = \mu(\sigma_{\rho 1} + \sigma_{\varphi 1}) - \alpha E \Lambda \dot{Q}_1 R\left(\frac{\rho}{r} - \frac{r}{\rho}\right) \Big/ \left(\frac{R}{r} + \frac{r}{R}\right)$$

The stresses are then given by

$$\sigma_\rho = \sigma_{\rho_0} + \sigma_{\rho_1}\cos\varphi$$
$$\sigma_\varphi = \sigma_{\varphi_0} + \sigma_{\varphi_1}\cos\varphi$$
$$\sigma_{\rho\varphi} = \sigma_{\rho\varphi_0} + \sigma_{\rho\varphi_1}\sin\varphi \tag{32}$$

By use of equations (11), (12), (28) and (29) the transient stresses can be evaluated. Figure 3 shows the circumferential stresses σ_φ as a function of the normalized time for a thick-walled tube ($R=2r$). Maximum stress values are reached in the stationary case. Figure 4 represents the stationary stress distributions for σ_φ, σ_ρ, $\sigma_{\rho\varphi}$ and different values of φ. It is well known that temperature components T_n with $n > 1$ have no influence on stationary stresses in tubes [7]. Thus it is possible to compose all stationary stress distributions by use of the components with $n=0$ and $n=1$.

3 Summary

The Fourier components of the instationary temperature in an asymmetrically radiation heated tube have been evaluated. As a result the temperature can be composed for an arbitrary heat flux distribution. The situation of the heat radiating half-space was outlined. From the temperature distributions related thermal stresses are calculated. It was found that maximum stresses occur in the stationary case.

References

1 Carslaw, H. S., and Jaeger, J. C., *Conduction of Heat in Solids*, Oxford University Press, 1959.

2 Holms, A. G., "A Biharmonic Relaxation Method for Calculating Thermal Stress in Cooled Irregular Cylinders," NACA-TN-2434, 1951.

3 Kettleborough, C. F., "Non-Axial-Symmetric Thermal Stresses in Circular Discs or Cylinders," *Int. Journ. for Num. Methods in Engineering*, Vol. 3, 1971, pp. 53–61.

4 Chern, J. M., and Pai, D. H., "Inelastic Analysis of a Straight Tube Under Combined Bending, Pressure and Thermal Loads," *Journ. of Pressure Vessel Technology*, Vol. 97, 1975, pp. 155–162.

5 Gupta, G. D., Rao, M. S. M., Narayanan, T. V., and Gangadharan, A. C., "Thermoelastic Analysis of Nonaxisymmetrically Heated Thick Cylindrical Shells," *J. of Pressure Vessel Technology*, Vol. 100, 1978, pp. 107–111.

6 Fett, T., and Munz, D., "Stress and Lifetime Calculations for First Wall and Blanket Structual Components," Part I, KfK 3875, Kernforschungszentrum Karlsruhe, Jan. 1985.

7 Melan, E., and Parkus, H., *Warmespannungen infolge stationarer Temperaturfelder*, Springerverlag, Vienna, 1953.

J. L. Nowinski
H. Fletcher Brown Professor Emeritus,
Department of Mechanical
and Aerospace Engineering,
University of Delaware,
Newark, Del. 19716

On the Wave Propagation in an Elastic Hollow Cylinder With Long-Range Cohesion Forces

After a brief derivation of the formula for the nonlocal moduli, Fourier transforms of the stress components in their nonlocal aspect are established. Satisfaction of the traction-free boundary conditions leads to the frequency equation of the problem. A particular case involving longitudinal Lamé modes is analyzed in more detail. A numerical example solved shows a considerable decrease of the speed and the frequency of the short waves as compared with those of long waves studied in the classical theory.

Introduction

Wave propagation in cylindrical rods of circular cross section is, as is well known, one of the classical subjects of elastodynamics, initiated and rigorously treated in the papers of Pochhammer and Chree about a hundred years ago. Since then, various facets of problem were examined in numerous publications, the reviews of which may be found in the books and articles of Kolsky [1], Ewing, Jardetsky and Press [2], Abramson, Plass and Ripperger [3], Miklovitz [4, 5], and Achenbach [6], to name only a few. All of this work was done within the framework of the conventional (whether linear or nonlinear) theory of elasticity, one of whose principal postulates states that the interactions of particles in solids represent contact forces with the range of action limited to zero.

Unlike the conventional (local) theory, the nonlocal theory of elasticity, or more broadly the nonlocal continuum theory of deformable bodies, developed in the third quarter of this century[1], asserts that the cohesion forces are long-range forces, and on account of this the thermomechanical state at a point of the body is influenced by all of the particles of the body. More explicitly this means that the stress at the observation point **X**, which in the classical elasticity is a function of the deformation at **X**, in the nonlocal theory becomes a functional of the deformation field at every point **X**′ of the body. This standpoint is not as academic as it may seem at the first sight, inasmuch as according to the experimental evidence interactions of particles reach occasionally their fifteenth closest neighbors. In addition to that, solutions of problems treated in the context of the non-local elasticity display occasionally an impressive agreement with the data of experiments and observations. As illustrations of this fact, the following may serve as a few examples:

1 Contrary to the predictions of the classical theory of elasticity, and in agreement with experimental evidence, the nonlocal theory concludes that elastic waves in unbounded media are dispersive. This new kind of dispersion has a strictly constitutive character and was found for plane waves in an infinite medium [8], for the Rayleigh waves [21], and for the Love waves [11].

2 Likewise, the stress singularities at the tips of the Griffith cracks predicted by the local elastic theory (as a result of the inadequacy of this theory to describe physical phenomena on a submicroscopic scale) become removed in the nonlocal theory, and replaced by regular stress distributions. The latter display a striking agreement with the conclusions of the atomistic theory [22].

3 Similar corrections are suggested by the nonlocal theory with regard to the classical singularities at the cores of dislocations. Here not only the infinite stress concentration is smoothed out, but also some interesting byproducts are obtained like, for example, the theoretical shear strength of the materials that compares favorably with the data of experiments [23].

4 In the case of seismic waves, the nonlocal elasticity predicts a lower bound for the speed of Love waves that agrees better with the seismological observation than the bound inferred from the classical theory [11].

5 So far as the fluid dynamics is concerned, the secondary flow pattern in rectangular pipes found experimentally agrees well qualitatively with the pattern calculated on the basis of the nonlocal dynamics of a viscous flow [24].

6 Finally, it is worth noting that the nonlocal theory of elasticity serves as a rather natural link between the classical local elasticity and the atomistic theory of lattice dynamics of Born and Von Kármán. In fact, by bringing the equations of nonlocal elasticity into a discrete form, one arrives at the equations of the lattice dynamics (cf., e.g., [18], equation (2.5)).

---

[1] Mostly due to the efforts of Edelen, Eringen, Kroener and Kunin (cf., e.g., [7] through [10]).

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

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

Although a final verdict on the usefulness of the nonlocal elasticity may be brought in only after further theoretical and experimental studies, this author – prompted by the foregoing and similar facts – examined recently the nonlocal aspects of wave propagation in a solid circular cylinder [12]. This problem is, in a certain sense, a special case of the problem analyzed in the present paper but with different boundary conditions. It was found in [12] that in addition to the known configurational dispersion, the non-local theory predicts a second type of dispersion associated with the constitution of the matter and determined by the constitutive equations.

The present note intends to extend the study [12] to *hollow* circular cylinders. To save on space, some calculations common to both cases are not reproduced in the present text. When the need arises, however, we refer to the appropriate place in [12].

With respect to the hollow cylinders, the conventional features of the wave propagation in such cylinders were examined in the comprehensive articles by Greenspon [13], Gazis [14], [15], Bird et al. [16], and Armenakas et al. [17],[2] among others.

We first recall the derivation of the dispersion equation in the case of longitudinal waves in an infinite space that leads to the formula for the nonlocal moduli in terms of the wave number. We then derive the equations for the Fourier transforms of the stress components for a hollow cylinder with nonlocal material characteristics. Satisfaction of the boundary conditions prescribing traction-free surfaces of the cylinder leads to the heavily transcendental frequency equation of the problem.

Two particular cases are considered, one of which involving Lamé's longitudinal modes is treated in more detail. A numerical example, showing a considerable decrease of the wave speed and frequency of short waves as compared with those of long waves studied in the classical theory, is illustrated graphically.

1 Nonlocal Moduli

In order to determine the elastic nonlocal moduli, we imagine a plane longitudinal wave to propagate in an infinite elastic nonlocal medium. Since the nonlocality of the medium brings about a (constitutional) dispersion of the wave, the associated dispersion equation may be compared with the corresponding equation derived on the basis of the Born-von Kármán lattice dynamics. A longer calculation (see [12], pp. 191–192) yields

$$\frac{2\bar{\mu}'(k)+\bar{\lambda}'(k)}{2\mu+\lambda}=\frac{1}{\sqrt{2\pi}}\left[\frac{\sin^2\dfrac{ka}{2}}{\left(\dfrac{ka}{2}\right)^2}-1\right], \tag{1.1}$$

where μ and λ are Lamé's constant, $\bar{\mu}'(k)$ and $\bar{\lambda}'(k)$ the Fourier transforms of the nonlocal moduli, k the wave number, and a the interatomic spacing. The foregoing equation rests on the assumption that the interactions of particles are limited to the closest-neighbor atomic planes (cf., [18], equation (1.3)).

2 A Hollow Circular Cylinder

Let us now consider an infinitely long hollow circular cylinder whose inner and outer radii are c and d, respectively. The cylinder is referred to a cylindrical r, ϕ, z coordinate system, and a plane longitudinal wave is transmitted in the z-direction along the axis of the cylinder. Referring for details of calculations to the paper [12], pp. 192–194, we note that in the Kroener-Eringen representation the nonlocal constitutive equations become:

$$\tau_{rr}=\lambda\theta+2\theta+2\mu\frac{\partial u}{\partial r}+\int_V\Big[\lambda'(|\mathbf{r}'-\mathbf{r}|,$$

$$|z'-z|)\theta'+2\mu'(|\mathbf{r}'-\mathbf{r}|,\ |z'-z|)\frac{\partial u'}{\partial r'}\Big]dv',$$

$$\tau_{\phi\phi}=\lambda\theta+2\mu\frac{u}{r}+\int_V\Big[\lambda'(|\mathbf{r}'-\mathbf{r}|,$$

$$|z'-z|)\theta'+2\mu'(|\mathbf{r}'-\mathbf{r}|,\ |z'-z|)\frac{u'}{r'}\Big]dv',$$

$$\tau_{zz}=\lambda\theta+2\mu\frac{\partial w}{\partial z}+\int_V\Big[\lambda'(|\mathbf{r}'-\mathbf{r}|,$$

$$|z'-z|)\theta'+2\mu'(|\mathbf{r}'-\mathbf{r}|,\ |z'-z|)\frac{\partial w'}{\partial z'}\Big]dv',$$

$$\tau_{rz}=\mu\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial r}\right)+\int_V\mu'(|\mathbf{r}'-\mathbf{r}|,$$

$$|z'-z|)\left(\frac{\partial u'}{\partial z'}+\frac{w'}{r'}\right)dv'. \tag{2.1}$$

where V is the volume of the rod, $\mathbf{r}$ the position vector, $u'\equiv u'(r',z';t)$ and $w'\equiv w'(r',\ z';\ t)$ the displacement components, and θ the dilatation. We subject the just-written equations to the double Fourier transformation (designated by the overbar) and obtain,

$$\begin{aligned}\bar{\tau}_{rr}&=a_1\bar{u}_{,r}+a_2\frac{\bar{u}}{r}-ika_2\bar{w},\\ \bar{\tau}_{\phi\phi}&=a_1\frac{\bar{u}}{r}+a_2\bar{u}_{,r}-ika_2\bar{w},\end{aligned} \tag{2.2.1}$$

$$\bar{\tau}_{zz}=a_2\left(\bar{u}_{,r}+\frac{\bar{u}}{r}\right)-ika_1\bar{w}, \tag{2.2.2}$$

$$\bar{\tau}_{rz}=a_3(\bar{w}_{,r}-ik\bar{u}), \tag{2.2.3}$$

with

$$\begin{aligned}a_1&=\lambda+2\mu+(2\pi)^{1/2}(\bar{\lambda}'+2\bar{\mu}'),\\ a_2&=\lambda+(2\pi)^{1/2}\bar{\lambda}',\\ a_3&=\mu+(2\pi)^{1/2}\bar{\mu}'.\end{aligned} \tag{2.3}$$

As shown in detail in Section 2 in [12], pp. 194–196, the governing equations of the problem are

$$\bar{\theta}_{,rr}+\frac{1}{r}\bar{\theta}_{,r}+\gamma_1^2\bar{\theta}=0, \tag{2.4.1}$$

$$\bar{\Omega}_{,rr}+\frac{1}{r}\bar{\Omega}_{,r}+\left(\gamma_2^2-\frac{1}{r^2}\right)\bar{\Omega}=0, \tag{2.4.2}$$

where

$$\bar{\theta}=\bar{u}_{,r}+\frac{\bar{u}}{r}-ik\bar{w}, \tag{2.5.1}$$

$$\bar{\Omega}=-(ik\bar{u}+\bar{w}_{,r}), \tag{2.5.2}$$

$\bar{u}$ and $\bar{w}$ are Fourier transforms of the radial and longitudinal displacements, respectively, and

$$\gamma_1^2=\frac{\rho\omega^2}{a_1}-k^2,\ \gamma_2^2=\frac{\rho\omega^2}{a_3}-k^2, \tag{2.5.3}$$

A lengthy but standard calculation leads to the following solution of the system (2.4.1) and (2.4.2):

---

[2] Mainly from the point of view of shells.

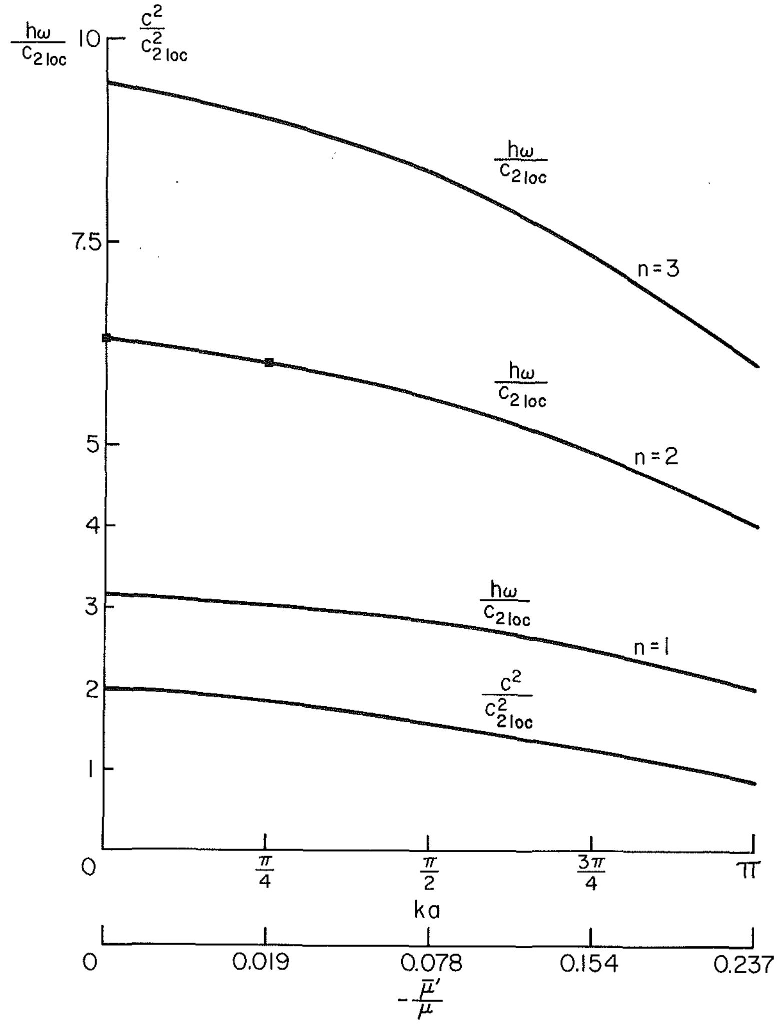


Fig. 1 Nondimensional frequency and wave speed versus the nondimensional wave number and the nondimensional nonlocal modulus $\bar{\mu}'/\mu$.

$$\bar{u}=C_1J_1(\gamma_1 r)+C_2Y_1(\gamma_1 r) + C_3J_1(\gamma_2 r)+C_4Y_1(\gamma_2 r), \tag{2.6.1}$$

$$\bar{w}=C_1\frac{ik}{\gamma_1}J_o(\gamma_1 r)+C_2\frac{ik}{\gamma_1}Y_o(\gamma_1 r) - C_3\frac{i\gamma_2}{k}J_o(\gamma_2 r)-C_4\frac{i\gamma_2}{k}Y_o(\gamma_2 r), \tag{2.6.2}$$

where J_i, Y_i, $i=0$, 1, are Bessel functions of the first and second kind, respectively. Substitution of the above equations into the equations of the stress components relevant to the present investigation gives,

$$\bar{\tau}_{rr}=C_1\left[a_4J_o(\gamma_1 r)-2a_3\frac{J_1(\gamma_1 r)}{r}\right] + C_2\left[a_4Y_o(\gamma_1 r)-2a_3\frac{Y_1(\gamma_1 r)}{r}\right]+2a_3C_3\left[\gamma_2 J_o(\gamma_2 r) - \frac{J_1(\gamma_2 r)}{r}\right]+2a_3C_4\left[\gamma_2 Y_o(\gamma_2 r)-\frac{Y_i(\gamma_2 r)}{r}\right], \tag{2.7}$$

$$\bar{\tau}_{rz}=ia_3[-2kC_1J_1(\gamma_1 r)-2kC_2Y_1(\gamma_1 r) + a_5C_3J_1(\gamma_2 r)+a_5C_4Y_1(\gamma_2 r)] \tag{2.8}$$

where

$$a_4=a_1\gamma_1+a_2\frac{k^2}{\gamma_1},\ a_5=\frac{\gamma_2^2}{k}-k. \tag{2.9}$$

The assumption that the inner and outer surfaces of the cylinder are free from external tractions, that is, in symbols,

$$\begin{aligned}\bar{\tau}_{rr}(c)&=\bar{\tau}_{rr}(d)=0,\\ \bar{\tau}_{rz}(c)&=\bar{\tau}_{rz}(d)=0,\end{aligned} \tag{2.10}$$

furnishes a system of four homogeneous linear algebraic equations for the four unknown coefficients C_i, $i=1, 2, 3, 4$. A nontrivial solution of the system requires, naturally, the vanishing of the principal determinant of the system.

A general solution of the associated frequency equation, which even in the classical case requires a comprehensive study (cf., e.g., [14]), is not of our primary interest here. Instead, we wish to confine our attention to a brief examination of two particular cases of the problem. The first of these is rather trivial; the second one permits us to make some observations as to the impact of the nonlocal treatment on the dispersion of a certain type of wave. (*a*) We first note that if the inner radius of the cylinder $c\to 0$, then in order for the stresses to remain bounded at $r=0$, one has to set $C_2=C_4=0$. The equations (2.7) and (2.8) then (after an appropriate translation of notation) reduce to the equation (2.12) derived directly in [12] for a solid cylinder. A more detailed analysis of this case is found in [12]. (*b*) The second case deals with longitudinal waves known as the equivoluminal Lamé modes ([19], p. 795). These waves are associated with the value of the Lamé constant $\lambda = 0$[3] and $\gamma_2^2 = k^2$. Combining this with the second pair of the boundary conditions (2.10), we find that $a_2 = a_5 = 0$, $a_1 = 2a_3$, $\rho\omega^2/a_1 = k^2$, $\gamma_1 = 0$, $C_2 = 0$, and $\bar{\tau}_{rz} = 0$. The remaining two of the boundary conditions now provide the characteristic equation of the problem,

$$J_1'(\gamma_2 c) Y_1'(\gamma_2 d) - J_1'(\gamma_2 d) Y_1'(\gamma_2 c) = 0. \tag{2.11}$$

The foregoing transcendental equation has the same form as the one derived by Gazis ([14], equation (34)) in the framework of the classical elasticity; the only difference between the equations lies in the presence of the nonlocal modulus $\bar{\mu}'$ in the coefficient a_1. Proceeding further, we find from the relation $\rho\omega^2/a_1 = k^2$ (after appeal to the equation (1.1)) for very long waves $c_{\text{nonloc}}/c_{2\text{ loc}} = 1.41$, and for admissibly shortest[4] waves $c_{\text{nonloc}}/c_{2\text{ loc}} = 0.91$, where c_{nonloc} is the actual wave speed and $c_{2\text{ loc}}$ is the speed of conventional equivoluminal waves.

For a thin cylindrical shell for which $h/c << 1$, we may assume that $\gamma_2 c >> 1$ and, consequently, $\gamma_2 d >> 1$ where $d = h + c$. With this in mind, we are permitted to approximate the development of the Bessel functions appearing in (2.11) by the first two terms of their Hankel asymptotic series (cf., e.g., [20], formulas 808.4 and 812.4). A lengthy calculation leads to the equation

$$\sin\gamma_2 h - \frac{7\gamma_2 h}{8\gamma_2^2 cd} \cos \gamma_2 h = 0, \tag{2.12}$$

which, after assuming that $\gamma_2 h = n\pi + \epsilon$ and $\epsilon << 1$, $n = 1, 2, 3, \ldots$, (cf., [15]) gives finally the value of the frequency ω in terms of the ratio h/c:

$$\omega = \sqrt{\frac{2a_3}{\rho}} \frac{n\pi}{h} \left[1 + \frac{\pi}{8(n\pi)^2} \left(\frac{h}{c}\right)^2\right], \qquad n = 1, 2, 3, \ldots, \tag{2.13}$$

where again

$$\frac{a_3}{\rho} = c_{2\text{ loc}}^2 \left(1 + \sqrt{2\pi} \frac{\bar{\mu}'}{\mu}\right). \tag{2.13a}$$

The graphs in Fig. 1 depict the nondimensional wave speed as well as the nondimensional wave frequency (the latter for the first three modes) as functions of the product ka (k = wave number, a = lattice constant) for the first Brillouin zone,[5] $0 \le ka \le \pi$. The value of the thickness h is assumed bo be equal 0.1c. It is seen that for the waves of the minimal length $2a$, the wave frequency drops to 0.64 of its value established in the conventional treatment. The same decrease is observed in the value of the wave speed.

3 Concluding Remarks

The following few remarks sum up the main differences between the local and the nonlocal aspects of the problem under discussion:

(*a*) Apart from the configurational dispersion, registered by the classical theory, and associated with the presence of the boundaries, the nonlocal theory discloses a new type of dispersion – the constitutive dispersion – distinct from the former, and associated with the internal structure, or constitution, of the matter.

(*b*) The nonlocal theory plays the role of a link between the classical elasticity and the atomistic approach: it converts into the conventional theory in the limit of the (infinitely) long waves, but turns into an atomic scale theory in the limit of the short waves and the discretization of the equations of motion.

(*c*) With regard to the wave velocity, in the range of short waves (beyond the reach of the classical theory), the nonlocal theory predicts the wave velocity less than that of the conventional long waves. In particular, in the case discussed in Section 2, for short waves the wave velocity (and the wave frequency) drops to 0.64 of its value established by the conventional theory.

Acknowledgment

This paper was prepared with the partial support of the University of Delaware.

References

1 Kolsky, H., *Stress Waves in Solids*, Oxford University Press, London, 1953.

2 Ewing, W. M., Jardetsky, W. S., and Press, F., *Elastic Waves in Layered Media*, McGraw-Hill, New York, 1957.

3 Abramson, H. N., Plass, H. J., and Ripperger, E. A., "Stress Propagation in Rods and Beams," *Advances of Applied Mechanics*, Vol. 5, 1958, pp. 111-194.

4 Miklovitz, J., "Elastic Wave Propagation," in *Applied Mechanics Surveys*, H. N. Abramson, H. Liebowitz, J. N. Crowley, S. Juhasz eds., Spartan Books, Washington, 1966, pp. 809-839.

5 Miklovitz, J., *Elastic Waves and Waveguides*, North-Holland, Amsterdam, 1980.

6 Achenbach, J. D., *Wave Propagation in Elastic Solids*, North-Holland, Amsterdam, 1976.

7 Edelen, D. G. B., "Nonlocal Field Theories," in *Continuum Physics*, Vol. 4, Eringen, A. C., ed., Academic Press, New York, 1976, pp. 75-204.

8 Eringen, A. C., "Nonlocal Polar Field Theories," *Continuum Physics*, Vol. 4, Eringen, A. C., ed., Academic Press, New York, 1976, pp. 205-267.

9 Kroener, E., "Continuum Mechanics and Range of Atomic Cohesion Forces," *Proceedings of International Conference on Fracture*, Vol. 1, Japanese Society of Fracture of Materials, Sendai, 1966.

10 Kumin, I. A., *Elastic Media with Microstructure*, Vols. 1 and 2, Springer-Verlag, Berlin, 1982 and 1983.

11 Nowinski, J. L., "On the Nonlocal Aspects of Propagation of Love Waves," *International Journal of Engineering Science*, Vol. 22, 1984, pp. 383-392.

12 Nowinski, J. L., "On the Nonlocal Theory of Longitudinal Waves in an Elastic Circular Bar," *Acta Mechanica*, Vol. 52, 1984, pp. 189-200.

13 Greenspon, J. E., "Flexural Vibrations of a Thick Walled Circular Cylinder," *Proceedings of the U.S. National Third Congress of Applied Mechanics*, ASME, New York, 1958, p. 163.

14 Gazis, D. C., "Three-Dimensional Investigation of the Propagation of Waves in a Hollow Circular Cylinder. Part I: Analytical Foundation," *The Journal of the Acoustical Society of America*, Vol. 31, 1959, pp. 568-573.

15 Gazis, D. C., "Exact Analysis of the Plane-Strain Vibrations of Thick-Walled Hollow Cylinders," *The Journal of the Acoustical Society of America*, Vol. 30, 1958, pp. 786-794.

16 Bird, J. F., Hart, R. W., and McClure, F. T., "Vibration of Thick-Walled Hollow Cylinder," *Journal of the Acoustical Society of America*, Vol. 32, 1961, p. 1404.

17 Armenakas, A. E., Gazis, D. C., and Hermann, G., *Free Vibrations of Circular Cylinder Shells*, Pergamon, Oxford, 1969.

18 Nowinski, J. L., "On the Nonlocal Effects in an Elastic Layer Subject to Gravity and Compression," *International Journal of Engineering Science*, Vol. 23, 1985, pp. 555-560.

19 Eringen, A. C., and Suhubi, E. S., *Elastodynamics*, Vol. 2, Academic Press, New York, 1975.

20 Dwight, H. B., *Tables of Integrals and other Mathematical Data*, MacMillan, New York, 1961.

21 Eringen, A. C., "On Rayleigh Surface Waves with Small Wave Lengths," *Letters in Applied and Engineering Sciences*, Vol. 1, 1973, pp. 11-17.

22 Ari, N., and Eringen, A. C., "Nonlocal Stress Field at Griffith Crack," *Crystal Lattice Defects and Amorphous Materials*, Vol. 10, 1983, pp. 33-38.

23 Eringen, A. C., "On Differential Equations of Nonlocal Elasticity and Solutions of Screw Dislocations and Surface Waves," *Journal of Applied Physics*, Vol. 54, No. 9, 1983, pp. 4703-4710.

24 Eringen, A. C., "Continuum Mechanics at the Atomic Seale," *Crystal Lattice Defects*, Vol. 7, 1977, pp. 109-130.

25 Brillouin, L., *Wave Propagation in Periodic Structures*, McGraw-Hill, New York, 1946.

---

[3] We extend this assumption to the nonlocal material, that is, we also set $\lambda' = 0$.

[4] As shown in lattice dynamic, the wave frequency being a function of the wave number, the wave length is not completely determined for the given frequency. This ambiguity is avoided if the wave length 1 is restricted to the interval $2a \le 1 \le \infty$, where a is the lattice parameter (cf., e.g., [23], p. 6).

[5] As noted in the preceding footnote, there is $2a \le 1 \le \infty$. Since $1 = 2\pi/k$ (k = wave number), then the ambiguity mentioned before disappears if $0 \le ka \le \pi$; the interval $[0, \pi]$ of the accepted variability of ka is known as the first Brillouin zone.

J. Takezaki

Y. Y. Hung

School of Engineering,
Oakland University,
Rochester, Mich. 48063

Direct Measurement of Flexural Strains in Plates by Shearography

Strains in a flexed plate are directly related to the second derivatives of the plate deflections. A technique is developed which enables shearography to measure second derivatives of deflection directly. The technique is based on generating fictitious carrier fringes in the form of periodic linear fringes of uniform spacing. The carrier fringes are distorted when the plate is deformed. The change in the pitch of the carrier fringes due to deformation yields the flexural strain directly. The pitch of the carrier fringes is equivalent to gage length.

Introduction

Plates are common structures used in engineering to resist bending loads. Experimental techniques are often used to measure strains in plates of complex geometry and loading conditions. Mechanical and electrical strain gages indicate strain at the two faces of the plate. From these strains, moments can be computed. However, this information can only be obtained on a point-by-point basis. Optical methods are more interesting since they can give whole-field information and no contact with the plate is required. These optical methods are classified into two categories: noncoherent light methods and coherent light methods. The noncoherent light methods are the moire method [1, 2] and the classical optical method [3]; the coherent light methods included holography [4, 5] and speckle interferometry [6]. The practical use of the moire method is limited by their deficiency in sensitivity. Holography, one the other hand, is too sensitive for many practical applications. Furthermore, holography measures deflections; therefore, it is necessary to differentiate the measured deflections twice [7, 8]. The work of differentiation is not only laborious but also creates a major source of error. Speckle photography [9, 10] and speckle-shearing interferometry [11–13], including shearography [14, 15], measure the slopes of deflections. Thus, one differentiation is still required to obtain the flexural strains.

Several methods have been suggested for obtaining the flexural strain directly [16–17]. However, all the proposed methods possess one or more severe limitations. Consequently, these methods demand complex experimental apparatus, complicated fringe interpretation, or apply only to special surfaces such as a mirror-like surface. Shearography seems to be a practical tool for the flexural strain measurement, as the experimental design is very simple, no surface treatment is needed and the fringe patterns produced are of high quality.

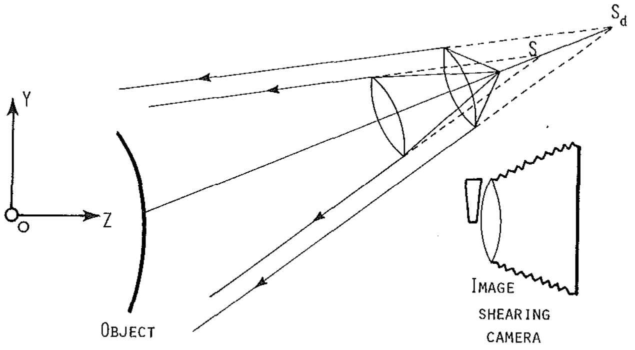


Fig. 1 Schematic for generating carrier fringes by shifting virtual light source

This paper describes a new technique of shearography that allows the second derivatives of deflections in plates to be measured directly.

Method

The experimental procedure is illustrated in Fig. 1. The object under study is illuminated by a point source of coherent light through a lens as shown, and it is imaged by an image-shearing camera. If a photographic plate at the image plane of the image-shearing camera is doubly exposed with the illuminating lens being translated along the illumination direction, a "fictitious fringe carrier" not due to the deformation is formed. The carrier fringes are in the form of linear fringes of equal spacing. If, in addition to the lens translation, the plate is deformed between exposures, a perturbed linear fringe carrier is obtained. The local perturbation depends on the local deflection gradients. The change of the spacing is directly related to the second derivatives of the plate deflections. Therefore, the flexural strains are directly determined by measuring the spacing of the fringe carrier.

Generation of Carrier Fringes

The linear carrier fringes are generated by translating the illumination lens. When a point light source is located inside

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami Beach, Fla., November 17–22, 1985 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. 10077, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November, 1983; final revision, March, 1985. Paper No. 85-WA/APM-18.

the focal length of the lens, a virtual image of the point source is formed. This is equivalent to a spherical wave diverging from the virtual point in [18]. If the lens is translated, the virtual point source is also translated. The translation of the virtual point source induces a quadratic phase variation of wavefront across the object's surface. The image-shearing camera measures derivatives of the phase variation, and the derivative of a quadratic phase change is a linear variation. Therefore, a set of linear fringes are produced by the lens translation and the fringe lines are perpendicular to the direction of shearing. Since this set of fringes is not due to deformation of object, it is referred to as "fictitious carrier fringes."

Analysis of Carrier Fringes

As illustrated in Fig. 1, shearing in the x-direction through the image-shearing camera is considered. The intensity of the total light field on the image plane, I_x has been obtained in the general form in [14, 15] as

$$I_x = 2a^2(1+\cos\phi_x) \tag{1}$$

where ϕ_x is a random phase angle. Equation (1) represents a random intensity variation pattern commonly known as speckle pattern.

When the light source position is shifted from S to S_d, as shown in Fig. 1, the relative phase shift Δ_x between the two sheared wave fronts induced by the relative optical path change δl (see the Appendix) can be obtained as

$$\Delta_x = \frac{2\pi}{\lambda}(\delta l) = \frac{2\pi}{\lambda}(k_1 x \delta x) \tag{2}$$

where λ is the wave length of a coherent light, $k_1 = (R_s - R_d)/R_s R_d$ (see Appendix) and δx is the magnitude of shearing. Therefore, with the translation of the light source, equation (1) is modified as

$$I_x' = 2a^2[1+\cos(\phi_x + \Delta_x)] \tag{3}$$

If a photographic film at the image plane is double exposed sequentially to I_x and I_x', the total energy recorded can be expressed as

$$\begin{aligned} I_T &= I_x + I_x' \\ &= 2a^2[2+\cos\phi_x + \cos(\phi_x + \Delta_x)] \end{aligned} \tag{4}$$

where the energy is recorded linearly. Equation (4) represents a frequency variable fringe; this type of fringes is not readily visible as explained in [14, 15]. However, equation (4) can be rewritten as

$$I_T = 4a^2\left[1 + \cos\frac{2\phi_x + \Delta_x}{2}\cos\frac{\Delta_x}{2}\right] \tag{5}$$

The second term in equation (5) represents the amplitude cos $(2\phi_x + \Delta_x/2)$ of high-frequency random carriers modulated by the low-frequency factor cos $\Delta_x/2$. Fringe lines are areas where the high-frequency carriers are nulled. This corresponds to

$$\cos\frac{\Delta_x}{2} = 0 \tag{6}$$

that is

$$\Delta_x = (2n+1)\pi = \frac{2\pi}{\lambda}(k_1 x)\delta x \quad (n = 0,1,2) \tag{7}$$

where n is the fringe order. Therefore, fringes are lines of equal Δ_x.

It should be emphasized that this type of fringe pattern is different from the conventional fringe patterns. Conventional fringes are loci at which intensity is minimal. The conventional fringes are readily visible to the eye, whereas the present frequency variation type fringes are not visible. However, these invisible fringes can be converted to the visible intensity variation type by an optical high-pass Fourier filtering process.

Equation (7) shows that the fringe pattern produced by the translation of the illumination source consists of periodic linear fringes. The fringe lines are perpendicular to the shearing direction x and their pitch is $\lambda/k_1\delta x$. This fringe pattern is referred to as "carrier fringes."

When the object is deformed in addition to the point light source translation, the relative phase shift Δ'_x induced by the relative optical path change for the deformation case (Appendix), is obtained as

$$\begin{aligned} \Delta_x' &= \frac{2\pi}{\lambda}(\delta l^*) \\ &= \frac{2\pi}{\lambda}\left(k_1 x + A\frac{\partial u}{\partial x} + B\frac{\partial v}{\partial x} + C\frac{\partial w}{\partial x}\right)\delta x \end{aligned} \tag{8}$$

where A, B, and C are sensitivity factors which can be varied (see Appendix). Equation (3), therefore, is modified as

$$I_x' = 2a^2[1+\cos(\phi_x + \Delta_x')] \tag{9}$$

If a photographic film at the image plane that is doubly exposed to an object before deformation and again after deformation with the light source position shifted, it records the total energy given by

$$\begin{aligned} I_T' &= I_x + I_x' \\ &= 2a^2[2+\cos\phi_x + \cos(\phi_x + \Delta_x')] \end{aligned} \tag{10}$$

Equation (10) is the same type as equation (4), thus we can rewrite equation (10) as

$$I_T' = 4a^2\left[1 + \cos\left(\phi_x + \frac{\Delta_x'}{2}\right)\cos\frac{\Delta_x'}{2}\right] \tag{11}$$

Again, fringe lines are areas where the high frequency carriers are nulled. That is

$$\cos\frac{\Delta_x'}{2} = 0 \tag{12}$$

or

$$\begin{aligned} \Delta_x' = (2n+1)\pi = \frac{2\pi}{\lambda}\Bigg(k_1 x + A\frac{\partial u}{\partial x} + B\frac{\partial v}{\partial x} \\ + C\frac{\partial w}{\partial x}\Bigg)\delta x (n = 0,1,2) \end{aligned} \tag{13}$$

Equation (13) represents a perturbed carrier fringe. The local perturbation is related to the second derivatives of displacement. For measuring flexural strains in plate structures in which the out-of-plane displacement dominates, equation (13) becomes

$$(2n+1)\pi = \frac{2\pi}{\lambda}\left(k_1 x + C\frac{\partial w}{\partial x}\right)\delta x \tag{14}$$

Direct Flexural Strain Measurement

Flexural strains in a flexed plate are related to the second derivatives of the plate deflection by

$$\epsilon_x = h\frac{\partial^2 w}{\partial x^2}, \epsilon_y = h\frac{\partial^2 w}{\partial y^2}, \gamma_{xy} = 2h\frac{\partial^2 w}{\partial x \partial y} \tag{15}$$

where h is the distance from the neutral plane to the plate surface.

The second derivatives of the plate deflection can be determined directly by measuring the change in pitch in the deformed carrier fringes.

If we consider the difference between the fringe order $i+1$ at position x_{i+1} and i at position x_i in equation (7), we can define the fringe pitch as follows:

$$p = x_{i+1} - x_i = \frac{\lambda}{k_1 \delta x} \tag{16}$$

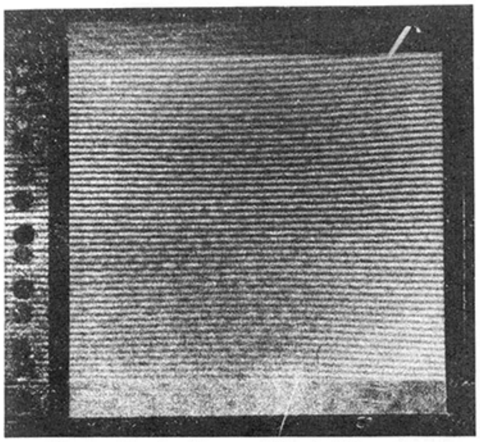


Fig. 2 Carrier fringes without deformation

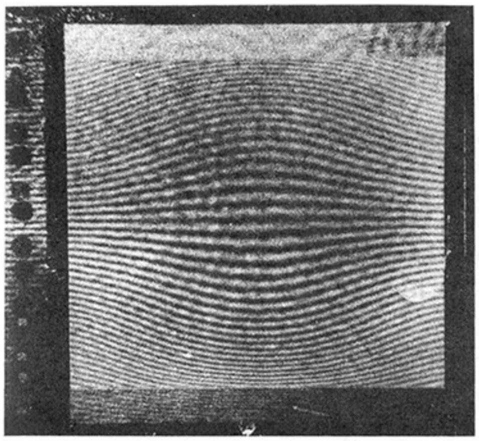


Fig. 3 Carrier fringes with deformation

In equation (16), λ, k_1, and δx are known constants. Therefore, the pitch of the carrier fringe for the case where there is no deformation can be obtained as the constant pitch in the direction of shearing.

Considering the carrier fringe positions x_i and x_{i+1} for the deformed carrier fringes, the fringe spacing p^* is defined as the same manner as equation (16). It can be written as

$$p^* = x_{i+1} - x_i = \frac{\lambda}{k_1 \delta x} - \frac{C}{k_1}\left(\frac{\partial w_{i+1}}{\partial x} - \frac{\partial w_i}{\partial x}\right) \quad (17)$$

For simplicity, let us assume that the object is illuminated and observed from a large distance and both along the z-direction. The sensitivity factors of equation (17) can be approximated by $C=2$, and equation (17) becomes

$$\left(\frac{\partial w_{i+1}}{\partial x} - \frac{\partial w_i}{\partial x}\right) = \frac{k_1}{2}\left(\frac{\lambda}{k_1 \delta x} - p^*\right) = \frac{k_1}{2}(p - p^*) \quad (18)$$

We can consider p^* as the small increment Δx, then

$$\left(\frac{\partial w_{i+1}}{\partial x} - \frac{\partial w_i}{\partial x}\right)/\Delta x = \frac{k_1}{2}\,\frac{(p-p^*)}{p^*} \quad (19)$$

If Δx is small, the right-hand side of the above equation approximates as $\partial^2 w/\partial x^2$, the second derivative of deflections.

Therefore

$$\frac{\partial^2 w}{\partial x^2} \simeq \left(\frac{\partial w_{i+1}}{\partial x} - \frac{\partial w_i}{\partial x}\right)/\Delta x \simeq -\frac{k_1}{2}\left(\frac{p^*-p}{p^*}\right) \quad (20)$$

Equation (20) shows that it is possible to obtain the second derivative of deflection by measuring the deformed carrier fringe pitch p^* and the undeformed carrier fringe pitch p.

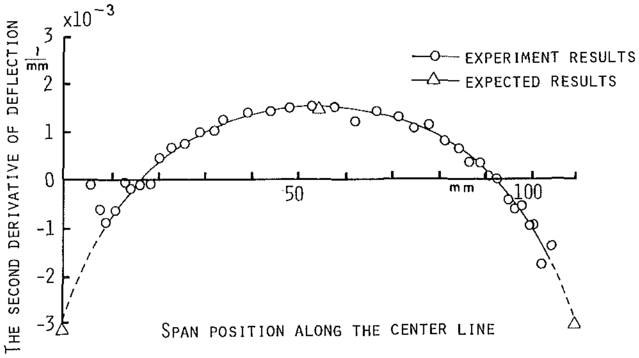


Fig. 4 The second derivative distribution of deflection along the centerline

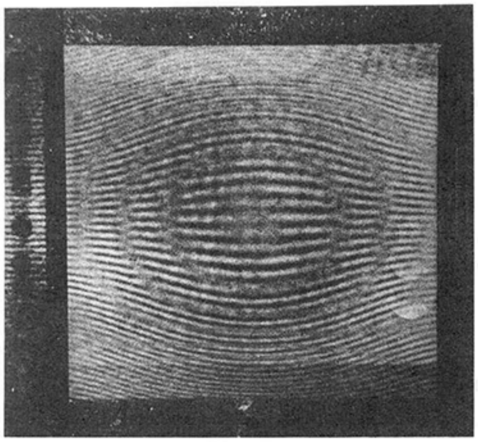


Fig. 5 The moire fringe pattern of the second derivative of deflection

Should the shearing be in the y-direction, equation (20) becomes

$$\frac{\partial^2 w}{\partial y^2} = -\frac{k_1}{2}\,\frac{(p^*-p)}{p^*} \quad (21)$$

To determine $\partial^2 w/\partial x \partial y$, shearing in a third direction is needed, say in a direction inclined at an angle α with the y-axis. This "rosette" information allows $\partial^2 w/\partial x \partial y$ to be calculated with the following equation

$$\frac{\partial^2 w}{\partial x_\alpha{}^2} = \frac{\partial^2 w}{\partial x^2}\sin^2\alpha + \frac{\partial^2 w}{\partial x \partial y}\sin 2\alpha + \frac{\partial^2 w}{\partial y^2}\cos^2\alpha \quad (22)$$

Experimental Demonstration

To verify the validity of the technique, a square plate made of aluminum and of dimension 110 mm × 110 mm × 2 mm was chosen for the study. The plate was built in along all edges and subjected to a uniform pressure of 3.45 KPa. An argon laser emitting light of wavelength $\lambda = 0.515 \mu m$ was employed as the coherent light source.

The aluminum plate was illuminated and observed at a large distance normal to the plate so that the sensitivity factor $A \simeq B \simeq 0$ and $C=2$.

Figure 2 shows an undeformed carrier fringe pattern obtained by translating the illuminating source between the two exposures. The fringes were made visible by a high-pass optical Fourier filtering technique. Figure 3 shows the deformed carrier fringes. In both fringe patterns shearing in the y-direction was used. $\partial^2 w/\partial y^2$ along the centerline of the plate was deduced with equation (21). The comparison of the experimental results with the theoretically expected results is shown in Fig. 4. Good agreement between the results of the two approaches is seen. Hence, the validity of the proposed technique is confirmed.

Another avenue of obtaining the second derivatives of deflections is to form a secondary moire fringe with two identical deformed carrier fringe patterns, one shifted relative to the other. Figure 5 shows a moire fringe pattern depicting $\partial^2 w/\partial y^2$ distribution in the square plate. Here, the carrier fringes enchance the visibility of the moire fringe pattern.

Conclusion

A technique employing shearography is developed that allows direct determination of the second derivatives of plate deflections and thus the flexural strains. Though the outputs of the technique are fringe pattern, it does not require identifying fringe orders. Instead, only the measurement of the carrier fringe spacing is needed to deduce the data. It should be pointed out that the differentiation process of shearography is a finite difference approach.

The removal of the need to identify fringe orders lets shearography acquire a high potential of being developed into a computer based automatic data reduction system.

Acknowledgment

This investigation was supported by the National Science Foundation (Grant CME 79-20097). The authors are grateful to Dr. C. C. Astill of NSF for his support.

References

1 Theocaris, P. S., *Moire Fringes in Strain Analysis*, Pergamon, New York, 1969.

2 Durelli, A. J., and Parks, V. J., *Moire Analysis of Strain*, Prentice Hall, 1970.

3 Ducan, J. P., and Sabin, P. G., "Determination of Curvatures in Flexed Elastic Plates by the Martinelli-Rouchi Technique," *Experimental Mechanics*, Vol. 3, No. 12, Dec. 1963, pp. 285-293.

4 *Holographic Nondestructive Testing*, Eff, R. J., ed., Academic Press, New York, 1974.

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

6 Edited by Erf, R. J., *Speckle Metrology*, Academic Press, New York, 1978.

7 Taylor, L. H., and Brandt, G. B., "An Error Analysis of Holographic Strains Determined by Cubic Splines," *Experimental Mechanics*, Vol. 12, No. 12, Dec. 1972, pp. 543-548.

8 Sollid, J. E., and Stetson, K. A., "Strains from Holographic Data," *Experimental Mechanics*, Vol. 18, No. 6, June 1978, pp. 208-214.

9 Chiang, F. P., and Juang, R. M., "Laser Speckle Interferometry for Plate Bending Problems," *Applied Optics*, Vol. 15, No. 9, Sept. 1976, pp. 2199-2204.

10 Hung, Y. Y., Daniel, I. M., and Rowlands, R. E., "Full-Field Optical Strain Measurement having Postrecording Sensitivity and Direction Selectivity," *Experimental Mechanics*, Vol. 18, No. 2, Feb. 1976, pp. 56-60.

11 Leendertz, J. A., and Butters, J. N., "An Image-Shearing Speckle Patterns Interferometer for Measuring Bending Moment," *J. of Physics E: Scientific Instruments*, Vol. 6, 1973, pp. 1107-1110.

12 Boon, D. M., "Determination of Slope and Strain Contours by Double-Exposure Shearing Interferometry," *Experimental Mechanics*, Vol. 15, No. 8, Aug. 1975, pp. 295-302.

13 Assa, A. Politch, J., and Bester, A. A., "Slope and Curvature Measurement by a Double Frequency-Grating Shearing Interferometer," *Experimental Mechanics*, Vol. 19, No. 4, Dec. 1979, pp. 129-137.

14 Hung, Y. Y., and Liang, C. Y., "Image-Shearing camera for Direct Measurement of Surface-Strains," *Applied Optics*, Vol. 10, No. 7, Apr. 1979, pp. 1046-1050.

15 Hung, Y. Y., "Shearography: A New Optical Method for Strain Measurement and Nondestructive Testing," *Optical Engineering*, Vol. 21, No. 3, May/June 1982, pp. 391-395.

16 Chiang, F. P., and Bailaugadi, M., "A Method for Direct Determination of Small Curvatures," *J. of Applied Mechanics*, Vol. 42, No. 1, 1978, pp. 29-31.

17 Chao, Y. J., Sutton, M. A., and Taylor, C. E., "Interferometric Methods for Measurement of Curvature and Twist in Thin Plate," *Proceeding of the 1982 Joint Conference on Experimental Mechanics*, Oahu-Maui, Hawaii, May 1982, pp. 514-519.

18 Goodman, J. W., *Introduction to Fourier-Optics*, McGraw-Hill, 1968.

19 Timoshenko, S. P., and Woinowsky-Krieger, S., *Theory of Plates and Shell*, McGraw-Hill, New York.

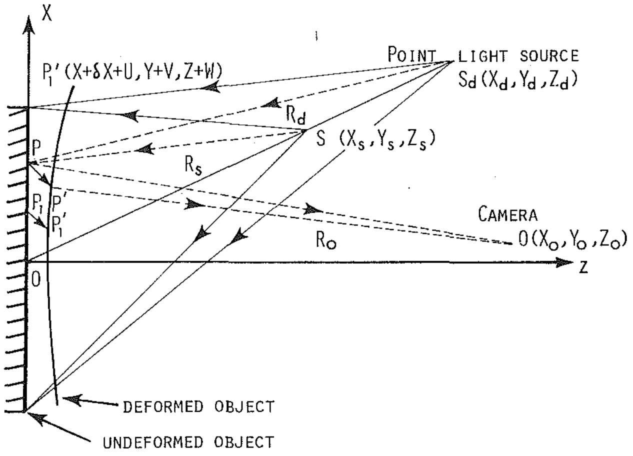


Fig. 6 Optical path diagram for analyzing carrier fringes

APPENDIX

Analysis of Carrier Fringes

The optical path difference for the ray travelling from the light source $S_d(x_d, y_d z_d)$ and $S(x_s, y_s, z_s)$ to the camera $O(x_0, y_0, z_0)$ via the point P on the object, is obtained as follows:

Before Deformation.

$$
\begin{aligned}
SP &= [(x_s-x)^2+(y_s-y)^2+(z_s-z)^2]^{1/2} \qquad \text{(A1)}\\
&\simeq R_s\left[1+\frac{x^2+y^2+z^2-2xx_s-2yy_s-2zz_s}{2R_s^2}\right]
\end{aligned}
$$

$$
\begin{aligned}
S_dP &= [(x_d-x)^2+(y_d-y)^2+(z_d-z)^2]^{1/2} \qquad \text{(A2)}\\
&\simeq R_d\left[1+\frac{x^2+y^2+z^2-2xx_d-2yy_d-2zz_d}{2R_d^2}\right]
\end{aligned}
$$

Where $R_s^2=x_s^2+y_s^2+z_s^2$ and $R_d^2=x_d^2+y_d^2+z_d^2$. The foregoing equations are obtained by neglecting high-order terms in the series expansion. Therefore, the optical path change l from S and S_D to the point on the object is

$$
\begin{aligned}
l(x,y,z) &= (S_dP+OP)-(SP+OP)\\
&\simeq R_d-R_s+\frac{1}{2}k_1R^2 \qquad \text{(A3)}
\end{aligned}
$$

where OP is a constant due to no deformation

$$k_1=\frac{R_s-R_d}{R_sR_d},\ R^2=x^2+y^2+z^2$$

For the neighboring point $P_1(x+\delta x, y, z)$, the optical path change $l_1(x+\delta x, y, z)$ can be obtained as follows:

$$
\begin{aligned}
l_1(x+\delta x,y,z) &= (S_dP_1+OP_1)-(SP_1+OP_1)\\
&= R_d-R_s+\frac{1}{2}k_1\{(x+\delta x)^2+y^2+z^2\} \qquad \text{(A4)}
\end{aligned}
$$

Small terms in the foregoing equation are neglected. Therefore, the relative optical path change is expressed as

$$
\begin{aligned}
\delta l = l_1(x+\delta x,y,z)-l(x,y,z) &= \frac{1}{2}k_1(2x\delta x+\delta x^2)\\
&= k_1x\delta x \qquad \text{(A5)}
\end{aligned}
$$

where δx^2 can be neglected if the magnitude of shearing δx is small. Consequently, the relative optical path change is given by the linear function in the direction of shearing. Similarly, in the y-direction,

$$\delta l = k_1y\delta y \qquad \text{(A6)}$$

After Deformation. $P(x,y,z)$ is displaced to $P'(x+u, y+v, z+w)$ and $P_1(x+\delta x,y,z)$ is displaced to $P_1'(x+\delta x+u+\delta u, y+v+\delta u, z+w+\delta w)$. By the similar derivation, it can be shown that the relative optical path change δl^* between the two neighboring points due to the deformation plus lens translation is

$$\begin{aligned}\delta l^* &\simeq \frac{1}{2}k_1(2x\delta x)+A\delta u+B\delta v+C\delta w\\ &= \left[k_1 x+A\frac{\delta u}{\delta x}+B\frac{\delta v}{\delta y}+C\frac{\delta w}{\delta x}\right]\delta x \qquad \text{(A7)}\end{aligned}$$

where

$$A=\frac{x}{R_d}+\frac{x}{R_o}-\frac{x_o}{R_o}-\frac{x_d}{R_d}, B=\frac{y}{R_d}+\frac{y}{R_o}-\frac{y_o}{R_o}-\frac{y_d}{R_d},$$

$$C=\frac{z}{R_d}+\frac{z}{R_o}-\frac{z_o}{R_o}-\frac{z_d}{R_d}$$

If the magnitude of shearing δx is small, the foregoing equation can be approximated as

$$\delta l^* \simeq \left(k_1 x+A\frac{\partial u}{\partial x}+B\frac{\partial v}{\partial x}+C\frac{\partial w}{\partial x}\right)\delta x \qquad \text{(A8)}$$

Similarly, in the shearing direction y

$$\delta l^* \simeq \left(k_1 y+A\frac{\partial u}{\partial y}+B\frac{\partial v}{\partial y}+C\frac{\partial w}{\partial y}\right)\delta y \qquad \text{(A9)}$$

R. H. Plaut
Professor,
Department of Civil Engineering,
Virginia Polytechnic Institute
and State University,
Blacksburg, Va 24061
Mem. ASME

L. W. Johnson
Professor,
Department of Mathematics,
Virginia Polytechnic Institute
and State University,
Blacksburg, Va 24061

N. Olhoff
Professor,
Aalborg University Center
DK-9220 Aalborg, Denmark

Bimodal Optimization of Compressed Columns on Elastic Foundations

We consider columns attached to elastic foundations and compressed by axial end loads. Pinned-pinned, clamped-clamped, and pinned-clamped boundary conditions are treated. The columns have rectangular sandwich cross sections with a fixed lightweight core and identical face sheets of variable thickness. For given total volume, we optimize the variation of the thickness along the column so as to maximize the buckling load. In most cases, the optimal design is bimodal (i.e., associated with two buckling modes). The optimal designs depend on the foundation stiffness, and the largest increase in buckling load relative to a column with constant thickness is 22 percent.

Introduction

Optimization of columns for maximum buckling load has been a topic of widespread interest (e.g., see the reviews by Haug and Olhoff in [1]). Part of this interest has been due to the existence of bimodal solutions [2], in which the optimal design has a double eigenvalue and two distinct eigenfunctions. However, only a few examples have been treated for columns which are attached to elastic foundations.

Kamat considered columns with solid cross sections of similar shape in his Ph.D. dissertation [3]. The ends were pinned and the area was assumed to be constant in 10 or 16 equal-length sections of the column. For two values of foundation stiffness, a unimodal formulation was used to maximize the buckling load. Similar problems were analyzed in Turner's Ph.D. dissertation [4], except that 20 sections were used and columns with clamped ends were treated as well. The results for the clamped-clamped columns were published in [5].

The dual problem of minimizing the total volume for a given buckling load was considered by Kiusalaas [6]. He assumed that the column had a sandwich cross section with 20 sections of constant face-sheet thickness. His analysis allowed for multiple eigenvalues. One example was presented, involving a pinned-pinned column with a given foundation stiffness, and his iterative solution procedure led to a bimodal optimal design.

Since the previous work on this problem is sparse, and some of it is restricted to a unimodal analysis, we decided to carry out the present study. We use a bimodal formulation and consider a range of foundation stiffnesses and three sets of boundary conditions: pinned-pinned, clamped-clamped, and pinned-clamped. An idealized sandwich cross section is assumed, for which the moment of inertia is proportional to the effective area. For given length and total volume of the column, the buckling load is maximized.

Formulation

We consider an elastic column of length L, which is attached to an elastic foundation of the Winkler type with constant stiffness coefficient K and is subjected to a compressive axial end load P. The horizontal coordinate is denoted X, with $0 \leq X \leq L$. The cross section is assumed to be rectangular with constant width B, consisting of a lightweight core of constant height H and Young's modulus E_c, covered by identical thin face sheets, each with variable thickness $T(X)$ and Young's modulus E_f (see Fig. 1). The bending stiffness of the cross section is given by

$$EI = \frac{1}{2} E_f B H^2 T \left(1 + \frac{2T}{H} + \frac{4T^2}{3H^2} + \frac{E_c H}{6E_f T} \right). \qquad (1)$$

We assume that the ratios T/H and $E_c H/(E_f T)$ are sufficiently small so that the last three terms inside the parentheses in (1) can be neglected.

The volume of each face sheet, denoted V_f, is fixed in our optimization problem. We denote the face-sheet cross-sectional area and the bending stiffness of a reference uniform column (having the same total volume) as A_u and EI_u, respectively. It follows that

$$A_u = V_f/L, EI_u = E_f H^2 V_f/(2L). \qquad (2)$$

The buckling load is the lowest of the eigenvalues $P = P_i$. If $Y_i(X)$ denotes the corresponding mode, the governing equilibrium equations are

---

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Mechanics Division, March 23, 1984; final revision, July 15, 1984.

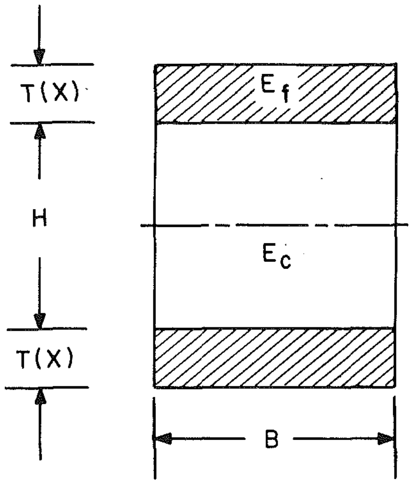


Fig. 1 Geometry of cross section

$$[EI(X)Y_i''(X)]'' + P_iY_i''(X) + KY_i(X) = 0,\ i=1,2,\ \ldots \tag{3}$$

We introduce the nondimensional quantities

$$x = X/L,\ y_i(x) = Y_i(xL)/L, \alpha(x) = A_f(xL)/A_u,$$
$$p_i = P_iL^2/(EI_u),\ k = KL^4/(EI_u). \tag{4}$$

The design variable is $\alpha(x)$, which is the nondimensional face-sheet area (or thickness). With the use of (4), the governing equations become

$$[\alpha(x)y_i''(x)]'' + p_iy_i''(x) + ky_i(x) = 0, i = 1,2,\ \ldots \tag{5}$$

and the constraint of constant volume becomes

$$\int_0^1 \alpha(x)dx = 1. \tag{6}$$

At a pinned end, $y_i = 0$ and $\alpha y_i'' = 0$. The latter condition will lead to $\alpha = 0$ at a pinned end of an optimal design. At a clamped end, $y_i = 0$ and $y_i' = 0$.

If we multiply (5) by $y_i(x)$, integrate, and carry out appropriate integration by parts, we obtain the standard Rayleigh quotient for the eigenvalue p_i. We normalize the modes such that the denominator of this quotient is unity, i.e.,

$$\int_0^1 (y_i')^2dx = 1,\ i = 1,2,\ldots, \tag{7}$$

and then we obtain

$$p_i = \int_0^1 \alpha(y_i'')^2dx + k\int_0^1 y_i^2dx,\ i = 1,2\ \ldots \tag{8}$$

For a given set of boundary conditions and a given foundation stiffness k, we wish to determine the design function $\alpha(x)$ which satisfies (6) and maximizes the minimum eigenvalue (i.e., the buckling load). The optimal solution may be multimodal, and it is convenient to utilize a bound formulation [7] in which we maximize a lower bound β on p_i. We construct the following augmented functional $\mathcal{L}$ to be minimized:

$$\mathcal{L} = -\beta + \sum_{i=1}^{\infty} \gamma_i(\beta - p_i + \theta_i^2)$$
$$+ \sum_{i=1}^{\infty} \eta_i\left[\int_0^1 (y_i')^2dx - 1\right] + \mu\left[\int_0^1 \alpha dx - 1\right] \tag{9}$$

where the p_i are given by (8), θ_i are slack variables, and γ_i, η_i, and μ are constant Lagrange multipliers.

Stationarity of $\mathcal{L}$ with respect to γ_i and θ_i leads to the fact that $\gamma_i = 0$ if p_i is not the minimum eigenvalue, while stationarity with respect to β implies that the sum of the γ_i values is unity. Stationarity of $\mathcal{L}$ with respect to $y_i(x)$ leads to (5). Finally, stationarity with respect to $\alpha(x)$ leads to the optimality condition

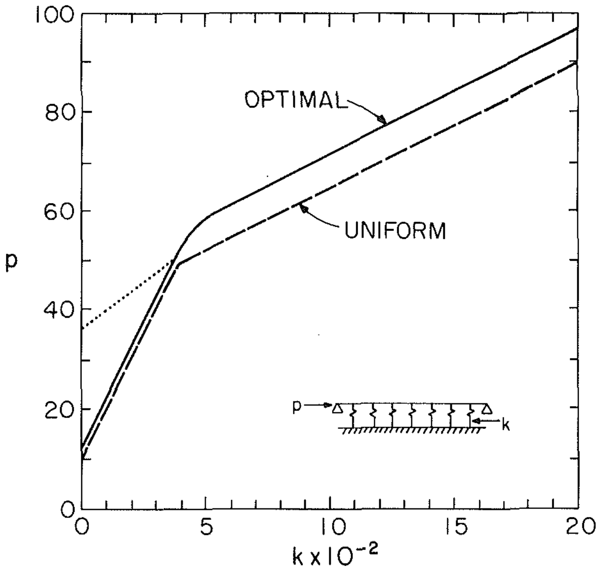


Fig. 2 Buckling loads for pinned-pinned columns (dotted line is second buckling load for optimal column when unimodal)

$$\sum_{i=1}^{\infty} \gamma_i[y_i''(x)]^2 = \mu. \tag{10}$$

We seek solutions to the set of equations (5), (6), (7), and (10), along with the derived conditions on the γ_i. Equation (10) is a necessary condition for stationarity of buckling load, and the numerical procedure converges to a maximum value of the buckling load. We will discuss the results separately for the three sets of boundary conditions.

Pinned-Pinned Column

For a uniform pinned-pinned column ($\alpha = 1$) attached to an elastic foundation, the buckling load and mode are as follows [8]: for the integer n such that

$$(n-1)^2n^2\pi^4 \le k \le n^2(n+1)^2\pi^4, \tag{11}$$

we have

$$p_{\text{unif}} = n^2\pi^2 + \frac{k}{n^2\pi^2}, y_1(x) = \frac{\sqrt{2}}{n\pi}\sin(n\pi x). \tag{12}$$

The mode shape changes from symmetric to antisymmetric (or vice versa) as k is increased past a transition value $(n-1)^2n^2\pi^4$, and at these values the buckling load is a double eigenvalue and is associated with two modes. In Fig. 2, the dashed lines depict p_{unif} for $0 \le k \le$ 2,000, with a transition value from $n = 1$ to $n = 2$ at $k = 389.6$.

If the foundation stiffness k is zero, the optimal design and corresponding buckling load and mode are given by [9]

$$\alpha(x) = 6x(1-x),\ p_{\text{opt}} = 12, y_1(x) = \sqrt{3}x(1-x). \tag{13}$$

The design $\alpha(x)$ is shown in Fig. 3(*a*), along with the reference uniform column having the same volume. This optimal solution is unimodal. The second buckling load, $p = 36$, corresponds to an antisymmetric mode.

In the range $0 < k < 376$, we find that the optimal solution remains unimodal. The governing mode $y_1(x)$ is given in (13), and using this mode in (5), along with the conditions $\alpha(0) = 0$, $\alpha(1) = 0$, and (6), leads to the optimal design

$$\alpha(x) = \left(6 + \frac{k}{120}\right)x - \left(6 + \frac{k}{20}\right)x^2 + \frac{k}{12}x^3 - \frac{k}{24}x^4 \tag{14}$$

and the corresponding buckling load

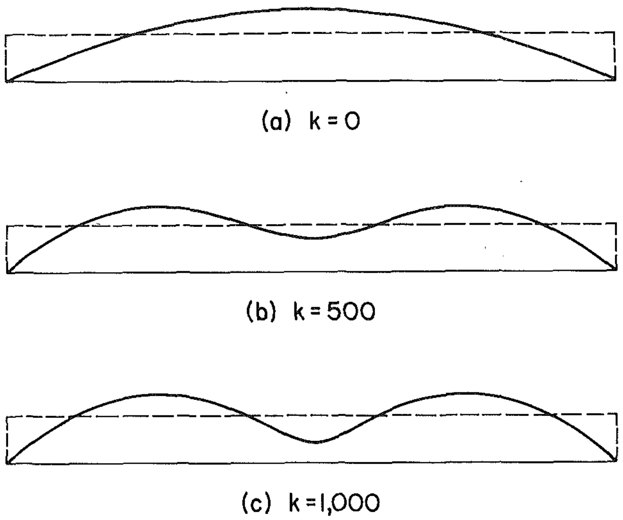


Fig. 3 Optimal designs for pinned-pinned columns

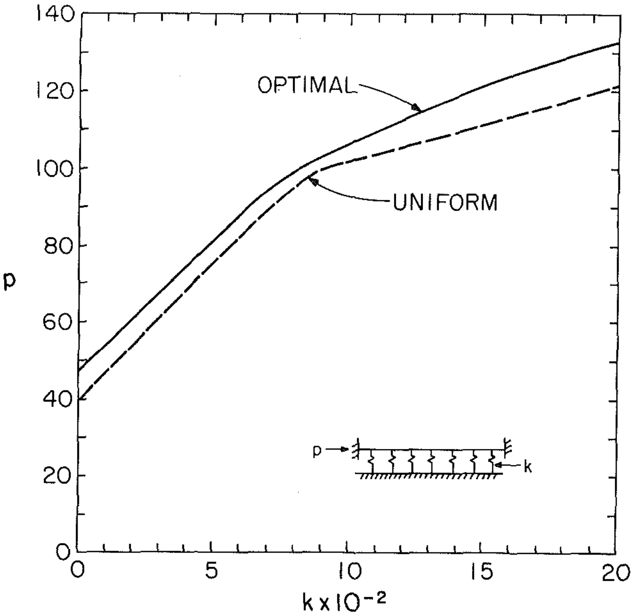


Fig. 4 Buckling loads for clamped-clamped columns

$$p_{\text{opt}} = 12 + (k/10). \tag{15}$$

As with the uniform column, this load increases linearly with k. The second buckling mode associated with the design (14) was computed by applying a finite-difference method to (5), and the resulting values are depicted by the dotted line in Fig. 2. At $k = 376$, p_{opt} (shown by the solid line) and the second buckling load coalesce.

If $k > 376$, the optimal solution is bimodal. We obtained a numerical solution to the governing equations with the use of finite differences. We assumed $\alpha(x)$ and $y_1(x)$ to be symmetric and $y_2(x)$ to be antisymmetric, and we used 16 knots in the region $0 < x \leq 1/2$. Then (5), (6), (7), and (10) become a system of nonlinear algebraic equations in the following quantities: p, c_1, c_2, and the unknown values of α, y_1, and y_2 at the knots. In order to avoid solutions involving negative values of α, we replaced α by ϕ^2 and used the ϕ values as unknowns. The resulting system of equations was solved directly with the use of a standard computer program.

The maximum buckling load obtained from this procedure, p_{opt}, is plotted in Fig. 2 for the range $376 \leq k \leq 2000$ and listed in Table 1 for $k = 500$, 1000, 1500, and 2000. The corresponding optimal designs $\alpha(x)$ for $k = 500$ and $k = 1000$ are shown in Figs. 3(b) and (c), respectively. The optimal designs for $k = 1500$ and $k = 2000$ are almost identical to that in Fig. 3(c). We also note that the design obtained by Kiusalaas [6] (for the dual problem) corresponds to $k = 443$ and has a local minimum in the center as do the designs shown in Figs. 3(b) and (c).

Table 1 Critical loads

k	Pinned-Pinned P_{unif}	Pinned-Pinned P_{opt}	Clamped-Clamped P_{unif}	Clamped-Clamped P_{opt}	Pinned-Clamped P_{unif}	Pinned-Clamped P_{opt}
0	9.9	12.0	39.5	48.0	20.2	24.7
500	52.1	58.6	75.4	81.1	55.8	64.1
1,000	64.8	71.9	101.2	106.1	74.5	86.5
1,500	77.5	84.4	111.2	121.3	87.0	99.8
2,000	90.1	97.0	121.1	132.8	98.1	111.4

The governing modes change shape as k is increased. As for the uniform column, the modes become more wavy and have more nodes with increasing k. The optimal designs also become more wavy as k grows larger.

Clamped-Clamped Column

For a uniform column with both ends clamped, the buckling load p_{unif} is the lowest root of the equation [8]

$$\sqrt{p - 2\sqrt{k}}\sin\left(\frac{1}{2}\sqrt{p + 2\sqrt{k}}\right) = \pm\sqrt{p + 2\sqrt{k}}\sin\left(\frac{1}{2}\sqrt{p - 2\sqrt{k}}\right) \tag{16}$$

Transitions between symmetric and antisymmetric modes occur when

$$k = (n^2 - 1)^2\pi^4, p_{\text{unif}} = 2(n^2 + 1)\pi^2, n = 2, 3, \ldots \tag{17}$$

Some other values of p_{unif} are listed in Table 1. In Fig. 4, p_{unif} is plotted versus k (dashed curve), and a transition from symmetric to antisymmetric buckling occurs at $k = 9\pi^4$. The buckling load is almost linear in k between transition points.

The optimal designs are found to be bimodal for all values of $k \geq 0$. At $k = 0$, the maximum buckling load for a sandwich column with clamped ends is $p_{\text{opt}} = 47.956$ [10]. Optimal buckling loads for the range $0 \leq k \leq 2000$ are plotted in Fig. 4, and some values are listed in Table 1. The optimal designs $\alpha(x)$ for $k = 0$, 500, 1000, and 1500 are shown in Fig. 5, and the design for $k = 2000$ is almost identical to that in Fig. 5(d). These results were obtained by the procedure described in the previous section, except twice as many knots were used in this case. Again, the optimal design and the associated modes become more wavy as k increases.

Pinned-Clamped Column

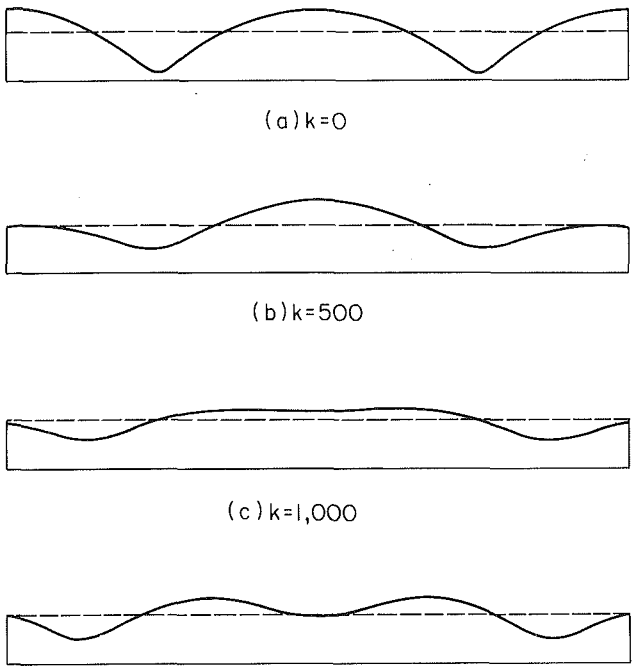
Here we assume the column is pinned at $x = 0$ and clamped at $x = 1$. For the uniform design ($\alpha = 1$), the buckling load p_{unif} is the lowest root of the equation

$$\sqrt{p - 2\sqrt{k}}\sin\sqrt{p + 2\sqrt{k}} = \sqrt{p + 2\sqrt{k}}\sin\sqrt{p - 2\sqrt{k}} \tag{18}$$

It is plotted in Fig. 6, and some values are listed in Table 1. Some other solutions are

$$p_{\text{unif}} = (n^2 + 1)\pi^2/2 \text{ at } k = (n^2 - 1)^2\pi^4/16, n = 2, 3, \ldots \tag{19}$$

Since the boundary conditions are not symmetric, this case does not involve symmetric and antisymmetric modes.


(a) k=0

(b) k=500

(c) k=1,000

(d) k=1,500

Fig. 5 Optimal designs for clamped-clamped columns

For $k=0$, the optimal solution is bimodal and is given by [11, 12]

$$\alpha(x)=\begin{cases} r^2x(1-\sqrt{2}x)/(2\sqrt{2}) & \text{if } 0\le x\le 1/\sqrt{2},\\ r^2(\sqrt{2}x-1)(\sqrt{2}-x)/(2\sqrt{2}) & \text{if } 1/\sqrt{2}\le x\le 1,\end{cases} \tag{20}$$

$$y_1(x)=\begin{cases} r(1-\sqrt{2})x+(r/2)x^2 & \text{if } 0\le x\le 1/\sqrt{2},\\ -(r/2)(1-x)^2 & \text{if } 1/\sqrt{2}\le x\le 1,\end{cases} \tag{21}$$

$$y_2(x)=\begin{cases} r[1-(1/\sqrt{2})]x-(r/2)x^2 & \text{if } 0\le x\le 1/\sqrt{2},\\ -(r/2)(1-x)^2 & \text{if } 1/\sqrt{2}\le x\le 1,\end{cases} \tag{22}$$

where

$$r^2=p_{\text{opt}}=6/(3\sqrt{2}-4)=24.73. \tag{23}$$

The design (20), depicted in Fig. 7(a), has a hinge at $x=1/\sqrt{2}$ (where $\alpha=0$). The mode $y_1(x)$ in (21) has continuous slope, while the mode $y_2(x)$ in (22) has a discontinuity in slope at the hinge and can be obtained from $y_1(x)$ by reflection of the section from $x=0$ to $x=1/\sqrt{2}$ across a chord connecting the pinned end with the hinge [12]. Both modes have constant $(y_i'')^2$, $0\le x\le 1$, and thus satisfy the unimodal optimality condition as well as the bimodal one.

When $k>0$, we find that the optimal solution is bimodal and does not contain an inner hinge (a unimodal solution with a hinge can be obtained analytically, but it has a lower buckling load). The boundary conditions are not symmetric, and therefore we cannot separate $y_1(x)$ and $y_2(x)$ into symmetric and antisymmetric functions. We used 32 knots in the region $0<x\le 1$ and again solved the governing equations directly.

Results for the pinned-clamped case are presented in Table 1 and Figs. 6 and 7. They exhibit similarities to the previous results in the form of the curve of p_{opt} versus k and in the increased waviness of the optimal design and associated modes with increasing k.

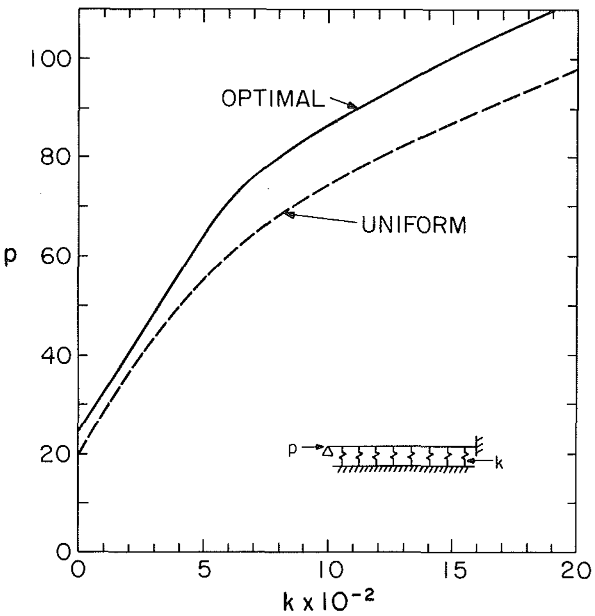
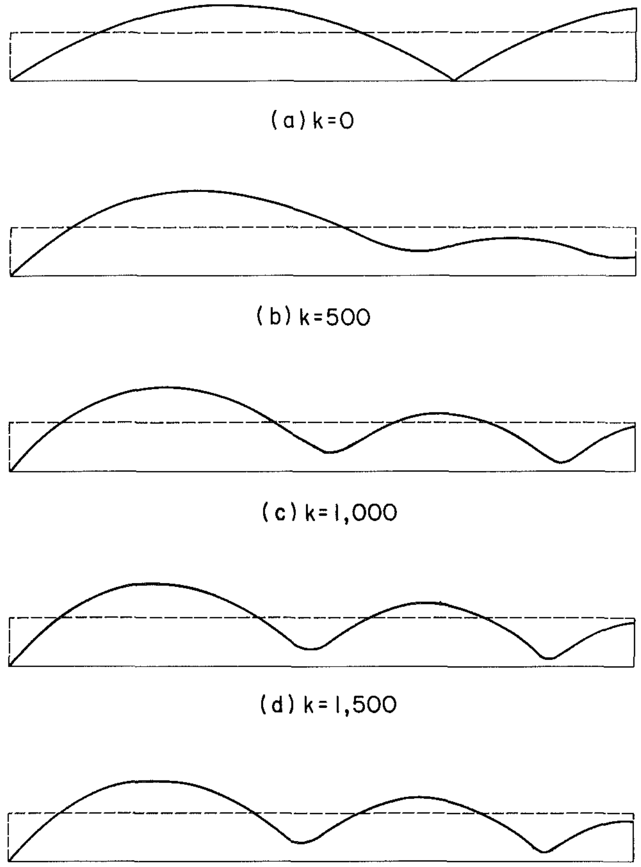


Fig. 6 Buckling loads for pinned-clamped columns


(a) k=0

(b) k=500

(c) k=1,000

(d) k=1,500

(e) k=2,000

Fig. 7 Optimal designs for pinned-clamped columns

Concluding Remarks

We have determined optimal designs for pinned-pinned, clamped-clamped, and pinned-clamped sandwich columns attached to elastic foundations. For pinned-pinned columns with a sufficiently small foundation stiffness coefficient ($0 \leq k < 376$), the optimal solution is unimodal (i.e., associated with a single buckling mode). In all other cases, the solution is bimodal. The percentage increase in buckling load over that of a uniform column is 22 percent when $k = 0$ and less than that when the foundation is present.

For uniform columns, the governing modes become more wavy as k increases (e.g., see (12)). This is also true for the optimally designed columns. In addition to the results presented here for $0 \leq k \leq 2000$, we carried out a few computations for values of k up to 20,000, and the trend continued: the solutions were bimodal and the modes and designs had increasing numbers of local minima and maxima.

The columns were assumed to have ideal rectangular sandwich cross sections of fixed width, so that the moment of inertia was proportional to the effective area. The results can also be applied to some other cross sections, such as ideal I-beams. For the case of a solid rectangular cross section of fixed height and varying width, however, the results will only be applicable if the foundation stiffness is constant; for example, if the restoring force were proportional to the contact area, then the stiffness would be a function of the width and would vary along the column [13].

Finally, we note that the occurrence of bimodal optimal designs for columns is not restricted to symmetric boundary conditions where the modes associated with the lowest eigenvalue can be assumed to be a symmetric mode and an antisymmetric mode. This was demonstrated in the case of a pinned-clamped column, previously in [12] for no elastic foundation (where the optimal design possesses an inner hinge), and here in the presence of an elastic foundation (where no inner hinges occur).

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant CEE-8210222.

References

1 Haug, E. J., and Cea, J., *Optimization of Distributed Parameter Structures*, Volume I, Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands, 1981.

2 Olhoff, N., and Rasmussen, S. H., "On Single and Bimodal Optimum Buckling Loads of Clamped Columns," *International Journal of Solids and Structures*, Vol. 13, 1977, pp. 605-614.

3 Kamat, M. P., "Optimization of Structural Elements for Stability and Vibration," Ph.D. Dissertation, Georgia Institute of Technology, 1972.

4 Turner, H. K., "Optimal Design of Elastic Structures for Stability Under Multiple Loads," Ph.D. Dissertation, Virginia Polytechnic Institute and State University, 1979.

5 Turner, H. K., and Plaut, R. H., "Optimal Design for Stability Under Multiple Loads," *Journal of the Engineering Mechanics Division, Proceedings, ASCE*, Vol. 106, 1980, pp. 1365-1382.

6 Kiusalaas, J., "Optimal Design of Structures with Buckling Constraints," *International Journal of Solids and Structures*, Vol. 9, 1973, pp. 863-878.

7 Bendsφe, M. P., Olhoff, N., and Taylor, J. E., "A Variational Formulation for Multicriteria Structural Optimization," *Journal of Structural Mechanics*, Vol. 11, 1983, pp. 523-544.

8 Hetenyi, M., *Beams on Elastic Foundation*, The University of Michigan Press, Ann Arbor, 1946.

9 Prager, W., and Taylor, J. E., "Problems of Optimal Structural Design," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, 1968, pp. 102-106.

10 Masur, E. F., "Optimal Structural Design Under Multiple Eigenvalue Constraints," *International Journal of Solids and Structures*, Vol. 20, 1984, pp. 211-231.

11 Banichuk, N. V., and Karihaloo, B. L., "On the Solution of Optimization Problems with Singularities," *International Journal of Solids and Structures*, Vol. 13, 1977, pp. 725-733.

12 Olhoff, N., and Niordson, F. I., "Some Problems Concerning Singularities of Optimal Beams and Columns," *Zeitschrift für angewandte Mathematik und Physik*, Vol. 59, 1979, pp. T16-T26.

13 Adali, S., "Optimal Circular Ring Sector Subject to Inequality Constraints," *Journal de Mécanique Appliquée*, Vol. 4, 1980, pp. 131-154.

R. H. Plaut
Professor,
Department of Civil Engineering.
Mem. ASME

L. W. Johnson
Professor,
Department of Mathematics.

Virginia Polytechnic Institute
and State University,
Blacksburg, Va. 24061

Optimal Forms of Shallow Cylindrical Panels With Respect to Vibration and Stability

Thin, shallow, elastic, cylindrical panels with rectangular planform are considered. We seek the midsurface form which maximizes the fundamental frequency of vibration, and the form which maximizes the buckling value of a uniform axial load. The material, surface area, and uniform thickness of the panel are specified. The curved edges are simply supported, while the straight edges are either simply supported or clamped. For the clamped case, the optimal panels have zero slope at the edges. In the examples, the maximum fundamental frequency is up to 12 percent higher than that of the corresponding circular cylindrical panel, while the buckling load is increased by as much as 95 percent. Most of the solutions are bimodal, while the rest are either unimodal or trimodal.

Introduction

Cylindrical panels are used in a variety of structures, such as airplanes, rocket boosters, ships, and submarines [1]. They are often subjected to disturbances which cause vibrations, or to axial compressive forces, which can cause buckling. Hence, there has been extensive research on the vibration and buckling of cylindrical panels. In most cases, the form of the panel perpendicular to its generators has been taken to be circular. Here, however, we consider shallow cylindrical panels with arbitrary form, and we seek the form which maximizes the fundamental vibration frequency or the buckling load.

A few representative papers dealing with cylindrical panels can be mentioned. Various sets of boundary conditions have been treated in the literature [2–4]. The vibration of noncircular cylindrical panels was analyzed in [5–7]. References 8 and 9 considered nonuniform compressive forces, while postbuckling behavior was discussed in [10–13]. Elastic-plastic behavior [1, 12], stiffened panels [14], and composite panels [15] have received attention, and some optimization studies have been carried out [1, 16, 17].

In the present work, the governing equations are based on Donnell's theory. The curved edges of the panel are assumed to be simply supported with no tangential displacement, while the straight edges are assumed to be either simply supported or clamped, with no normal displacement. For the case of buckling, a compressive, normal load is applied uniformly along the curved edges.

In the optimization formulation, the boundary conditions, material, thickness, and surface area of the panel are specified, and the form is variable. Hence, the total volume and mass of the panel are given. The optimality conditions for maximum fundamental frequency and maximum buckling load are derived with the use of the calculus of variations, allowing for multimodal solutions. To obtain numerical results, an iterative technique based on the optimality condition is applied to get close to the optimal solution and determine its governing modes, followed by a direct solution of the governing equations. The results depend on the boundary conditions, aspect ratio, and surface area parameter. Twelve cases are solved for both the vibration problem and the buckling problem. Three of the resulting optimal solutions are unimodal, 18 are bimodal, and three are trimodal.

We note that a similar type of investigation was carried out in [18, 19] for shallow, axisymmetric shells with a circular boundary.

Shell Equations

We consider a thin, shallow, elastic, cylindrical panel with constant thickness h, density ρ, Young's modulus E, and Poisson's ratio ν. It has a rectangular planform with $X = 0, a$ along its curved edges, and $Y = 0, b$ along its straight edges. The height of the middle surface above the base plane (i.e., the form) is $Z(Y)$, which will be chosen in an optimal manner. At the straight edges $Y = 0, b$, the slopes $Z'(Y)$ are not specified.

The area S of the middle surface can be written as

$$S = a\left[b + \frac{1}{2}\int_0^b (Z')^2 dY\right] \tag{1}$$

under our shallowness assumption. We define the nondimensional surface area parameter β^2 by

$$\beta^2 = 2(S - ab)a/(q^2 b) \tag{2}$$

where

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 15, 1984; final revision, April 15, 1985.

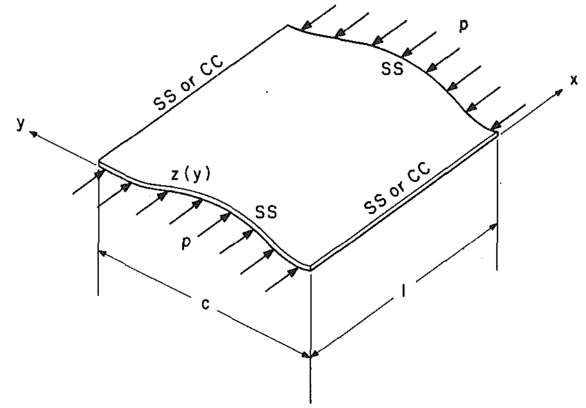


Fig. 1 Geometry of cylindrical panel (in nondimensional terms)

$$q = h/\sqrt{12(1-\nu^2)}. \tag{3}$$

For shallow, circular cylindrical panels with height H and radius of curvature R, we have $Z(Y) = 4HY(b-Y)/b^2$ and $\beta^2 = 16a^2H^2/(3q^2b^2) = a^2b^2/(12q^2R^2)$.

With T denoting time, we let $W(X, Y, T)$ be the upward deflection of the panel and $F(X, Y, T)$ be the Airy stress function. In-plane inertia is neglected. The linear equations of motion, based on Donnell's theory with uniform axial prestress, are [20]

$$D\nabla^4 W + PW_{XX} - Z''F_{XX} + \rho h W_{TT} = 0, \tag{4}$$

$$\nabla^4 F + EhZ'' W_{XX} = 0, \tag{5}$$

where P is a uniform, normal, compressive force per unit length on the edges $X = 0, a$,

$$D = \frac{Eh^3}{12(1-\nu^2)}, \nabla^4(\ ) = \left(\frac{\partial^4}{\partial X^4} + 2\frac{\partial^4}{\partial X^2 \partial Y^2} + \frac{\partial^4}{\partial Y^4}\right)(\ ), \tag{6}$$

and subscripts X, Y, and T denote partial derivatives.

To study vibrations with frequency Ω, we let

$$W(X,Y,T) = \tilde{W}(X,Y)\cos\Omega T, F(X,Y,T) = \tilde{F}(X,Y)\cos\Omega T. \tag{7}$$

We then define the nondimensional quantities

$$x = X/a, y = Y/a, c = b/a, z = Z/q, w = \tilde{W}/q,$$
$$f = \tilde{F}/D, p = Pa^2/D, \omega = \Omega a^2\sqrt{\rho h/D} \tag{8}$$

where $0 \leq x \leq 1$, $0 \leq y \leq c$ (see Fig. 1), and we define the design function $g(y)$ by

$$g(y) = z'(y). \tag{9}$$

In nondimensional terms, equation (2) becomes

$$\frac{1}{c}\int_0^c g^2 dy = \beta^2 \tag{10}$$

and the governing equations for $w(x, y)$, $f(x, y)$ become

$$\nabla^4 w + pw_{xx} - g'f_{xx} - \omega^2 w = 0, \tag{11}$$

$$\nabla^4 f + g' w_{xx} = 0, \tag{12}$$

where

$$\nabla^4(\ ) = \left(\frac{\partial^4}{\partial x^4} + 2\frac{\partial^4}{\partial x^2 \partial y^2} + \frac{\partial^4}{\partial y^4}\right)(\ ) \tag{13}$$

and subscripts x and y denote partial derivatives.

Classical simply supported boundary conditions are assumed on the curved edges $x = 0,1$. Normal displacement is allowed, but not tangential displacement, and [13]

$$w = w_{xx} = f = f_{xx} = 0. \tag{14}$$

On the straight edges $y = 0,c$, we assume either simply supported or clamped conditions, with tangential displacement but no normal displacement. These conditions are denoted $SS2$ and $CC2$ in [4], and lead to

$$w = w_{yy} = f_{xy} = f_{yyy} = 0 \tag{15}$$

and

$$w = w_y = f_{xy} = f_{yyy} = 0, \tag{16}$$

respectively (see the Appendix).

We now consider the two optimization problems separately.

Maximum Fundamental Frequency

For the case of free vibrations, we set $p = 0$ in (11). Further, we denote the nondimensional natural frequencies by ω_n, $n = 1, 2, \ldots$, and the corresponding vibration modes and stress functions by $w_n(x, y)$ and $f_n(x, y)$, respectively. Multiple frequencies may exist; for example, we may have a double frequency $\omega_1 = \omega_2$ with corresponding eigenfunction pairs w_1, f_1 and w_2, f_2.

From (11) and (12), we can obtain a Rayleigh quotient for ω_n^2 which becomes

$$\omega_n^2 = \int_0^c\int_0^1 [(\nabla^2 w_n)^2 - (\nabla^2 f_n)^2 + 2(w_n)_x(f_n)_x g']dxdy \tag{17}$$

if we normalize its denominator by

$$\int_0^c\int_0^1 w_n^2 dxdy = 1, \tag{18}$$

where

$$\nabla^2(\ ) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)(\ ). \tag{19}$$

Our objective is to determine the design function $g(y)$ which maximizes the fundamental (i.e., lowest) vibration frequency for a given value of the surface area parameter β^2. The optimal solution may be multimodal, and it is convenient to utilize a bound formulation [21, 22] in which we maximize a lower bound B on ω_n^2. With the use of Lagrange multipliers γ_n, α_n, and μ, and slack variables θ_n, we construct the following augmented functional:

$$L = -B + \sum_{n=1}^{\infty} \gamma_n(B - \omega_n^2 + \theta_n^2) + \sum_{n=1}^{\infty} \alpha_n\left[\int_0^c\int_0^1 w_n^2 dxdy - 1\right] + \mu\left[\frac{1}{c}\int_0^c g^2 dy - \beta^2\right] \tag{20}$$

where ω_n is given by (17) and we have incorporated the constraints (10) and (18).

Stationarity of L with respect to γ_n and θ_n leads to the fact that $\gamma_n = 0$ if ω_n is not the fundamental frequency, while stationarity with respect to B implies that the sum of the γ_n values is unity. Stationarity of L with respect to w_n and f_n leads to (11) and (12), respectively. Finally, stationarity with respect to the design function g furnishes the optimality condition

$$g(y) = \frac{c}{\mu}\sum_{n=1}^{\infty} \gamma_n \int_0^1 [(w_n)_x(f_n)_x]_y dx. \tag{21}$$

If the straight edges of the shell are clamped, it follows from (9), (16), and (21) that the optimal form has $z' = 0$ (i.e., zero slope) at these edges.

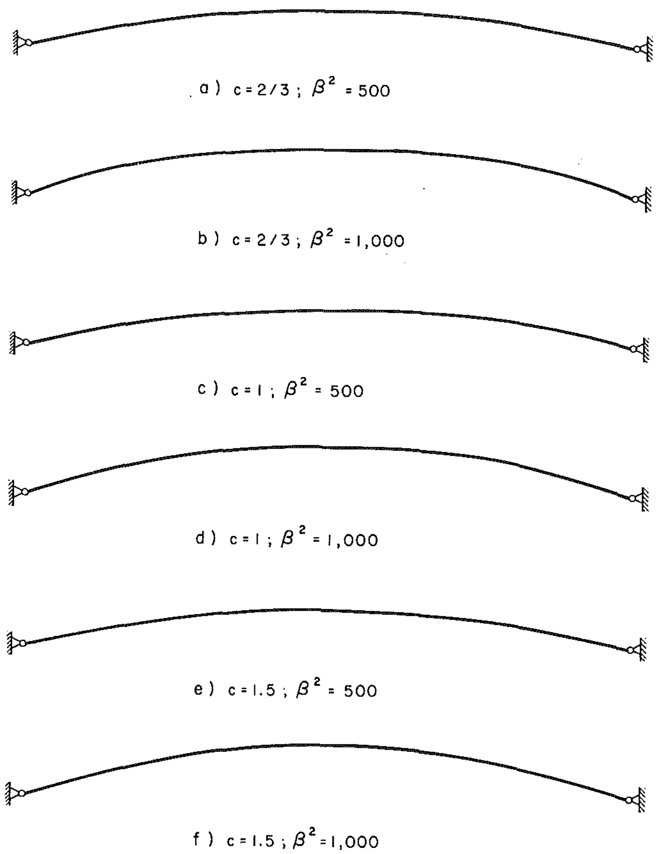
Maximum Buckling Load

For the case of buckling, we set $\omega = 0$ in (11) and denote the eigenvalues by p_n, $n = 1, 2, \ldots$, and the corresponding eigenfunctions by $w_n(x, y)$ and $f_n(x, y)$. If we now normalize the modes by

$$\int_0^c\int_0^1 [(w_n)_x]^2 dxdy = 1, \tag{22}$$

the appropriate Rayleigh quotient becomes

$$p_n = \int_0^c\int_0^1 [(\nabla^2 w_n)^2 - (\nabla^2 f_n)^2$$


a) c = 2/3 ; $\beta^2 = 500$

b) c = 2/3 ; $\beta^2 = 1{,}000$

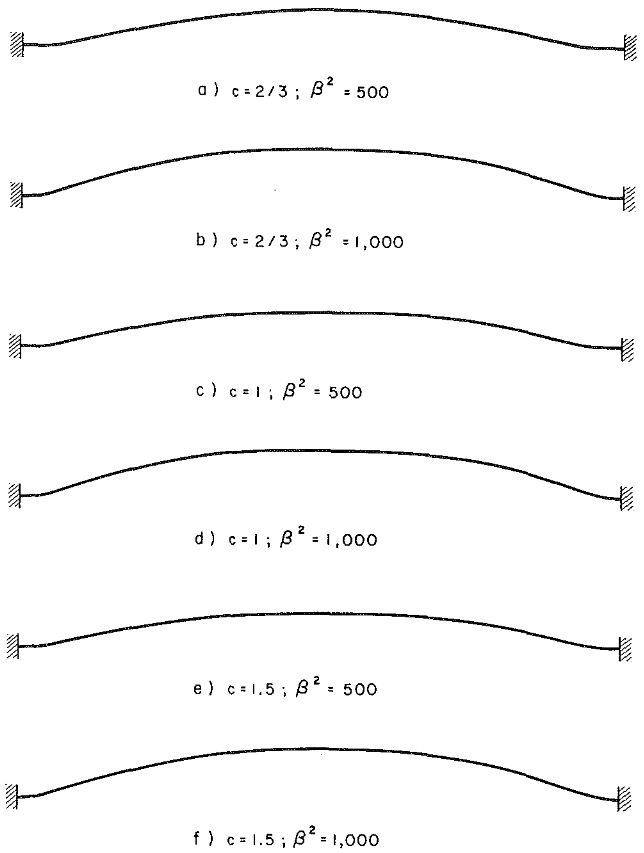
c) c = 1 ; $\beta^2 = 500$

d) c = 1 ; $\beta^2 = 1{,}000$

e) c = 1.5 ; $\beta^2 = 500$

f) c = 1.5 ; $\beta^2 = 1{,}000$

Fig. 2 Optimal forms for maximum fundamental frequency with simply supported edges


a) c = 2/3 ; $\beta^2 = 500$

b) c = 2/3 ; $\beta^2 = 1{,}000$

c) c = 1 ; $\beta^2 = 500$

d) c = 1 ; $\beta^2 = 1{,}000$

e) c = 1.5 ; $\beta^2 = 500$

f) c = 1.5 ; $\beta^2 = 1{,}000$

Fig. 3 Optimal forms for maximum fundamental frequency with straight edges clamped

$$+2(w_n)_x(f_n)_x g']dxdy. \tag{23}$$

The optimization formulation is similar to that in the previous section. In (20), ω_n^2 is replaced by p_n (which has the same formula) and the normalization condition (18) is replaced by (22). The optimality condition again is given by (21).

Solution Procedure

For both optimization problems, we let

$$w_n(x,y) = w_{mn}(y)\sin(m\pi x),$$
$$f_n(x,y) = f_{mn}(y)\sin(m\pi x). \tag{24}$$

Then (11) and (12) become

$$w''''_{mn}n - 2m^2\pi^2 w''_{mn} + (m^4\pi^4 - m^2\pi^2 p_{mn}$$
$$-\omega_{mn}^2)w_{mn} + m^2\pi^2 g' f_{mn} = 0, \tag{25}$$

$$f''''_{mn} - 2m^2\pi^2 f''_{mn} + m^4\pi^4 f_{mn} - m^2\pi^2 g' w_{mn} = 0, \tag{26}$$

where we denote the corresponding frequency and load by ω_{mn} and p_{mn}, respectively, with $p_{mn} = 0$ when the frequency is optimized and $\omega_{mn} = 0$ when the buckling load is optimized. The optimality condition (21) can be written in the form

$$g(y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\Gamma_{mn}[w_{mn}(y)f_{mn}(y)]' \tag{27}$$

where the Γ_{mn} are constants which are zero except for pairs m, n associated with the minimum eigenvalue. For example, if the lowest frequency is bimodal with $\omega_{11} = \omega_{12}$, then all coefficients Γ_{mn} in (27) are zero except Γ_{11} and Γ_{12}.

The boundary conditions (either (15) or (16)), aspect ratio c, and surface area parameter β^2 are specified. Two phases are used in the solution procedure. In the first phase, we choose an initial design $g(y)$ and solve (25) and (26) for different values of m to find the lowest eigenvalue (either the fundamental frequency or buckling load) and the associated eigenfunction pair (or pairs) w_{mn}, f_{mn}. Equation (27) is employed to modify the design, and (25), (26) are solved using the new design. This is repeated to increase the lowest eigenvalue until the multiplicity of the optimal eigenvalue and the forms of the associated eigenfunctions become clear. Next, in the second phase, we solve (27) directly for $g(y)$, using (25) and (26) to furnish $w_{mn}(y)$ and $f_{mn}(y)$. When the solution is trimodal, for example, this procedure involves a set of seven nonlinear differential equations. The constraints (10) and either (18) or (22) are incorporated to scale the functions.

To obtain the numerical solutions, we divide the length $0 \le y \le c$ into 32 sections and apply finite differences at the knots. A standard eigenvalue program is used in the first phase of the solution procedure, and a quasi-Newton method in the second phase.

A number of locally optimal solutions may exist. We attempt to determine the global optimum for each case by starting from several initial designs, converging to a local optimum, and then selecting the best of these as the optimal design.

Results

We computed optimal solutions for aspect ratios $c = 2/3$, 1, and 1.5, and surface area parameter values $\beta^2 = 500$ and 1,000. In all cases, the optimal forms $z(y)$ turned out to be symmetric about the center, $y = c/2$.

First, consider the free vibration problem. The optimal forms are depicted in Figs. 2 and 3 when the straight edges are simply supported and clamped, respectively. In the figures, we plot $z(y)/c$ versus y/c. The optimal panels in Figs. 2 and 3 lie above the base plane. As mentioned earlier, at a clamped edge the optimal form $z(y)$ has a zero slope.

Some frequencies associated with these optimal designs are listed in Table 1, along with frequencies of a circular cylin-

Table 1 Vibration frequencies for circular and optimal forms

	c	β^2	Circular panel				Optimal panel				Frequency ratio
			ω_1	ω_2	ω_3	ω_4	ω_1	ω_2	ω_3	ω_4	
SS	2/3	500	103.2	121.8	154.5	210.8	104.5	118.6	151.8	208.7	1.01
SS	2/3	1,000	107.8	157.6	187.2	237.2	116.0	132.6	166.8	222.7	1.08
SS	1	500	59.7	86.4	123.0	183.7	63.4	81.6	120.1	182.1	1.06
SS	1	1,000	68.4	109.3	142.9	198.2	74.7	97.2	135.1	194.0	1.09
SS	1.5	500	41.3	65.5	105.8	170.0	42.6	64.4	105.4	169.9	1.03
SS	1.5	1,000	50.8	80.5	117.0	177.4	50.8	80.5	117.0	177.4	1.00
CC	2/3	500	111.1	126.3	160.2	216.8	124.7[1]	140.6	175.6	231.6	1.12
CC	2/3	1,000	144.4	158.8	189.8	241.2	150.8	164.6	195.5	247.5	1.04
CC	1	500	69.6	87.3	124.8	185.5	74.8	92.1	129.4	189.7	1.07
CC	1	1,000	80.5	109.3	143.7	199.6	85.1	104.3	141.8	200.9	1.06
CC	1.5	500	44.7	65.8	106.4	170.6	46.2	67.3	108.6	172.7	1.04
CC	1.5	1,000	52.1	80.5	117.4	177.9	54.8	80.6	119.4	180.2	1.05

[1] unimodal; all other optimal solutions are bimodal

Table 2 Buckling loads for circular and optimal forms

	c	β^2	Circular panel				Optimal panel				Load ratio
			p_1	p_2	p_3	p_4	p_1	p_2	p_3	p_4	
SS	2/3	500	1080.	375.7	$\underline{268.7}$	281.3	547.1	$\underline{288.9}$	$\underline{288.9}$	327.4	1.08
SS	2/3	1,000	1178.	629.0	394.6	$\underline{356.4}$	687.2	$\underline{414.4}$	$\underline{414.4}$	436.3	1.16
SS	1	500	361.0	189.1	$\underline{170.3}$	213.6	$\underline{201.4}$	$\underline{201.4}$	219.1	259.1	1.18
SS	1	1,000	474.4	302.8	$\underline{229.8}$	248.8	$\underline{298.2}$	$\underline{298.2}$	$\underline{298.2}$	317.2	1.30
SS	1.5	500	172.9	$\underline{108.6}$	126.1	183.1	$\underline{167.3}$	185.4	208.5	249.6	1.54
SS	1.5	1,000	261.6	164.2	$\underline{154.1}$	199.3	$\underline{296.9}$	$\underline{296.9}$	321.4	350.8	1.93
CC	2/3	500	1251.	404.0	$\underline{289.1}$	297.5	966.9	416.3	$\underline{369.1}$	390.7	1.28
CC	2/3	1,000	2113.	639.0	405.5	$\underline{368.6}$	914.4	$\underline{527.6}$	$\underline{527.6}$	552.8	1.43
CC	1	500	491.1	192.9	$\underline{175.3}$	218.0	$\underline{252.1}$	$\underline{252.1}$	276.9	317.1	1.44
CC	1	1,000	655.9	302.8	$\underline{232.4}$	252.3	$\underline{394.9}$	$\underline{394.9}$	$\underline{394.9}$	409.4	1.70
CC	1.5	500	202.2	$\underline{109.5}$	127.5	184.3	$\underline{185.4}$	$\underline{185.4}$	196.5	237.9	1.69
CC	1.5	1,000	274.7	164.3	$\underline{155.1}$	200.4	$\underline{302.6}$	$\underline{302.6}$	$\underline{302.6}$	318.1	1.95

drical panel having the same boundary conditions and values of c and β^2. For each value of m in (24), there are an infinite number of vibration frequencies, and we denote the lowest of these by ω_m in Table 1. The lowest frequencies for $m = 1,2,3$, and 4 are presented in Table 1 for the optimal panels and the corresponding circular forms. (We note that the four frequencies listed for each case do not necessarily include the second, third, or fourth lowest frequencies of the panel; for example, the second lowest frequency for $m = 1$ may be lower than ω_2, ω_3, and ω_4.)

In these 12 examples, the fundamental frequency is associated with $m = 1$, i.e., one half-sine wave in the x direction. For the circular cylindrical panels, the lowest eigenvalue is distinct. This is also true for the optimal design in the case when the edges are clamped with $c = 2/3$ and $\beta^2 = 500$. In the other 11 optimal solutions, however, ω_1 is a double eigenvalue and the solution is bimodal. The vibration mode associated with ω_1 in the unimodal case is symmetric about $y = c/2$ (with no nodes). For the other optimal solutions, the governing modes can be written as a symmetric and anti-symmetric pair, or any linear combination of those functions. In the circular case the governing vibration mode is antisymmetric about $y = c/2$ (with one node) for the first six rows and last three rows of Table 1, and symmetric for the remaining three rows.

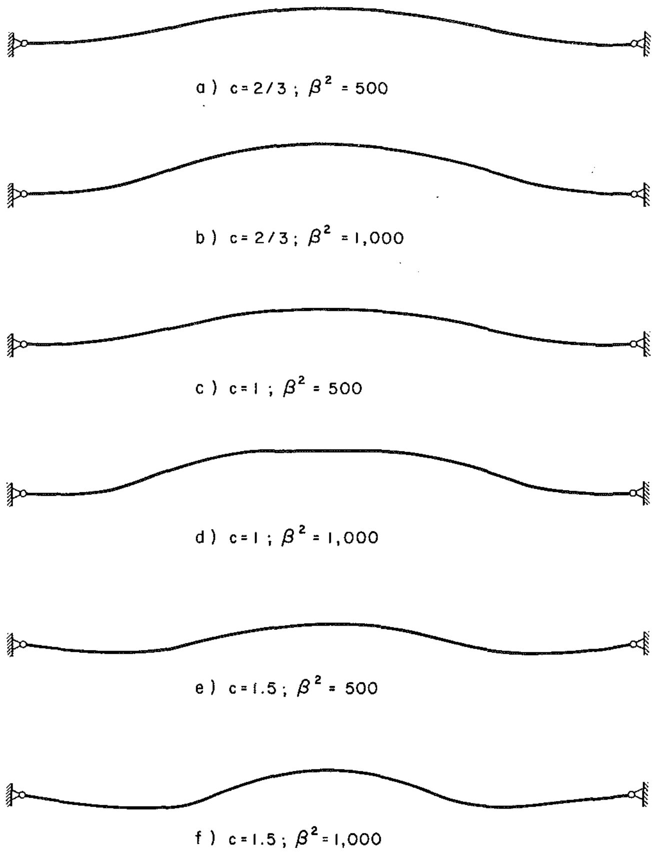
The last column of Table 1 lists the ratio of the optimal fundamental frequency to that of the corresponding circular cylindrical panel. In one case the optimal form is nearly circular and, therefore, the improvement is negligible. The largest increase shown is 12 percent, which occurs for the panel having a unimodal solution.

Results for the buckling problem are presented in Table 2 and Figs. 4 and 5. For most of these optimal forms, the panel does not lie completely above the base plane. In the case of clamped edges, the slope there is zero again.

Similarly to the notation in Table 1, we denote the lowest eigenvalue for a given value of m by p_m, and we list p_1, p_2, p_3, and p_4 for the optimal and circular cylindrical panels. In contrast to the vibration results, here none of the cases exhibits a multiple eigenvalue for a given value of m. Instead, eigenvalues for different values of m are sometimes equal to each other.

In Table 2, the buckling loads (i.e., the lowest eigenvalues over all values of m) are underlined. For the circular cylindrical panel, the buckling loads are distinct eigenvalues and the associated mode for each case has either two, three, or four half-sine waves in the x direction. For the optimal panels, modes with lower values of m tend to govern or join in a multimodal solution. As an example, in the first row of Table 2, the buckling mode is associated with $m = 3$ for the circular case, while the optimal solution is bimodal with $m = 2$ and $m = 3$. In the second row, p_4 is the buckling load for the circular form, while p_2 and p_3 again govern for the optimal panel.

We note that there are three trimodal solutions in Table 2, corresponding to the designs in Figs. 4(*d*), 5(*d*) and 5(*f*). In


a) $c = 2/3$; $\beta^2 = 500$

b) $c = 2/3$; $\beta^2 = 1{,}000$

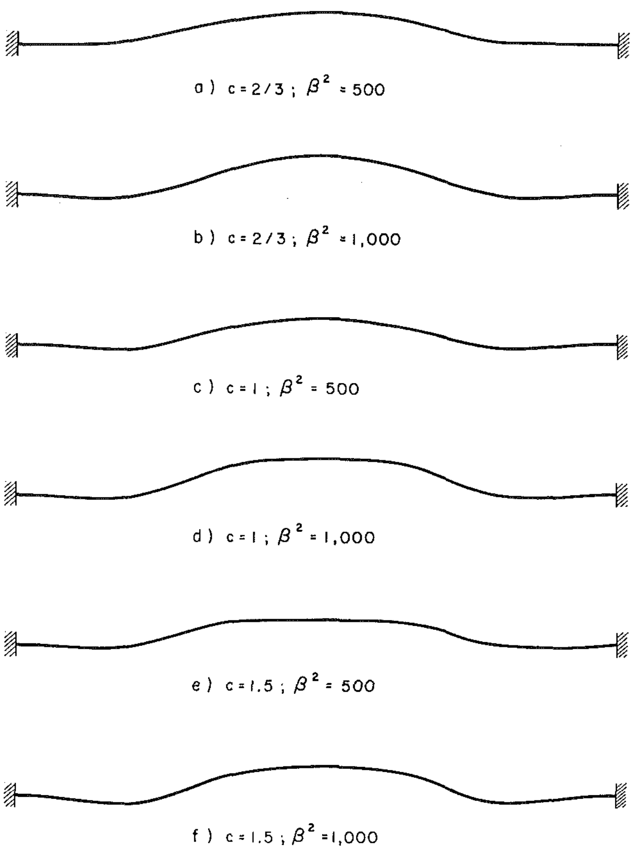
c) $c = 1$; $\beta^2 = 500$

d) $c = 1$; $\beta^2 = 1{,}000$

e) $c = 1.5$; $\beta^2 = 500$

f) $c = 1.5$; $\beta^2 = 1{,}000$

Fig. 4 Optimal forms for maximum buckling load with simply supported edges


a) $c = 2/3$; $\beta^2 = 500$

b) $c = 2/3$; $\beta^2 = 1{,}000$

c) $c = 1$; $\beta^2 = 500$

d) $c = 1$; $\beta^2 = 1{,}000$

e) $c = 1.5$; $\beta^2 = 500$

f) $c = 1.5$; $\beta^2 = 1{,}000$

Fig. 5 Optimal forms for maximum buckling load with straight edges clamped

these cases, there are buckling modes with one, two, and three half-sine waves in the x direction.

In some of the examples, the second eigenvalues for one or more of the governing mode numbers m are close to the buckling load. For example, if the straight edges are clamped with aspect ratio $c = 1$ and surface area parameter $\beta^2 = 1{,}000$, then the optimal buckling load is $p_1 = p_2 = p_3 = 394.9$, while the second frequencies for $m = 2$ and $m = 3$ are 396.8 and 396.2, respectively.

In all of the circular cases and most of the optimal cases listed in Table 2, the buckling modes associated with the buckling loads are symmetric about $y = c/2$ (with no nodes). The only anti-symmetric modes (with one node) are associated with the following eigenvalues of the optimal forms: $m = 2$ in the fourth, sixth, and last two rows, and $m = 3$ in the tenth row. When the optimal solution is multimodal, any linear combination of the individual buckling modes is also a solution of (11) and (12) at the buckling load.

The last column of Table 2 lists the ratio of the optimal buckling load to that of the corresponding circular cylindrical panel. The increase ranges from 8 percent for the first case to 95 percent for the last case.

During the buckling optimization procedure for a given case, different initial designs sometimes led to different solutions. In other words, there may be a number of local optima. For example, in the case of clamped edges with $c = 1$ and $\beta^2 = 500$, the maximum buckling load is $p_1 = p_2 = 252.1$. There is another solution of the governing equations (25), (26), and the optimality condition (27), for which $p_1 = 359.8$, $p_2 = p_3 = 248.5$, and $p_4 = 282.0$. Hence, this suboptimal design is also bimodal. Its form is similar to that in Fig. 5(c) except that its values $z(y)$ are somewhat higher.

An interesting feature occurred in the case of clamped edges with $c = 2/3$ and $\beta^2 = 500$. The optimal solution is unimodal with buckling load $p_3 = 369.1$. However, there is a suboptimal solution which is bimodal, with $p_1 = 1392$, $p_2 = 483.8$, $p_3 = p_4 = 364.8$, $p_5 = 413.1$, and $p_6 = 494.7$. Its form lies slightly above that shown in Fig. 5(a). This demonstrates that it is possible to have a suboptimal design with higher modality (i.e., a larger number of coincident eigenvalues) than the optimal design.

Concluding Remarks

We have considered the problems of maximizing the fundamental vibration frequency and the buckling load of a shallow, cylindrical panel by varying its form. The boundary conditions, material, thickness, and surface area of the panel are specified. Numerical results are obtained for a number of cases. The optimal solutions are unimodal, bimodal, or trimodal. In comparison with the corresponding circular cylindrical panel, the increase in fundamental frequency is fairly small (less than 12 percent in the examples), while the increase in buckling load can be substantial (8–95 percent in the examples).

When the straight edges of the panel are assumed to be clamped, the optimal forms have zero slope there. In Figs. 2 and 3, for maximum fundamental frequency of natural vibrations, the optimal panel lies above the base plane. In Figs. 4 and 5, for maximum buckling load, the optimal panels usually lie below the base plane near the straight edges and above it in the central region, and tend to resemble corrugated plates.

In the numerical solution procedure, we use a finite difference method to transform the differential equations into nonlinear algebraic equations in the displacement, stress function, and design function. After some iterations, the equations (including the optimality condition) are solved by a quasi-Newton method. It appears that there are many local maxima, especially in the buckling problem. By starting from a variety of initial shapes, we have attempted to obtain the globally optimal solution for each example. In some cases, the best solution does not have the highest modality of all the local solutions.

We did not investigate the imperfection-sensitivity of the optimal panels. For shells, small imperfections may either decrease or increase natural vibration frequencies [23]. In the buckling problem, when the optimal solutions are bimodal or trimodal, we expect them to be more imperfection-sensitive than the corresponding circular cylindrical panels [24]. However, the increase in buckling load may be more advantageous than a greater sensitivity to imperfections.

Acknowledgments

The first author was supported by the National Science Foundation under Grant CEE-8210222.

References

1 Bushnell, D., "PANDA – Interactive Program for Minimum Weight Design of Stiffened Cylindrical Panels and Shells," *Computers and Structures*, Vol. 16, 1983, pp. 167–185.

2 Rehfield, L. W., and Hallauer Jr., W. L., "Edge Restraint Effect on Buckling of Compressed Curved Panels," *AIAA Journal*, Vol. 6, 1968, pp. 187–189.

3 Hsueh, P. S., and Chajes, A., "Buckling of Axially Loaded Cylindrical Panels," *Journal of the Engineering Mechanics Division, Proceedings, ASCE*, Vol. 97, 1971, pp. 919–933.

4 Sobel, L. H., Weller, T., and Agarwal, B. L., "Buckling of Cylindrical Panels Under Axial Compression," *Computers and Structures*, Vol. 6, 1976, pp. 29–35.

5 Kurt, C. E., and Boyd, D. E., "Free Vibrations of Noncircular Cylindrical Shell Segments," *AIAA Journal*, Vol. 9, 1971, pp. 239–244.

6 Srinivasan, R. S., and Bobby, W., "Free Vibration of Non-Circular Cylindrical Shell Panels," *Journal of Sound and Vibration*, Vol. 46, 1976, pp. 43–49.

7 Massalas, C., Soldatos, K., and Tzivanidis, G., "Free Vibrations of Non-Circular Cylindrical Panels With Arbitrary Boundary Conditions," *Journal of Sound and Vibration*, Vol. 69, 1980, pp. 491–495.

8 Durban, D., and Singer, J., "Buckling of Cylindrical Panels Under Non-Uniform Axial Compression," *Israel Journal of Technology*, Vol. 11, 1973, pp. 9–16.

9 Krishnamoorthy, G., and Narang, B. S., "Buckling of Open Cylindrical Shells Under Combined Compression and Bending Stress," *AIAA Journal*, Vol. 15, 1977, pp. 322–328.

10 Koiter, W. T., "Buckling and Post Buckling Behaviour of Cylindrical Panel Under Axial Compression," Report S476, Vol. 20, National Luchtvaartlaboratorium, Amsterdam, 1956.

11 Brush, D. O., and Almroth, B. O., *Buckling of Bars, Plates, and Shells*, McGraw-Hill, New York, 1975.

12 Tvergaard, V., "Buckling of Elastic-Plastic Cylindrical Panel Under Axial Compression," *International Journal of Solids and Structures*, Vol. 13, 1977, pp. 957–970.

13 Knightly, G. H., and Sather, D., "Nonlinear Buckling and Stability of Cylindrical Panels," *SIAM Journal on Mathematical Analysis*, Vol. 10, 1979, pp. 389–403.

14 Weller, T., "Combined Effects of In-Plane Boundary Conditions and Stiffening on Buckling of Eccentrically Stringer-Stiffened Cylindrical Panels," *Computers and Structures*, Vol. 14, 1981, pp. 427–442.

15 Bauld Jr., N. R., and Khot, N. S., "A Numerical and Experimental Investigation of the Buckling Behavior of Composite Panels," *Computers and Structures*, Vol. 15, 1982, pp. 393–403.

16 Patnaik, S. N., and Sankaran, G. V., "Optimum Design of Stiffened Cylindrical Panels With Constraint on the Frequency in the Presence of Initial Stresses," *International Journal of Mechanical Sciences*, Vol. 10, 1976, pp. 283–299.

17 Zimmerman, R., "Mass Minimization of Composite Material Cylindrical Shells and Curved Panels With Buckling Constraints," *Optimization Methods in Structural Design*, H. Eschenauer and N. Olhoff, eds., Bibliographisches Institut, Mannheim, Federal Republic of Germany, 1983, pp. 438–443.

18 Plaut, R. H., Johnson, L. W. and Parbery, R., "Optimal Forms of Shallow Shells With Circular Boundary. Part 1: Maximum Fundamental Frequency," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, 1984, pp. 526–530.

19 Plaut, R. H., and Johnson, L. W., "Optimal Forms of Shallow Shells With Circular Boundary. Part 2: Maximum Buckling Load," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, 1984, pp. 531–535.

20 Bisplinghoff, R. L., and Ashley, H., *Principles of Aeroelasticity*, Wiley, New York, 1962.

21 Bendsøe, M. P., Olhoff, N., and Taylor, J. E., "A Variational Formulation for Multicriteria Structural Optimization," *Journal of Structural Mechanics*, Vol. 11, 1983, pp. 523–544.

22 Olhoff, N., and Taylor, J. E., "On Structural Optimization," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, 1983, pp. 1139–1151.

23 Rosen, A., and Singer, J., "Influence of Asymmetric Imperfections on the Vibrations of Axially Compressed Cylindrical Shells," *Israel Journal of Technology*, Vol. 14, 1976, pp. 23–36.

24 Thompson, J. M. T., and Hunt, G. W., *A General Theory of Elastic Stability*, Wiley, New York, 1973.

APPENDIX

The boundary conditions on the stress function along the straight edges are derived in this Appendix.

We denote strains by ϵ_X, ϵ_Y, γ_{XY}, midsurface displacements by U, V, W, and stress resultants by N_{XX}, N_{YY}, N_{XY}. In the buckling problem, these represent incremental quantities with respect to the uniform prestressed state.

If $Z(Y)$ is the form of the middle surface, the strain-displacement relations are [19]

$$\begin{aligned}&\epsilon_X = U_{,X} + (W_{,X})^2/2, \ \epsilon_Y = V_{,Y} + Z' W_{,Y} + (W_{,Y})^2/2,\\ &\gamma_{XY} = U_{,Y} + V_{,X} + Z' W_{,X} + W_{,X} W_{,Y}\end{aligned} \tag{28}$$

and the stress-strain relations are

$$\begin{aligned}&\epsilon_X = (N_{XX} - \nu N_{YY})/(Eh), \ \epsilon_Y = (N_{YY} - \nu N_{XX})/(Eh),\\ &\gamma_{XY} = 2(1+\nu) N_{XY}/(Eh),\end{aligned} \tag{29}$$

where subscripts X, Y following a comma denote partial derivatives. The Airy stress function F is defined by

$$N_{XX} = F_{,YY}, N_{YY} = F_{,XX}, N_{XY} = -F_{,XY}. \tag{30}$$

At the straight edges $Y = 0, b$, we assume simply supported or clamped boundary conditions, which implies

$$W = W_{,YY} = 0 \text{ or } W = W_{,Y} = 0, \tag{31}$$

respectively. For a third condition, we assume $N_{XY} = 0$, which allows tangential displacements. Hence, from (30),

$$F_{,XY} = 0. \tag{32}$$

Finally, we assume there are no normal displacements, i.e., $V = 0$.

It follows from $V = W = 0$ that $V_{,X} = W_{,X} = 0$ along the edges $Y = 0, b$. Since $N_{XY} = 0$ implies $\gamma_{XY} = 0$ from (29), we then obtain $U_{,Y} = 0$ from (28), and thus $U_{,YX} = 0$. Now, if we differentiate ϵ_X in (28) with respect to Y, we get $\partial\epsilon_X/\partial Y = 0$. We then differentiate ϵ_X in (29) with respect to Y and use (30) and the previous result, which leads to

$$F_{,YYY} - \nu F_{,XXY} = 0. \tag{33}$$

Differentiating (32) with respect to X yields $F_{,XXY} = 0$, and (33) then furnishes the fourth boundary condition in the form

$$F_{,YYY} = 0. \tag{34}$$

N. J. Salamon[1]
Professor,
Department of Engineering Science
and Mechanics,
The Pennsylvania State University,
University Park, Pa. 16802
Mem. ASME

T. P. Pawlak[1]
Engineer,
Swanson Analysis Systems, Inc.,
Houston, Pa. 15342

F. F. Mahmoud
Professor and Vice Dean,
Mechanical Engineering Department,
Zagazig University,
Cairo, Egypt

Plates in Unilateral Contact With Simple Supports: Pressure Loading

The response of a square plate, simply and unilaterally supported, to pressure loading is numerically treated. The support system consists of discrete elastic springs whose stiffnesses range from near-rigid to compliant character. It is found that, except for rather low support stiffnesses, the plate will lift off the foundation. After demonstrating good agreement with a recent analytical work, the deflections and shear fields are provided. The response mode changes dramatically as the supports approach rigidity.

Introduction

The response of plates unilaterally supported by foundations to transverse loading is complicated by the need to determine the contact region as well as the equilibrating reactions. In part this accounts for the scarcity of literature pertaining to the matter, particularly for the case where the foundation "simply" supports the plate periphery. To simplify matters, Keer and Mak (1981) treated an infinitely extended plate under a concentrated load near a corner. Unfortunately, their solution is not applicable to a centrally loaded square plate because the load is not sufficiently near to a corner. Dempsey, Keer, Patel, and Glasser (1984) provide a complicated solution for a square plate under pressure (a partial solution was presented by Dempsey, Keer, and Patel, 1983). Results are only given for the deflections, and it is not a trivial matter to generate those for the shearing force. In the aforementioned works, the methodology involves integral equation and series formulations, and the lineal supports are rigid. Although rigid support systems are commonly associated with civil engineering, elastic support systems may often be closer to reality and indeed applications cover a spectrum of support stiffnesses, for example, gait devices in biomechanical engineering, (i.e., Harris, Salamon, and Weber, 1981).

The present paper is a numerical study of plates under uniform pressure loads that are unilaterally supported around the periphery by discrete elastic springs. After formulating the problem, it is demonstrated that results for plates on near-rigid springs tally with those found by Dempsey et al., (1984). Then the behavior of elastic systems is discussed. From these discrete results extrapolations to continuously supported systems are made. It is found that the response mode changes dramatically as the elastic supports approach rigidity. Moreover, the shear load varies considerably across the supports. However, avenues are open to optimize the distribution. The problem is important because liftoff will occur even at low relative stiffnesses.

The Problem

Formulation. An iterative algorithm is used in conjunction with the finite element method for linear systems. The problem consists of a square plate of dimension L and thickness t resting on springs distributed about its periphery with spacing a. The plate is discretized into classic plate elements (called the ACM element in Cook, 1981) which are four node squares of spacing a; the foundation springs are linear with stiffness k_s. If $\{\mathbf{w}\}$ contains the deflections $\{\mathbf{d}\}$ and the rotations $\{\theta_x, \theta_y\}^T$, $\{\mathbf{R}\}$ is the vector of loads and m is an increment counter, then upon assembly the linear system can be written as

$$D[K^{(m)}]\{\mathbf{w}\}^m = \{\mathbf{R}\} \tag{1}$$

provided that the stiffness matrix $[K]$ posesses diagonal elements associated with degree of freeedoms $\{\mathbf{d}\}$ of the form

$$k^d_{ii} = k_{ii} + \alpha_i (k_s/D) \tag{2}$$

where $\alpha_i = 1$ or 0, depending on whether a spring is attached to the ith degree of freedom or not. Here $D = Et^3/12(1-\nu^2)$ is the plate rigidity with Young's modulus E and Poisson's ratio ν. The general k_{ij} are plate element stiffnesses with D factored out and are functions of element size and ν. Furthermore, if foundation/plate combinations are chosen such

---

[1] Formerly with the Mechanical and Aerospace Engineering Department, West Virginia University, Morgantown, W. Va.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami Beach, Fla., November 17-22, 1985, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October, 1984; final revision, April, 1985. Paper No. 85-WA/APM-21.

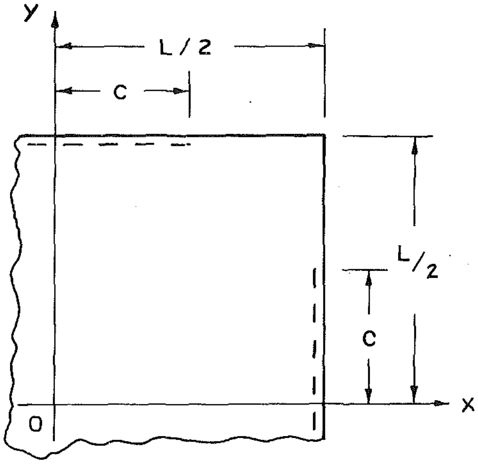


Fig. 1 Quarter of the plate modeled where c denotes the half-length of contact

that k_s/D is the same, then the product $\{\mathbf{wD}\}$ is the same for these combinations.

For foundation/plate combinations, further utility is gained by considering the foundation as continuous and including the plate geometry. In this way a nondimensional relative stiffness κ is defined as

$$\kappa = (k_s/a)\,(L^3/D) \tag{3}$$

Hence $[K^{(m)}] = [K(\nu, \kappa\,; a, L; m)]$ and in turn $\{\mathbf{wD}\}^{M}$, M the last iteration, is a more general solution than equation (1) at first glance implies. It may be noted that κ is analogous to the quantity $(k_s/4EIa)^{1/4}\ L$, I the second moment of area and L the beam length, which originates in the theory of beams on elastic (springs) foundations.

In the initial state, one physically views the weightless plate resting upon all of the springs. However for the simulation not all springs need be attached; the procedure only requires sufficient initial constraints, boundary conditions and spring attachments, to avoid rigid body motion.

For any iteration m, the procedure solves (1 for $\{\mathbf{wD}\}$ consistent with the current constraint prescription. As described below, the next site i for a contact or separation is chosen and α_i set to 1 or 0, which updates $[K^{(m)}]$ to $[K^{(m+1)}]$. If i is null, there are no new sites, hence $m = M$ and $\{\mathbf{wD}\}^M$ is the solution; otherwise the procedure repeats.

Since the springs are independent of each other, contacts and separations are sensed kinematically directly from the contents of $\{\mathbf{wD}\}$. For the next event, site i is chosen by satisfying the criterion

$$d_i = \mathrm{Max}(\pm d_j),\ 1 \le j \le NC \tag{4}$$

where NC is the number of candidate sites, $d_j > 0 \Rightarrow$ separation and $d_j \le 0 \Rightarrow$ contact, and the + sign is used for sites currently connected and the − sign for those currently disconnected. Should several sites yield the same value d_i, the minimum value i is chosen.

Control of the procedure is maintained through a data structure which keeps track of the contact status of each candidate constraint and through a decision process enhanced by theoretical precepts for receding contact problems, Dundurs and Stippes (1970). A detailed discussion of the algorithm as applied to the solution of beam and plate problems is provided in a journal dedicated to numerical methods by Mahmoud, Salamon, and Pawlak, (1985).

Numerical Model. One quarter of the square plate is modeled, Fig. 1, using 81 elements to fill a 10 × 10 nodal mesh. Symmetry is enforced at the centerlines by setting the rotations $\theta_x = 0$ along x and $\theta_y = 0$ along y. Springs underlie the plate edges $x = L/2$ and $y = L/2$, but only those within length c remain in contact under load over each half-side. Springs located at symmetry points $(x,y) = (0,L/2)$ and $(L/2,0)$ are connected initially and their stiffnesses halved.

The important field quantity generated by the computer is the deflection d, whence along a side the contact half-length c is computed as

$$c = (N-1)a + \xi \tag{5}$$

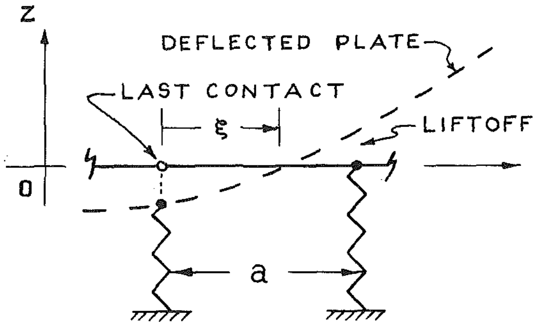


Fig. 2 The end-of-contact region

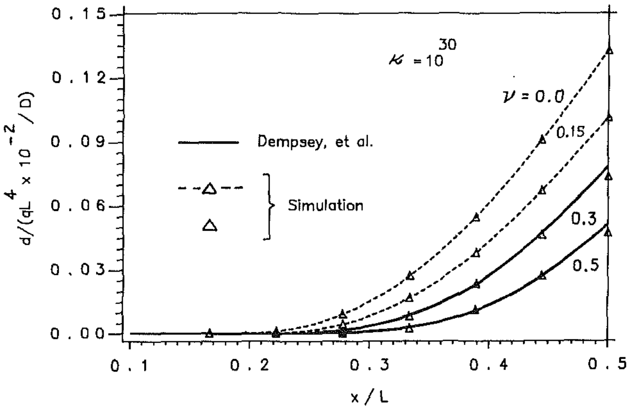


Fig. 3 Deflections along edge for plates on rigid supports under uniform pressure

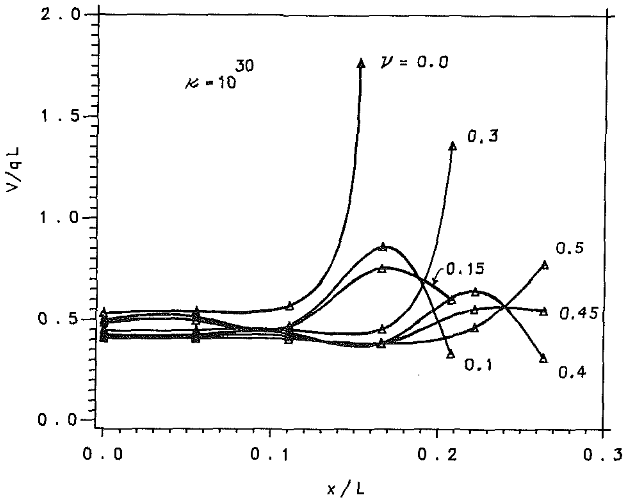


Fig. 4 Shear distributions along edge for plates on rigid supports under uniform pressure

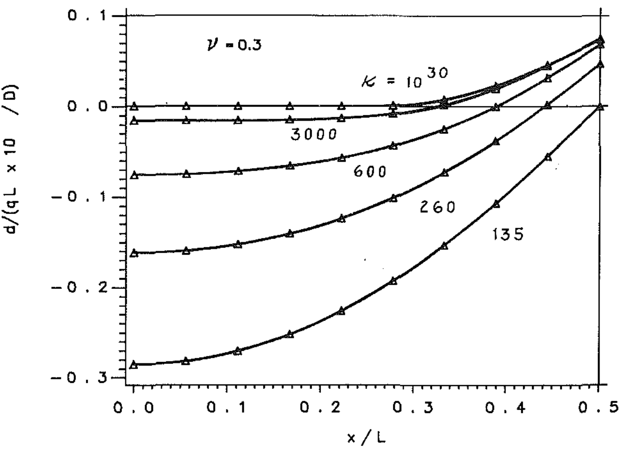


Fig. 5 Deflections along edge for plates on elastic supports under uniform pressure

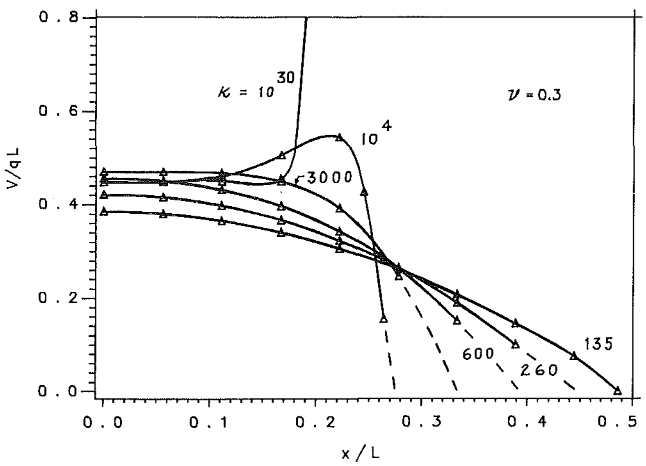


Fig. 6 Shear distributions along edge for plates on elastic supports under uniform pressure

where N is the number of nodes along a half-side in contact, and ξ is a linearly interpolated extension of contact between the last depressed spring and first lifted-off plate node, Fig. 2. In turn the loads in compressed springs are determined and converted into continuous shear distributions V by averaging the spring loads across an element span a or half-span $(a/2)$; specifically at the centroid of the distribution

$$V = k_s\, d/a \tag{6}$$

except at the end-of-contact and $\xi < a/2$; then it is taken as 2 $k_s\, d/a$. In this form the shear force per unit of length provides an approximation to the common case of continuous support. However if the discrete values are desired, one need only invert (6) or its end-of-contact counterpart to obtain the concentrated reactions, e.g., solve for $k_s d$.

Results

Two types of loading are studied: uniform pressure distributed over the entire plate and a concentrated load P located at the origin of coordinates (Fig. 1). In the subsequent figures, discrete points are plotted as triangles and connected together by a cubic spline curve-fit. Data points for shear are located at the centroids of their distribution. The focus here is on pressure loadings; a sequel will cover concentrated loading.

Rigid Supports. The behavior of a plate on rigid supports is approximated by setting the relative stiffness, and in effect the spring stiffness, to a very large value; in the present case $\kappa = 10^{30}$. It should be noted that despite this extreme, the foundation is not "infinitely" rigid and although the response to load is small, it still retains its deformable character. With the plate size, element discretization, and relative stiffness set, the deflection profiles and shear distributions along the plate edge are computed for the range of Poisson's ratio $0 \leq \nu \leq 0.5$.

Figures 3 and 4 show the deflection profiles and shear distributions generated by a uniform pressure load q. The deflections for $\nu = 0.3$ and 0.5 are seen to agree very well with those provided by Dempsey et al. (1984), that for $\nu = 0.1$ agrees equally well, but is not shown because it spoils the figure arrangement.

The effect of increasing Poisson's ratio is to stiffen the plate and in turn extend the length of contact (Fig. 3). Yet even with $\nu = 0.5$ almost half the plate edge ($\approx$ 45 percent) lifts off the supports. Although the displacement profiles illustrate the behavior, the shear distributions elucidate it. By concentrating upon the values of shear load near the ends-of-contact in Fig. 4, one can extract the following: (*i*) for $\nu = 0$ the plate bears down hard on the last contacting support, however for $\nu = 0.1$ the plate stiffens sufficiently to extend the contact and bear down lightly upon the next suppport; (*ii*) for ν increasing to 0.3, pressure increases on this new support until with $\nu = 0.4$ the plate siffness increases sufficiently to again cause the contact to grow and the process repeats.

Elastic Supports. In order to view the role relative stiffness plays in the behavior of the plate, the Poisson's ratio is set to 0.3. Then values for κ are selected from a range such that the contact length varies from least to full extent. In the figures for the shear distributions, those for intermediate values of κ are extrapolated from the last datum point to zero using dashed lines.

The deflections and shear distributions along an edge of the plate are shown in Figs. 5 and 6. For very soft (relative) support the plate is in full contact, but incipient liftoff occurs when $\kappa = 135$. With increasing κ plate liftoff grows asymptotically (as will be discussed subsequently) to a maximum limit. The shear distributions illustrate the increased burden sustained by inboard supports as the contact length decreases, displaying a maximum at $\kappa = 3000$. But for greater values of κ the distribution goes through a transition characterized by a decrease in shear near the midside of the plate and a compensatory increase near the end-of-contact; for $\kappa = 10^4$ the increase reverses, tapering to zero as the plate lifts off of the supports, however for $\kappa = 10^{30}$ the increase appears singular (but its finiteness is clearly visible in Fig. 4). Hence when the foundation is relatively very stiff the last support in contact bears the brunt of the load. On the other hand, for relatively soft foundations, greater load is borne by the inboard supports. Clearly near-rigid and rigid supports markedly change the character of the shear distribution.

Extent of Contact

The growth of contact for a discrete system occurs in jumps or steps with respect to continuous change of an influential parameter. This is clearly conveyed by changes in the lengths over which the shear is distributed in Fig. 4. The subsequent sections, with the exception of that on incipient liftoff, are concerned with the extent of contact with respect to change in either relative stiffness or Poisson's ratio; it is convenient, and perhaps more relevant, to present these results as continuous phenomena. Hence they are approximations for the problem of plates on continuous support. Where the extent of contact involves many discrete supports, the accuracy should be good,

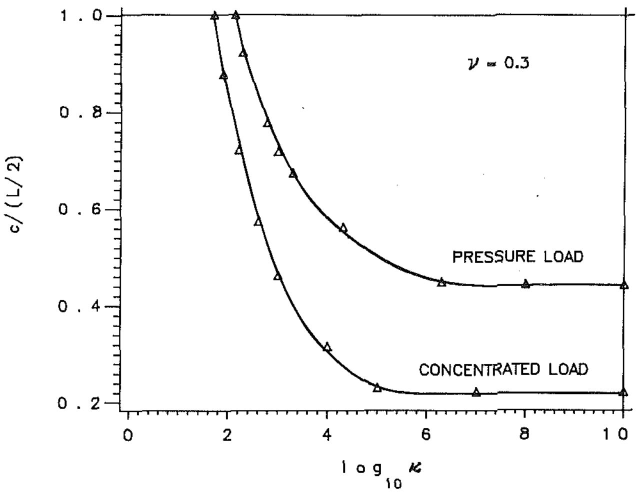


Fig. 7 Extent of contact with respect to the relative stiffness

otherwise in this context it should be held suspect. Clearly improved accuracy will result from use of higher densities of discrete supports.

Incipient Liftoff. It is of particular interest to know whether or not liftoff will occur because if it does not the problem is a linear one. The objective then is to find relative stiffness thresholds, denoted by κ_0, for the incipience of plate liftoff. This is done by iterating with respect to κ for each of five different Poisson's ratios. It turns out that the data can be approximated by simple polynomials to an accuracy of three digits.

Incipient liftoff will occur for pressure loading if

$$\kappa_0 = 97.0\,(1.0 + \nu + \nu^2) \tag{7}$$

Although κ_0 is nicely behaved with respect to ν, the contact length is very sensitive to changes in κ in this "soft support" regime as will be seen in the next section. Hence the determination of (7) required considerable effort.

Effect of Relative Stiffness. In situations where $\kappa > \kappa_0$, liftoff will occur and the problem becomes nonlinear. However it is helpful if the contact length is known. To this end Fig. 7 provides an estimate of the extent of contact with respect to the relative stiffness but, however, does so only for Poisson's ratio equal to 0.3.

From the figure it is seen that near the liftoff threshold ($c/(L/2) = 1$) the contact length is very sensitive to changes in relative stiffness. Alternatively as κ approaches the condition of rigid support ($\kappa \rightarrow \infty$), contact length decreases to an asymptotic value; $c/(L/2) = 0.444$ for pressure load and, for comparison, 0.222 for the concentrated load. Furthermore these asymptotes are reached when $\kappa \approx 5 \times 10^6$ for the pressure load and $\kappa \approx 5 \times 10^5$ for the concentrated load. A simple expression to fit the data was not found; the rather extensive range of relative stiffness is a contributing reason.

Effect of Poisson's Ratio: Rigid Support. For the important case when the supports can be considered rigid, it is possible to set bounds upon the contact length as a function of Poisson's ratio. This is done by using information from Figs. 4 and 6 together with the following argument: for continuous support the contact length will be greater than that achieved for values of Poisson's ratio that produce maximum values of shear force (for example, at $\nu = 0.0, 0.3, 0.5$ in Fig. 4). Alternatively the contact length will be less than that achieved for values of Poisson's ratio that produces minimum values of shear force (e.g., at $\nu = 0.1$ and 0.4 Fig 4). Taken in this manner, the extracted data are found to vary almost linearly in agreement with the results for pressure load given by Dempsey et al. (1984).

The data are easily fit to linear expressions. For the pressure load, these are

$$\frac{c}{L/2} = \left\{ \begin{array}{l} 0.370\nu + 0.408 \text{ (upper bound)} \\ 0.429\nu + 0.326 \text{ (lower bound)} \end{array} \right\}, 0 \leq \nu \leq 0.5 \tag{8}$$

A linear fit to the data of Dempsey et al (1984) gives

$$c/(L/2) = 0.415\nu + 0.343, 0 \leq \nu \leq 0.5 \tag{9}$$

so that indeed (8) brackets (9) and the maximum error between (8) and (9) is 5 percent for the lower and 19 percent for the upper bounds; the lower bound of (8) yields a maximum error of 4.4 percent with their exact data.

Conclusion

The results generated for plates on independent elastic supports are sufficiently general to permit design optimization of such systems. The results demonstrate accuracy to problems involving rigid support and provide insight into those involving elastic support. It is also possible to project the effect of continuous, lineal support.

Optimal design of unilaterally supported systems is achieved through rational location of supports and prescription of system properties. On the one hand, for simply-supported plates, particular advantage can be gained if the relative stiffness is large. Then high singular-like shear loads may be avoided by shifting support locations; alternatively, the shear load may be more evenly distributed by varying the Poisson's ratio of the plate. While the former is readily possible, the latter is not impracticable especially if one considers the plate to be of a designable composite material. On the other hand, if the relative stiffness is not large, optimal distribution of shear may be achieved by tailoring the stiffness of individual supports as well as by varying the spacing between them.

The importance of the effects of elastic support resides with the low value of relative sitffness necessary for full contact. Hence a wide range of practical problems exist for which it would be rather unusual for plate liftoff to tend not to occur. One interesting effect not directly portrayed in the results is that the curvature of the plate changes sign for high values of relative stiffness. As viewed on edge, for stiff supports the curvature at midside is convex (frowning) and at the ends concave (smiling), whereas for softer supports it is concave all along the edge. The implication is that the prescription of ideal, infinitely rigid supports prevents this natural mode of behavior and contributes to the generation of a mathematical singularity.

It is possible to project the effect of continuous, lineal support upon the system. Discontinuous jumps in contact length with changes in Poisson's ratio or relative stiffness inherent in the discrete system will be eliminated; the relationships will be smooth and monotonic. Moreover for large values of relative stiffness, the shear distribution along the edge would rise from moderate values at midside to a peak and smoothly, albeit sharply, drop to zero at the ends-of-contact. In contrast, if the continuous support is infinitely rigid as described above, the reaction fields would display a singular behavior at the ends-of-contact as discussed by Dempsey et al (1984).

The effect of a centrally located, concentrated load are similar in form to that for pressure, but considerably different in magnitude and the extent of resulting distributions. This is evident in Fig. 7 by the short extent of contact for large values of relative stiffness. Further details will be discussed in a sequel to this paper.

Acknowledgments

The results reported here were obtained in the course of research supported by the National Science Foundation. The authors are also grateful to John P. Dempsey for the timely communication of new results.

References

Cook, Robert D., 1981, *Concepts and Applications of Finite Element Analysis* (2nd ed.), Wiley & Sons, New York.

Dempsey, J. P., Keer, L. M., and Patel, N. B., 1983, "Contact Between Plates and Unilateral Supports," in *Recent Advances in Engineering Mechanics and Their Impact on Civil Engineering Practice*, W. F. Chen and A. D. M. Lewis, eds., *Proc. of the Fourth Engineering Mechanics Div. Speciality Conf.*, ASCE, Purdue Univ., West Lafayette, Indiana, 23-25, May.

Dempsey, J. P., Keer, L. M., Patel, N. B., and Glasser, M. L., 1984, "Contact Between Plates and Unilateral Supports," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 324-328.

Dundurs, J., and Stippes, M., 1970, "Role of Elastic Constants in Certain Contact Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 965-970.

Harris, G. F., Salamon, N. J., and Weber, R. C., 1981, "Effects of Subject Position on Balance Platform Measurements," ASME *Journal of Biomechanical Engineering*, 103, pp. 213-216.

Keer, L. M., and Mak, A. F., 1981, "Loss of Contact in the Vicinity of a Right-Angle Corner for a Simply Supported, Laterally Loaded Plate," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 597-600.

Mahmoud, F. F., Salamon, N. J., and Pawlak, T. P., 1985, "Simulation of Structural Elements in Receding/Advancing Contact," *Computers and Structures*, in press.

J. P. Dempsey
Assistant Professor.
Mem. ASME

Hui Li
Graduate Student.

Department of Civil and
Environmental Engineering,
Clarkson University,
Potsdam, N.Y. 13676

Rectangular Plates on Unilateral Edge Supports: Part 1—Theory and Numerical Analysis

The corners of a simply supported, laterally loaded rectangular plate must be anchored to prevent them from lifting off the supports. If no such anchors are provided, and the supports are unilateral or capable of exerting forces in one direction only, parts of the plate will bend away from the supports upon loading. The loss of contact when uplift of laterally loaded rectangular plates is not prevented is examined in this paper. Arbitrary centrally symmetric loading is considered. Finite integral transforms convert the coupled dual-series equations that result from the Levy-Nadai approach to two coupled singular integral equations. Different solution methods are applicable for sagged and unsagged supports; these two numerical approaches are discussed in detail.

Introduction

The corners of a simply supported, laterally loaded rectangular plate must be anchored to prevent them from lifting off the supports (Timoshenko and Woinowsky-Krieger, 1959). If no such anchors are provided, and the simple supports are unilateral or capable of exerting forces in one direction only, parts of the plate near the corners will bend away from the supports upon loading. The tendency of a laterally loaded, unilaterally constrained, rectangular plate to separate from its simple supports motivates one to consider the actual extent of contact.

The present paper examines the natural contact of centrally loaded rectangular plates resting on unilateral supports; each support is viewed as a unilateral constraint allowing only upward motion of the plate. It is assumed that two opposite supports have sagged by equal amounts. The amount of sag is taken as uniform, so that the sagged supports lie in a plane parallel to the other supports. The plate will touch the sagged supports only when the loading has reached a certain level. If the loading is increased further, the contact between the plate and the sagged supports will spread or advance. With no anchors at the corners, the contact between the plate and the unsagged supports will decrease with increasing load. The contact problem for the case of sagged supports is therefore one of advancing contact (Dundurs and Stippes, 1970). For the case of a zero amount of sag, the extent of contact will diminish immediately upon loading; it therefore follows that this contact problem is one of receding contact and hence the extent of contact between the plate and the supports is independent of the level of loading and the support reactions are proportional to the load. While both types of contact problem are treated here within the same formulation, the aforementioned characteristics mean that a different solution method applies for the case of a finite amount of sag than for the case of a zero amount of sag.

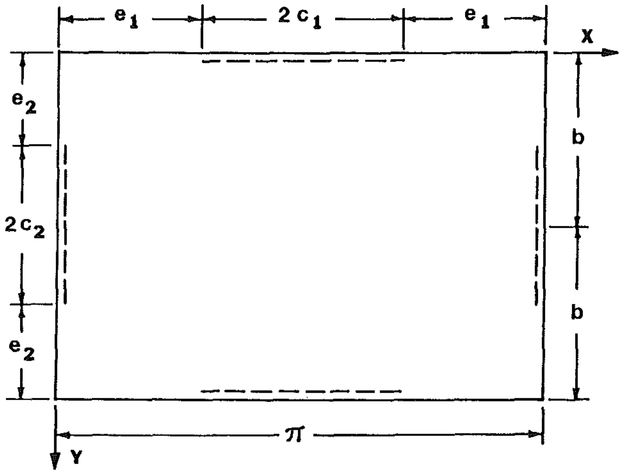


Fig. 1 Rectangular plate on unilateral edge supports

The plate geometry and coordinate system are shown in Fig. 1. The coordinates and dimensions shown are scaled by the factor π/a, where a is the actual plate length in the $\bar{x}$-direction. The actual (barred) coordinates are, for instance, $\bar{x} = ax/\pi$ and $\bar{y} = ay/\pi$. Without loss of generality it is assumed that the two sagged edge supports are those parallel to the x-axis. The amount of sag between these supports and the undeformed plate is defined by the constant distance δ. On contact, each simple support rigidly resists displacement in the positive z-direction only. The loading distributed over the surface of the plate is assumed to be symmetric about $x = \pi/2$, $y = b$ and, in terms of the transformed coordinates, is given here by the expression

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami Beach, Fla., November 17-22, 1984 of the 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. 10077, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March, 1985; final revision, June, 1985. Paper No. 85-WA/APM-23.

$$q(x,y)=Q\sum_{m=1,3,\ldots}^{\infty}\ \sum_{n=1,3,\ldots}^{\infty} a_{mn}\sin(mx)\sin(n\pi y/2b) \qquad (1)$$

where Q is an appropriate load factor for the particular loading being treated; $a_{mn}\equiv 0$ for m or n even. Because of the symmetry of the deflection function that results, there are only two unknown contact lengths: e_1 and e_2 (Fig. 1). The above loading is supplied in the positive z-direction, with the corresponding deflection given by $w(x, y)$. The differential equation for the deflection of the plate is

$$D\nabla^4 w=q(a/\pi)^4,\quad D=Eh^3/12(1-\nu^2) \qquad (2a,b)$$

the moments and supplemented, or Kirchhoff, shearing forces are given by

$$M_x=-(\pi/a)^2 D(w,_{xx}+\nu w,_{yy}) \qquad (3)$$

$$M_y=-(\pi/a)^2 D(w,_{yy}+\nu w,_{xx}) \qquad (4)$$

$$M_{xy}=-M_{yx}=(\pi/a)^2 D(1-\nu)\,w,_{xy} \qquad (5)$$

$$V_x=-(\pi/a)^3 D(w,_{xx}+(2-\nu)\,w,_{yy}),_x \qquad (6)$$

$$V_y=-(\pi/a)^3 D(w,_{yy}+(2-\nu)\,w,_{xx}),_y \qquad (7)$$

while the corner force at $x = 0, y = 0$ is given by

$$R=2M_{xy}|_{x=0,y=0} \qquad (8)$$

The bending of partially supported uniformly loaded rectangular plates with anchored corners has been considered by Kiattikomol et al. (1974) and Dundurs et al. (1974) for simple support conditions involving unsagged and sagged supports, respectively. The loss of contact in the vicinity of a right-angled unanchored corner for a simply supported uniformly loaded quarter infinite plate has been examined by Keer and Mak (1981). The extent of contact between a uniformly loaded square plate resting on simple supports was solved by Dempsey et al. (1984). The solution technique and mathematical manipulations used here overlap, to a large extent, those used in the latter paper. In order to avoid undue duplication of many expressions, the results in the foregoing work will be cited freely.

The approach used in the present paper gives rise to four coupled series equations that are solved using finite integral transforms. The particular type of transform used is governed by the strength of the singularity to be allowed at each point where the simple support changes to a free edge (hereafter called a transition point). The nature of the singularities in the bending fields at these points was first revealed by Williams (1952). Although this analysis treats finite plates, the infinite plate solution by Keer and Mak (1981) provides the correct singularity to be used in the vicinity of the mixed condition. The local behavior at the transition point must be the same for both finite and infinite plates, as the solution by Dempsey et al. (1984) verified. In conclusion, if the plate seeks its natural contact with either unsagged or sagged supports, no singularity in the moments can be allowed at the transition points.

Formulation

Because of the symmetry of the lateral load considered in (1), boundary conditions need only be written on one quadrant: $y = 0$ and $y = b$, $0 \le x \le \pi/2$; $x = 0$ and $x = \pi/2$, $0 \le y \le b$ (Fig. 1)

$$w=W_c \quad : x=0, y=0 \qquad (9)$$

$$w,_x=0,\ w=\delta \quad : y=0, e_1\le x\le \pi/2 \qquad (10a,b)$$

$$V_y=0 \quad : y=0, 0\le x<e_1 \qquad (10c)$$

$$M_y=0 \quad : y=0, 0\le x\le \pi/2 \qquad (10d)$$

$$V_y=0,\ w,_y=0 \quad : y=b, 0\le x\le \pi/2 \qquad (11a,b)$$

$$w,_y=0,\ w=0 \quad : x=0, e_2\le y\le b \qquad (12a,b)$$

$$V_x=0 \quad : x=0, 0\le y<e_2 \qquad (12c)$$

$$M_x=0 \quad : x=0, 0\le y\le b \qquad (12d)$$

$$V_x=0,\ w,_x=0 \quad : x=\pi/2, 0\le y\le b \qquad (13a,b)$$

In (9), W_c is the deflection of the corner and is to be determined. Furthermore, since the corners are unanchored and corner forces are therefore not permitted, (8) and (5) give

$$w,_{xy}=0:\ x=0, y=0 \qquad (14)$$

Utilizing the Levy-Nadai approach (Timoshenko and Woinowsky-Krieger, 1959), the lateral deflection satisfying (2a) and (9) is taken as

$$w(x,y)=\frac{Qa^4}{2D}\sum_{m=1,3,\ldots}^{\infty}[W_m^{(1)}(x,y)+W_m^{(2)}(y,x)]+W_c \qquad (15)$$

where

$$W_m^{(k)}(u,v)=[X_m^{(k)}(v)+Y_m^{(k)}(v)]\sin(m_k u) \qquad (16)$$

and

$$\pi^4 X_m^{(k)}(u)=\sum_{n=1,3,\ldots}^{\infty} a_{mn}(m_k^2+n_k^2)^{-2}\sin(n_k u) \qquad (17)$$

$$Y_m^{(k)}(u)=A_m^{(k)}\cosh(m_k u)+B_m^{(k)}m_k u\sinh(m_k u)$$
$$+C_m^{(k)}\sinh(m_k u)+D_m^{(k)}m_k u\cosh(m_k u) \qquad (18)$$

$$m_1=m,\ m_2=m\pi/2b,\ n_1=n\pi/2b,\ n_2=n \qquad (19)$$

Boundary conditions (10d), (11a,b) and (12d), (13a,b) lead to the following relations for $k = 1$ and $k = 2$, respectively

$$B_m^{(k)}=-(1-\nu)A_m^{(k)}/2 \qquad (20)$$

$$C_m^{(k)}=[\beta_k\operatorname{sech}^2(\beta_k)-2(1-\nu)^{-1}\tanh(\beta_k)](1-\nu)A_m^{(k)}/2 \qquad (21)$$

$$D_m^{(k)}=\tanh(\beta_k)(1-\nu)A_m^{(k)}/2 \qquad (22)$$

where

$$\beta_1=m_1 b,\ \beta_2=m_2(\pi/2) \qquad (23a,b)$$

The boundary conditions in (10a,c) and (12a,c), mixed as they are with respect to slope and shear, give rise to four coupled series equations. Following the same procedure used in a related analysis by Dempsey et al. (1984), (10c) and (12c) are integrated with respect to x and y, respectively, to give symbolically $[\int V_y(x,0)dx+C_y]$ and $[\int V_x(0,y)dy+C_x]$; C_y and C_x are constants of integration. To determine C_y and C_x it suffices to first add and then subtract the two equations obtained by letting $x = 0$ and $y = 0$ in the latter equations, respectively. The expressions for C_y and C_x are simplified by using an identity given by the corner force condition in (14). Equations (10a), (12a), $[\int V_y(x,0)dx+C_y]$ and $[\int V_x(0,y)+C_x]$ give finally, after substituting (15–23) and simplifying, the following coupled series equations

$$\sum_{m=1,3,\ldots}^{\infty} m_1 P_m^{(1)}\cos(m_1 x)=0,\ e_1\le x\le \pi/2 \qquad (24)$$

$$\sum_{m=1,3,\ldots}^{\infty} m_2 P_m^{(2)}\cos(m_2 y)=0,\ e_2\le y\le b \qquad (25)$$

$$\sum_{m=1,3,\ldots}^{\infty}[m_1^2 P_m^{(1)}\{(1+F_m^{(1)})\cos(m_1 x)-c_m^{(1)}\}$$
$$-m_2^2 P_m^{(2)}\{S_m^{(2)}(x)-c_m^{(2)}\}]$$
$$=\sum_{m=1,3,\ldots}^{\infty}[H_m^{(1)}\cos(m_1 x)-T_m^{(2)}(x)-d_m^{(1)}(0)+d_m^{(2)}(0)]$$
$$0\le x<e_1 \qquad (26)$$

$$\sum_{m=1,3,\ldots}^{\infty} [-m_1^2 P_m^{(1)} \{ S_m^{(1)}(y) - c_m^{(1)} \}$$

$$+ m_2^2 P_m^{(2)} \{ (1 + F_m^{(2)}) \cos(m_2 y) - c_m^{(2)} \}]$$

$$= \sum_{m=1,3,\ldots}^{\infty} [H_m^{(2)} \cos(m_2 y) - T_m^{(1)}(y) + d_m^{(1)}(0) - d_m^{(2)}(0)]$$

$$0 \leq y < e_2 \qquad (27)$$

where

$$P_m^{(k)} = (1-\nu) A_m^{(k)}/2 \qquad (28)$$

$$1 + F_m^{(k)} = \tanh(\beta_k) - \eta \beta_k \operatorname{sech}^2(\beta_k), \quad \eta = (1-\nu)/(3+\nu) \qquad (29a,b)$$

$$S_m^{(k)}(u) = \eta \operatorname{sech}(\beta_k) [\beta_k \operatorname{sech}(\beta_k) \cosh(m_k u)$$

$$- m_k u \cosh(\beta_k - m_k u) + \sinh(\beta_k - m_k u)] \qquad (30)$$

$$c_m^{(k)} = 2 \tanh(\beta_k)(3+\nu) \qquad (31)$$

$$H_m^{(k)} = d_m^{(k)}(0) + \eta b_m^{(k)}(0) \qquad (32)$$

$$T_m^{(k)}(u) = d_m^{(k)}(u) - \eta b_m^{(k)}(u) - g_m^{(k)}(u) + g_m^{(k)}(0) \qquad (33)$$

and

$$(3+\nu) \pi^4 m_k d_m^{(k)}(u)$$

$$= \sum_{n=1,3,\ldots}^{\infty} a_{mn} n_k (m_k^2 + n_k^2)^{-1} \cos(n_k u) \qquad (34)$$

$$\pi^4 b_m^{(k)}(u) = m_k \sum_{n=1,3,\ldots}^{\infty} a_{mn} n_k (m_k^2 + n_k^2)^{-2} \cos(n_k u) \qquad (35)$$

$$(3+\nu) \pi^4 m_k g_m^{(k)}(u) = \sum_{n=1,3,\ldots}^{\infty} a_{mn} n_k^{-1} \cos(n_k u) \qquad (36)$$

Equations (15-23), (28), and (35), together with (14) give the following expression for the corner force condition used in the derivation of (26) and (27)

$$\sum_{m=1,3,\ldots}^{\infty} [m_1^2 P_m^{(1)} a_m^{(1)} + m_2^2 P_m^{(2)} a_m^{(2)}]$$

$$= \sum_{m=1,3,\ldots}^{\infty} [b_m^{(1)}(0) + b_m^{(2)}(0)] \qquad (37)$$

where

$$(1-\nu) a_m^{(k)} = (1+\nu) \tanh(\beta_k) - (1-\nu) \beta_k \operatorname{sech}^2(\beta_k) \qquad (38)$$

The remaining boundary conditions (10*b*) and (12*b*) are satisfied at a later stage. The problem is therefore reduced to the determination of the constants $P_m^{(k)}$ $(k=1,2)$ such that the four coupled series (24-27) are satisfied. In this respect, the constant 1 in each of (26) and (27) serves to isolate the singularities at $x = e_1$ and $y = e_2$, respectively, since as $m \to \infty$ the functions $F_m^{(k)} \sim e^{-2\beta_k}$ $(k=1,2)$.

Singular Integral Equations

The coupled equations (24-27) may be reduced to two coupled singular integral equations by representing the unknown coefficients $P_m^{(k)}$ in (28) by finite Fourier transforms

$$m_k^2 P_m^{(k)} = \int_0^{e_k} \phi_k(t) \sin(m_k t)\, dt \qquad (k=1,2) \qquad (39)$$

in which the auxiliary functions, $\phi_k(t)$ $(k=1,2)$, remain to be determined.

The same procedure outlined by Dempsey et al. (1984) in an analysis of the contact between a uniformly loaded square plate and unilateral supports is now followed. It is readily shown that the above definitions for $P_m^{(1)}$ and $P_m^{(2)}$ identically satisfy the two uncoupled equations (24) and (25), respectively. The remaining coupled equations (26) and (27) can be reduced to coupled singular integral equations of the first kind

$$\frac{1}{\pi} \int_0^{e_1} \left(\frac{1}{t-x} + \frac{1}{t+x} \right) \phi_1(t)\, dt + \int_0^{e_1} M_1(x,t) \phi_1(t)\, dt$$

$$+ \int_0^{e_2} N_2(x,t) \phi_2(t)\, dt = f_1(x), \quad 0 \leq x < e_1 \qquad (40)$$

$$\frac{1}{\pi} \int_0^{e_2} \left(\frac{1}{t-y} + \frac{1}{t+y} \right) \phi_2(t)\, dt + \int_0^{e_1} N_1(y,t) \phi_1(t)\, dt$$

$$+ \int_0^{e_2} M_2(y,t) \phi_2(t)\, dt = f_2(y), \quad 0 \leq y < e_2 \qquad (41)$$

where $(j=1,2)$

$$\pi M_j(u,t) = (\pi/l_j)/\sin[\pi(t-u)/l_j] - 1/(t-u)$$

$$+ (\pi/l_j)/\sin[\pi(t+u)/l_j] - 1/(t+u)$$

$$+ 4(\pi/l_j) \sum_{m=1,3,\ldots}^{\infty} [F_m^{(j)} \cos(m_j u) - c_m^{(j)}] \sin(m_j t) \qquad (42)$$

$$\pi N_j(u,t) = -4(\pi/l_j) \sum_{m=1,3,\ldots}^{\infty} [S_m^{(j)}(u) - c_m^{(j)}] \sin(m_j t) \qquad (43)$$

$$\pi f_1(x) = 4 \sum_{m=1,3,\ldots}^{\infty} [H_m^{(1)} \cos(m_1 x) - T_m^{(2)}(x) - d_m^{(1)}(0) + d_m^{(2)}(0)] \qquad (44)$$

$$\pi f_2(y) = 4(\pi/l_2) \sum_{m=1,3,\ldots}^{\infty} [H_m^{(2)} \cos(m_2 y) - T_m^{(1)}(y) + d_m^{(1)}(0)$$

$$- d_m^{(2)}(0)] \qquad (45)$$

and

$$l_1 = \pi, \quad l_2 = 2b \qquad (46a,b)$$

The corner force condition in (37) together with (39) becomes

$$\int_0^{e_1} R_1(t) \phi_1(t)\, dt + \int_0^{e_2} R_2(t) \phi_2(t)\, dt = B \qquad (47)$$

where

$$R_k(u) = \sum_{m=1,3,\ldots}^{\infty} a_m^{(k)} \sin(m_k u) \qquad (k=1,2) \qquad (48a)$$

$$B = \sum_{m=1,3,\ldots}^{\infty} (b_m^{(1)}(0) + b_m^2(0)) \qquad (48b)$$

Equations (40) and (41) both reduce to the corner force condition in (47) for $x = 0$ and $y = 0$, respectively.

Physical Quantities

The edge displacements the plate undergoes near each corner as it lifts off the supports, using (15) together with (20-22) and (28), are given by

$$D(1-\nu) w(x,0) = 2Qa^4 \sum_{m=1,3,\ldots}^{\infty} P_m^{(1)} \sin(m_1 x)$$

$$+ W_c: \; 0 \leq x \leq \pi/2 \qquad (49)$$

$$D(1-\nu) w(0,y) = 2Qa^4 \sum_{m=1,3,\ldots}^{\infty} P_m^{(2)} \sin(m_2 y) + W_c: \; 0 \leq y \leq b \qquad (50)$$

The edge displacements in (49) and (50), together with (39) can be used to determine W_c and then expressions for the edge displacements on those parts of the plate that lose contact. That is

$$4D(1-\nu)\,w(x,0)=\pi Qa^4\int_x^{e_1}\phi_1(t)\,(x-t)\,dt$$
$$+4(1-\nu)\,\delta D:\ 0\le x<e_1 \tag{51}$$

$$2D(1-\nu)\,w(0,y)=bQa^4\int_y^{e_2}\phi_2(t)\,(y-t)\,dt:\ 0\le y<e_2 \tag{52}$$

Compatibility of displacements in (51), (52) imposes an additional condition on the solution to the coupled equations (40) and (41), since $w(x\to 0,\ 0)$ should equal $w(0,\ y\to 0)$. It follows that

$$(\pi/2)\int_0^{e_1}t\phi_1(t)\,dt-b\int_0^{e_2}t\phi_2(t)\,dt=2(1-\nu)\,\delta D/Qa^4 \tag{53}$$

The supplemented, or Kirchhoff, shearing forces V_y and V_x at the contact with the two supports $y=0$, $e_1<x<\pi/2$ and $x=0$, $e_2<y\le b$, respectively, are

$$V_y(e_1<x\le\pi/2,y=0)=-\pi^3Qa(3/2+\nu/2)\sum_{m=1,3,\ldots}^{\infty}$$
$$\{m_1^3P_m^{(1)}(1+F_m^{(1)})\sin(m_1x)+m_2^2P_m^{(2)}S_{m,x}^{(2)}(x)$$
$$-m_1H_m^{(1)}\sin(m_1x)-T_{m,x}^{(2)}(x)\} \tag{54}$$

$$V_x(x=0,e_2<y\le b)=-\pi^3Qa(3/2+\nu/2)\sum_{m=1,3,\ldots}^{\infty}$$
$$\{m_1^2P_m^{(1)}S_{m,y}^{(1)}(y)+m_2^3P_m^{(2)}(1+F_m^{(2)})\sin(m_2y)$$
$$-m_2H_m^{(2)}\sin(m_2y)-T_{m,y}^{(1)}(y)\} \tag{55}$$

Preliminary to discussing the solution of the coupled integral equations in (40), (41) and the additional condition in (53), it is necessary to clarify the end-point behavior of the auxiliary functions, $\phi_k(t)$ $(k=1,2)$, in (39). It follows from (15–28) that

$$D(1-\nu)\{w_{,xx};w_{,yy}\}|_{y=0}$$
$$=Qa^4\{-1;\nu\}\sum_{m=1,3,\ldots}^{\infty}m_1^2P_m^{(1)}\sin(m_1x) \tag{56}$$

$$D(1-\nu)\{w_{,xx};w_{,yy}\}|_{x=0}$$
$$=Qa^4\{\nu;-1\}\sum_{m=1,3,\ldots}^{\infty}m_2^2P_m^{(2)}\sin(m_2y) \tag{57}$$

for $0\le x\le\pi/2$ and $0\le y\le b$, respectively. Given the identity (Gorman, 1982)

$$2\sum_{m=1,3,\ldots}^{\infty}\sin(m\pi t/2)\sin(m\pi u/2)=\delta(u-t) \tag{58}$$

equations (56), (57) together with (39) provide that

$$M_x(x,0);\ M_y(0,y)=(1+\nu)Qa^2\pi^2\{\pi\phi_1(x);\ 2b\phi_2(y)\}/4 \tag{59a,b}$$

Clearly, the auxiliary functions $\phi_1(x)$ and $\phi_2(y)$ have the same behavior at $x=e_1$, $y=0$ and $x=0$, $y=e_2$, respectively, as do the moments there. The foregoing relationships establish that $M_x(x,\ 0)\sim\phi_1(x)$ and $M_y(0,\ y)\sim\phi_2(y)$; boundary conditions (10*d*) and (12*d*) quickly determine, therefore, that $\phi_k(0)=0$ $(k=1,\ 2)$. As discussed in the introduction, the moments are bounded at the transition points. The shearing forces in (54) and (55) should therefore be singular as the inverse square root with distance from $x=e_1$, $y=0$ and $x=0$, $y=e_2$; only the terms involving the constant 1 in these expressions contribute to the singularities at e_1, e_2. Using (39) together with (54) and (55), expressions analogous to that obtained by Dempsey et al. (1984) in (44) of that paper (in which a factor of 1/2 is missing) are found. It follows that $V_y(x,\ 0)\sim(x-e_1)^{-1/2}$ and $V_x(0,\ y)\sim(y-e_2)^{-1/2}$ as $x,\ y\to e_1^+,\ e_2^+$, respectively, as long as $\phi_k(t)\sim(e_k-t)^{1/2}$ $(k=1,2)$. As expected, the moments are bounded at the transition points from simple support to no contact. Furthermore, although each supplemented shearing force is singular at the ends of each contact interval, each is integrable because the singularities are of the inverse square-root type.

Numerical Analysis

Equations (40), (41), and (53) are prepared for numerical analysis by first extending the ranges of integration to $-e_1<x<e_1$, $-e_2<y<e_2$; it is useful to note also that $M_j(u,\ t)$ and $N_j(|u|,\ t)$ are odd in t but even in u. The functions $f_k(u)$ are also even in u. If it is assumed that the functions $\phi_j(t)$ $(j=1,\ 2)$ are odd functions, the above equations can be written in the form

$$\frac{1}{\pi}\int_{-e_1}^{e_1}\frac{\phi_1(t)}{t-x}dt+\frac{1}{2}\int_{-e_1}^{e_1}M_1(x,t)\,\phi_1(t)\,dt$$
$$+\frac{1}{2}\int_{-e_2}^{e_2}M_2(y,t)\,\phi_2(t)\,dt=f_2(y),$$
$$-e_1<x<e_1 \tag{60}$$

$$\frac{1}{\pi}\int_{-e_2}^{e_2}\frac{\phi_2(t)}{t-y}dt+\frac{1}{2}\int_{-e_1}^{e_1}N_1(|y|,t)\,\phi_1(t)\,dt$$
$$+\frac{1}{2}\int_{-e_2}^{e_2}M_2(y,t)\,\phi_2(t)\,dt=f_2(y),$$
$$-e_2<y<e_2 \tag{61}$$

$$(\pi/2)\int_{-e_1}^{e_1}t\phi_1(t)\,dt-b\int_{-e_2}^{e_2}t\phi_2(t)\,dt=4(1-\nu)\,\delta D/Qa^4 \tag{62}$$

Similarly, (47) can be written

$$\int_{-e_1}^{e_1}R_1(t)\,\phi_1(t)\,dt+\int_{-e_2}^{e_2}R_2(t)\,\phi_2(t)\,dt=2B \tag{63}$$

The Gauss-Chebyshev integration formulae (Erdogan et al., 1973) are applied to equations (60–63). First, however, let

$$t=e_1u,\ x=e_1s,\ \phi_1(t)=\bar\theta_1(u)\sqrt{1-u^2} \tag{64a}$$

$$t=e_2v,\ y=e_2z,\ \phi_2(t)=\bar\theta_2(v)\sqrt{1-v^2} \tag{64b}$$

where $\bar\theta_1(u)$ and $\bar\theta_2(v)$ are regular at $\pm$ 1. The foregoing definitions of $\phi_k(t)$ imply that the moments are bounded at $t=e_1,\ e_2$ for $k=1,\ 2$, respectively. Equations (60–62) can thus be written in the following form

$$\sum_{j=1}^{n}\frac{(1-u_j^2)}{n+1}\Bigg[\frac{\bar\theta_1(u_j)}{u_j-s_i}+\frac{\pi}{2}e_1M_1(e_1s_i,e_1u_j)\,\bar\theta_1(u_j)$$
$$+\frac{\pi}{2}e_2N_2(e_1|s_i|,e_2v_j)\,\bar\theta_2(v_j)\Bigg]=f_1(e_1s_i) \tag{65}$$

$$\sum_{j=1}^{n}\frac{(1-u_j^2)}{n+1}\Bigg[\frac{\bar\theta_2(v_j)}{v_j-z_i}+\frac{\pi}{2}e_1N_1(e_2|z_i|,e_1u_j)\,\bar\theta_1(u_j)$$
$$+\frac{\pi}{2}e_2M_2(e_2z_i,e_2v_j)\,\bar\theta_2(v_j)\Bigg]=f_2(e_2z_i) \tag{66}$$

$$\sum_{j=1}^{n}\frac{(1-u_j^2)}{n+1}\pi(1+u_j)\,[(\pi/2)e_1^2\bar\theta_1(u_j)$$
$$-be_2^2\bar\theta_2(u_j)]=4(1-\nu)\,\delta D/Qa^4 \tag{67}$$

where

$$u_j = v_j = \cos[j\pi/(n+1)], \; j = 1,2, \ldots, n \qquad (68a)$$

$$s_i = z_i = \cos[(2i-1)\pi/2(n+1)], \; i = 1,2, \ldots, n+1 \qquad (68b)$$

Equation (63) becomes

$$\sum_{j=1}^{n} \frac{(1-u_j^2)}{n+1}[e_1 R_1(e_1 u_j)\bar{\theta}_1(u_j) + e_2 R_2(e_2 v_j)\bar{\theta}_2(v_j)] = 2B \qquad (69)$$

Unsagged Supports. For the case of unsagged supports ($\delta = 0$) a receding contact problem results. That is, the extent of contact is independent of the load level but the support reactions are proportional to the load. For $\delta = 0$ in (67), the solution method is as follows. There are $(2n+3)$ equations arising from (65-67) from which we wish to solve for n values each of $\bar{\theta}_1(u_j)$, $\bar{\theta}_2(v_j)$ and one value each of e_1 and e_2. For $x = y = 0$ each of (60), (61) reduces to the corner force condition (63); the associated collocation points in (65), (66) are for $i = n/2+1$. The procedure is to solve the $2n$ equations in (65), (66) obtained by disregarding the two $i = n/2+1$ equations and by first assuming values for e_1 and e_2. Equation (67) and the corner force condition (69) then provide checking equations. A series of values of e_1 and e_2 are tried until the correct ones that simultaneously satisfy the checking equations are found.

Sagged Supports. For sagged supports ($\delta \neq 0$) an advancing contact problem results for a load level greater than Q°, where Q° denotes the load at which the plates starts to touch the sagged supports (in which case $e_1 = \pi/2$). In advancing contact problems the extent of contact depends upon the level of loading. The amount of sag δ and the intensity Q of the distributed loading determine for a given plate the extent of contact defined by e_1, e_2. As pointed out by Dundurs et al. (1974) it is more convenient to view e_1 and Q, rather than δ and Q, as the independent parameters. The solution procedure is to first set e_1 at some value, then solve the $2n$ equations in (65), (66) obtained by once again disregarding the two $i = n/2+1$ equations after assuming a value for e_2. The corner force condition (69) then provides a checking equation; a series of values of e_2 are tried until one that satisfies the checking equation to the desired accuracy is determined.

The load intensity at which the plate first touches the sagged supports Q° is found by setting $e_1 = \pi/2$. For $Q > Q^\circ$, $e_1 < \pi/2$. An examination of (65), (66), and (69) reveals that for a given aspect ratio $(2b/\pi)$, e_2 is a function solely of e_1. It then follows that (67) can be written in the form

$$\delta = \gamma(e_1;b)Qa^4/D \qquad (70)$$

where the definition of the function $\gamma(e_1;b)$ is obvious from (67). For a given aspect ratio and contact length e_1, therefore, the ratio Q/Q° is independent of the amount of sag δ since

$$Q/Q^\circ = \gamma(\pi/2;b)/\gamma(e_1;b) \qquad (71)$$

The edge displacements, on the other hand, are directly proportional to δ, given b and e_1; the latter observation can be verified by substituting (71) in (51) and (52).

Conclusions

The theory and solution methods required to determine the natural contact of rectangular plates resting on unilateral edge supports and subjected to centrally symmetric but otherwise arbitrary loading have been presented. The possibility that two opposite supports may have sagged was treated. The influence of different factors such as aspect ratio, amount of sag, and Poisson's ratio remain to be examined for a range of loading distributions and levels.

References

Dempsey, J. P., Keer, L. M., Patel, N. B., and Glasser, M. L., 1984, "Contact Between Plates and Unilateral Supports," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 324-328.

Dundurs, J., and Stippes, M., 1970, "Role of Elastic Constants in Certain Contact Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 965-970.

Dundurs, J., Kiattikomol, K., and Keer, L. M., 1974, "Contact Between Paltes and Sagged Supports," *ASCE Journal of the Engineering Mechanics Division*, Vol. 100, pp. 445-456.

Erdogan, F., Gupta, G. D., and Cook, T. S., 1973, "Numerical Solution of Singular Integral Equations," *Mechanics of Fracture*, Vol. 1, Sih, G. C., ed., Noordhoff, Leyden, pp. 368-425.

Gorman, D. J., 1982, *Free Vibration Analysis of Rectangular Plates*, Elsevier, New York, p. 308.

Keer, L. M., and Mak, A. F., 1981, "Loss of Contact in the Vicinity of a Right Angle Corner for a Simply Supported, Laterally Loaded Plate," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 597-600.

Kiattikomol, K., Keer, L. M., and Dundurs, J., "Application of Dual Series to Rectangular Plates," *ASCE Journal of the Engineering Mechanics Division*, Vol. 100, pp. 433-444.

Timoshenko, S. P., and Woinowsky-Krieger, S., *Theory of Plates and Shells*, McGraw-Hill, New York, 1959.

Williams, M. L., 1952, "Surface Stress Singularities Resulting from Various Boundary Conditions in Angular Corners of Plates under Bending," *Proceedings of the First U.S. National Congress of Applied Mechanics*, pp. 325-329.

J. P. Dempsey
Assistant Professor.
Mem. ASME

Hui Li
Graduate Student.

Department of Civil
and Environmental Engineering,
Clarkson University,
Potsdam, N.Y. 13676

Rectangular Plates on Unilateral Edge Supports: Part 2—Implementation; Concentrated and Uniform Loading

Rectangular plates in unilateral contact with sagged and unsagged supports laterally loaded by centrally concentrated loads and uniform pressure are examined. The loss of contact and the redistribution of deflections, moments, and support reactions are presented. Computer implementation aspects are discussed.

Introduction

In a companion paper Dempsey and Li (1986) presented the theory and solution methods required to determine the natural contact of centrally but otherwise arbitrarily laterally loaded rectangular plates on unilateral edge supports. In this paper, computer implementation aspects and two numerical examples, concentrated and uniform loading, are presented. Results for the loss of contact, uplift, and redistribution of deflections, moments, and support reactions are provided. The loadings chosen provide upper and lower bounds on the unilateral contact behavior that would result for intermediate classes of loading, such as circular or patch loads.

The plate geometry and coordinate system are shown in Fig. 1. The actual (barred) coordinates are given by $\bar{x}=ax/\pi$, $\bar{y}=ay/\pi$, $\bar{b}=ab/\pi$, etc. Without loss of generality it is assumed that the two sagged edge supports are those parallel to the x-axis. The amount of sag between these supports and the undeformed plate is defined by the distance δ. The support conditions are defined by (Fig. 1).

$$w,_x=0\ ,\ w=\delta : y=0 \text{ and } y=2b\ ,\ e_1\leq x\leq \pi-e_1 \qquad (1a)$$

$$w,_y=0\ ,\ w=0 : x=0 \text{ and } x=\pi\ ,\ e_2\leq y\leq 2b-e_2 \qquad (1b)$$

As revealed by (1), each support is level. Also each edge support rigidly resists displacement in the positive z-direction only. The loading treated in the companion paper (Dempsey and Li, 1986) is assumed to be symmetric about $x=\pi/2$, $y=b$ and, in terms of the transformed coordinates, is given there by the expression

$$q(x,y)=Q\sum_{m=1,3,\ldots}^{\infty}\ \sum_{n=1,3,\ldots}^{\infty} a_{mn}\sin(mx)\sin(n\pi y/2b) \qquad (2)$$

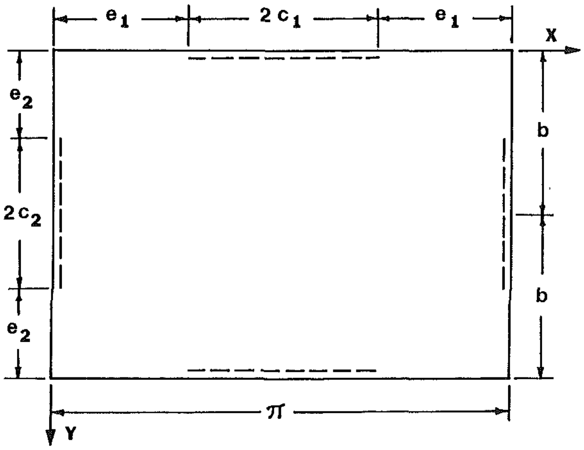


Fig. 1 Rectangular plate on unilateral edge supports

where Q is an appropriate load factor for the particular loading being treated.

Further details concerning the theory and numerical analysis are provided in the companion paper (Dempsey and Li, 1986). Equations in the latter paper will be cited freely and referred to using the notation (D:). It suffices to note here that the auxiliary functions $\phi_1(t)$, $\phi_2(t)$ introduced in (D:39) are solved for numerically from (D:65–D:69). By substituting these functions into (D:39), then (D:28) and (D:20–D:22) for $A_m^{(k)}$ and $B_m^{(k)}$, $C_m^{(k)}$, $D_m^{(k)}$, respectively, and into (D:51) for W_c, the deflection function $w(x,y)$ given in (D:15) can be determined for any position in the plate and so can the moments and shear forces.

Concentrated and Uniform Loading

For a uniform load of intensity q, the expansion in (2) is valid for

$$Q=q, a_{mn}=16/\pi^2 mn \quad (m \text{ or } n \text{ odd}) \qquad (3a,b)$$

for even m or n, $a_{mn}=0$. The numerical solution of

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Miami Beach, Fla., November 17–22, 1985, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March, 1985; final revision, June, 1985. Paper No. 85-WA/APM-24.

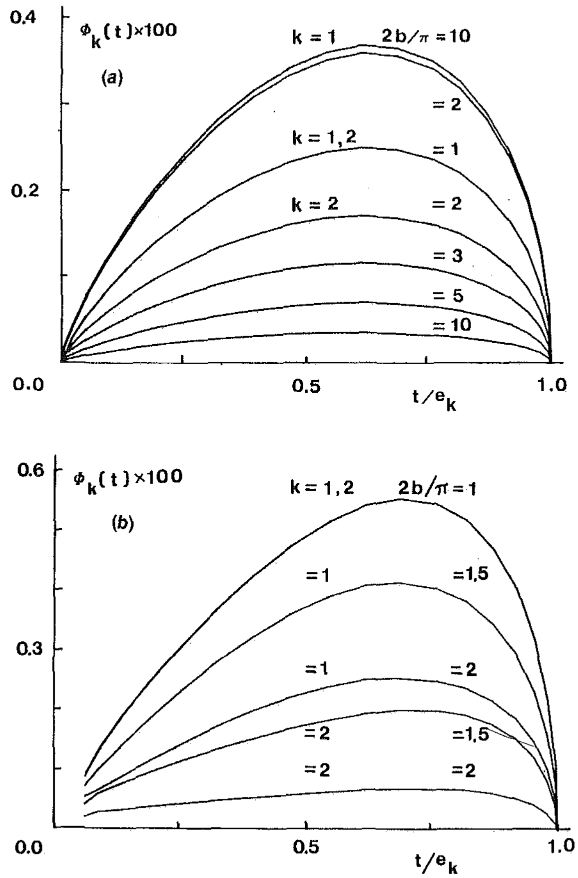


Fig. 2 The auxiliary functions $\phi_k(t)$ for a rectangular plate on unsagged supports ($\nu = 0.3$): (a) uniform pressure, (b) concentrated load

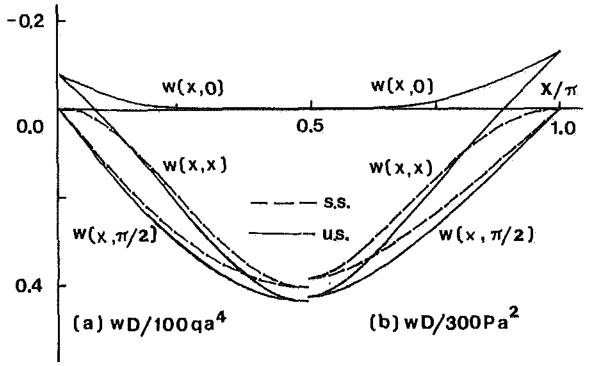


Fig. 3 Deflections $w(x,0)$, $w(x,x)$, and $w(x,\pi/2)$ for square plate on unsagged unilateral (u.s.) and simple (s.s) supports ($\nu = 0.3$): (a) uniform pressure; (b) concentrated load

(D:65–D:69) requires the summation of three expressions involving (3). Using Hansen (1975), it is found that (D:34–D:36) become, for $-l_k/2 < u < l_k/2$ $(l_1 = \pi, l_2 = 2b)$

$$(3+\nu)\pi^4 l_k m_k^3 d_m^{(k)}(u) = 4\,\mathrm{sech}(\beta_k)\sinh(\beta_k - m_k|u|) \tag{4}$$

$$\begin{aligned}\pi^4 l_k m_k^3 b_m^{(k)}(u) &= 2\,\mathrm{sech}(\beta_k)[\sinh(\beta_k - m_k|u|) \\ &\quad + m_k|u|\cosh(\beta_k - m_k|u|) \\ &\quad - \beta_k\,\mathrm{sech}(\beta_k)\cosh(m_k u)]\end{aligned} \tag{5}$$

$$(3+\nu)\pi^4 l_k m_k^3 g_m^{(k)}(u) = 4(\beta_k - m_k|u|) \tag{6}$$

In terms of the actual coordinates and dimensions, a concentrated load at the center of the plate can be expressed as $\bar{q}(\bar{x},\bar{y}) = P\delta(\bar{x} - a/2)\delta(\bar{y} - ab/\pi)$; in terms of the transformed coordinates, the same load is given by

Table 1 Rectangular plates on unsagged supports (= 0.3)

Uniform pressure: $Q = q$

$2b/\pi$	e_1/π $[R/Qa^2]$	e_2/π $[e_2/2b]$	$\frac{W_c\cdot10^2}{Qa^4/D}$	$\frac{w_{\max}\cdot10}{Qa^4/D}$	$\frac{(V_x)_{\mathrm{mid}}}{Qa}$	$\frac{(V_y)_{\mathrm{mid}}}{Qa}$	$\frac{(M_x)_{\max}}{Qa^2}$	$\frac{(M_y)_{\max}}{Qa^2}$
1.0	0.262 [0.065]	0.262 [0.262]	−0.0756	0.0438 (0.0406)	0.448 (0.421)	0.448 (0.421)	0.0509 (0.0479)	0.0509 (0.0479)
1.5	0.298 [0.085]	0.303 [0.202]	−0.131	0.0817 (0.0773)	0.499 (0.486)	0.525 (0.480)	0.0853 (0.0812)	0.0512 (0.0498)
2.0	0.308 [0.092]	0.316 [0.158]	−0.150	0.105 (0.101)	0.508 (0.503)	0.548 (0.496)	0.105 (0.102)	0.0464 (0.0464)
3.0	0.311 [0.093]	0.320 [0.107]	−0.157	0.123 (0.122)	0.505 (0.505)	0.555 (0.501)	0.120 (0.119)	0.0404 (0.0406)
Concentrated loading: $Q = P/a^2$								
1.0	0.400 [0.122]	0.400 [0.400]	−0.387	0.123 (0.116)	0.900 (0.650)	0.900 (0.650)	∞	∞
1.5	0.449 [0.096]	0.520 [0.347]	−0.363	0.163 (0.153)	0.737 (0.680)	0.716 (0.346)	∞	∞
2.0	0.477 [0.060]	0.630 [0.315]	−0.255	0.170 (0.165)	0.692 (0.678)	0.645 (0.182)	∞	∞
3.0	0.498 [0.018]	0.851 [0.284]	−0.098	0.170 (0.169)	0.676 (0.675)	1.305 (0.048)	∞	∞

Table 2 Square plate on unsagged supports $e_1 = e_2 = e$

ν	e^u/π	e^c/π
0.1	0.300	0.434
0.3	0.262	0.400
0.5	0.218	0.353

u = uniform pressure
c = concentrated load

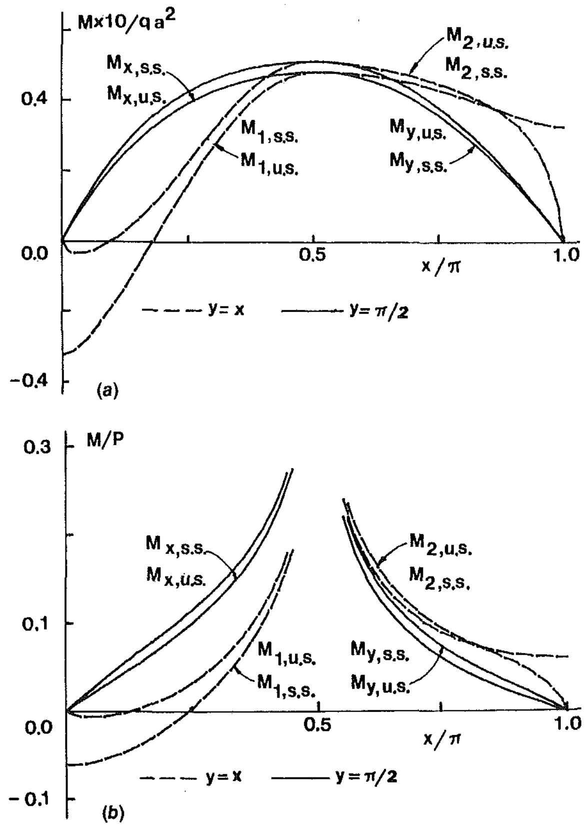


Fig. 4 Moments M_1 (x,x), $M_2(x,x)$, $M_x(x, \pi/2)$ and $M_y(x, \pi/2)$ for a square plate on unsagged unilateral (u.s.) and simple (s.s) supports ($\nu = 0.3$): (a) uniform pressure, (b) concentrated load

$$q(x,y) = P(\pi/a)^2\delta(x-\pi/2)\delta(y-b) \tag{7}$$

The equivalent double Fourier sine series in (2) is valid for

$$Q = P/a^2, a_{mn} = 4(\pi/2b)\sin(m\pi/2)\sin(n\pi/2) \tag{8a,b}$$

Equations (D:34–D:36) together with (8b) give, using Hansen (1975),

$$(3+\nu)\pi^3 d_m^{(k)}(u) = m^{-1}\sin(m\pi/2)\text{sech}(\beta_k)\cosh(m_k u) \tag{9}$$

$$\pi^4 m_k^2 b_m^{(k)}(u) = (\pi/2b)\sin(m\pi/2)\beta_k[\beta_k\tanh(\beta_k)\cosh(m_k u) - m_k u\sinh(m_k u)] \tag{10}$$

$$(3+\nu)\pi^4 m_k^2 g_m^{(k)}(u) = 2(\pi/2b)\sin(m\pi/2)\beta_k \tag{11}$$

The closed-form expressions in (4–6), (9–11) may now be used to define (D:32), (D:33), (D:37), and then (D:44), (D:45), (D:48b) and, finally, (D:65–D:69).

Implementation

The foregoing numerical analysis requires the summation of 58 different series. The authors established a uniform error tolerance for each summation. A global check on the accuracy was obtained by using the fact that the support reactions are integrable and that therefore the total load must balance. The final results are believed to be accurate to three significant digits.

Several of the series converge very slowly. Examples include

Table 3 Uniformly loaded square plate on unsagged supports: variation of physical quantities with Poisson's ratio

$\frac{x}{\pi}$	$\frac{w(x,0)}{\alpha}$	$\frac{w(x,x)}{\alpha}$	$\frac{w(x,\pi/2)}{\alpha}$	$\frac{M_x(x,\pi/2)}{\beta}$	$\frac{M_y(x,\pi/2)}{\beta}$	$\frac{M_1(x,x)}{\beta}$	$\frac{M_2(x,x)}{\beta}$	$\frac{V(x,0)}{\gamma}$
				$\nu = 0.1$				
0.0	−1.12	−1.12	0.000	0.000	0.000	0.000	0.000	0.000
0.1	−0.488	0.278	1.53	2.26	1.44	−0.201	3.82	0.000
0.2	−0.100	1.74	2.85	3.56	2.72	1.08	4.25	0.000
0.3	0.000	3.18	3.84	4.20	3.70	2.72	4.38	----
0.4	0.000	4.26	4.46	4.46	4.32	4.10	4.43	5.16
0.5	0.000	4.66	4.66	4.53	4.53	4.53	4.53	4.97
				$\nu = 0.3$				
0.0	−0.756	−0.756	0.000	0.000	0.000	0.000	0.000	0.000
0.1	−0.283	0.347	1.43	2.31	1.78	0.013	3.38	0.000
0.2	−0.030	1.64	2.67	3.75	3.21	1.33	4.20	0.000
0.3	0.000	2.98	3.61	4.57	4.16	3.15	4.66	5.28
0.4	0.000	4.00	4.19	4.97	4.88	4.60	4.95	4.48
0.5	0.000	4.38	4.38	5.09	5.09	5.09	5.09	4.47
				$\nu = 0.5$				
0.0	−0.479	−0.479	0.000	0.000	0.000	0.000	0.000	0.000
0.1	−0.135	0.394	1.37	2.42	2.09	0.204	2.92	0.000
0.2	−0.001	1.57	2.56	4.03	3.69	1.73	4.04	0.000
0.3	0.000	2.85	3.46	5.00	4.81	3.70	4.89	3.99
0.4	0.000	3.84	4.02	5.52	5.47	5.17	5.42	4.24
0.5	0.000	4.21	4.21	5.68	5.68	5.68	5.68	4.08

$\alpha = qa^4/10^3\text{D}$, $\beta = qa^2/10^2$, $\gamma = qa/10$

Table 4 Square plate on unsagged supports loaded by a concentrated load ($\nu = 0.3$)

$\frac{x}{\pi}$	$w(x,0)$ α	$w(x,x)$ α	$w(x,\pi/2)$ α	$M_x(x,\pi/2)$ β	$M_y(x,\pi/2)$ β	$M_1(x,x)$ β	$M_2(x,x)$ β	$V(x,0)$ γ
0.0	−0.387	−0.387	0.000	0.000	0.000	0.000	0.000	0.000
0.1	−0.225	−0.042	0.336	0.469	0.281	−0.041	0.535	0.000
0.2	−0.096	0.330	0.657	0.907	0.557	0.150	0.788	0.000
0.3	−0.019	0.727	0.945	1.40	0.940	0.534	1.13	0.000
0.4	0.000	1.09	1.18	2.16	1.63	1.26	1.81	----
0.5	0.000	1.29	1.29	∞	∞	∞	∞	9.00

$\alpha = Pa^2/10^2\text{D}, \beta = P/10, \gamma = P/10a$

Table 5 Square plate on sagged supports: uniform pressure $\nu = 0.3$

e_1/π	e_2/π	$-W_c/\delta$	q/q°
0.500	$<10^{-4}$	$<10^{-4}$	1.00
0.475	$<10^{-4}$	$<10^{-4}$	1.67
0.450	0.021	$<10^{-4}$	2.16
0.425	0.065	0.00229	2.73
0.400	0.102	0.0118	3.47
0.375	0.135	0.0368	4.50
0.350	0.166	0.0943	6.08
0.325	0.195	0.228	8.82
0.300	0.222	0.590	15.0
0.275	0.248	2.48	43.9

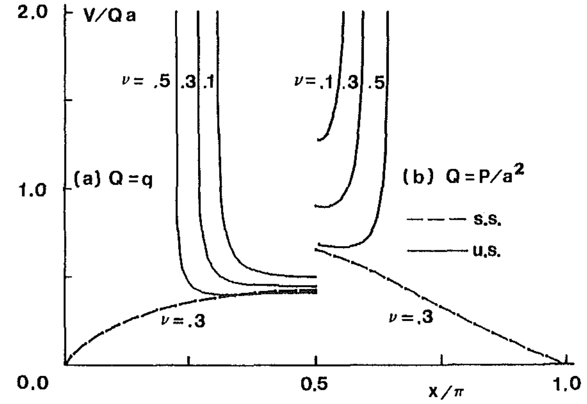


Fig. 5 Support reaction $V_y(x,0)$ for a square plate on unsagged unilateral (u.s.) and simple (s.s) supports: (*a*) uniform pressure, (*b*) concentrated load

$\sum_{m=1,3,\ldots}^{\infty} \sin(mx)/m^2, \sum_{m=1,3,\ldots}^{\infty} x^m/m^2, \sum_{m=1,3,\ldots}^{\infty} e^{-mx}/m^2$. The first was summed by noting that (Hansen, 1975)

$$\sum_{m=1,3,\ldots}^{\infty} \frac{\sin(mx)}{m^2} = -\int_0^{x/2} \ln(\tan t)dt \tag{12}$$

ln(tan$t$) can then be expanded into a Taylor series (using MACSYMA). The second and third pose difficulties when $x \to 1^-$ and $x \to 0^+$, respectively.

Many different methods that accelerate the convergence of series were tried; often, these methods worked well in special instances. However, crude term by term summation was used in most cases, with all possible check cases being evaluated to verify accuracy. The latter approach, unavoidably, led to undesirably lengthy computer runs.

Unsagged Supports. Rectangular plates in unilateral contact with unsagged edge supports subjected to either uniform pressure or concentrated lateral loads were examined as described above. In Figs. 2(*a*,*b*) the resulting auxiliary functions, $\phi_k(t)$ ($k = 1,2$), introduced in (D:39), and solved for in (D:65–D:67), are plotted. The requirements that $\phi_k(0) = 0$ and $\phi_k(t) \sim (e_k - t)^{1/2}$ as $t \to e_k^+$, are clearly satisfied. In Table 1, the extent of contact e_1 and e_2, the magnitude of corner uplift W_c, the maximum displacement $w_{\max} = w(\pi/2,b)$, the maximum moments $(M_x)_{\max} = M_x(\pi/2,b)$ and $(M_y)_{\max} = M_y(\pi/2,b)$, the mid-side support reactions $(V_x)_{\text{mid}} = V_x(0,b)$ and $(V_y)_{\text{mid}} = V_y(\pi/2,0)$, are tabulated for different aspect ratios. Poisson's ratio is taken as 0.3.

The values in parentheses in the columns of Table 1 for $w_{\max}$, $(M_x)_{\max}$, $(M_y)_{\max}$, $(V_x)_{\text{mid}}$ and $(V_y)_{\text{mid}}$ correspond to the associated simply supported solution provided by Timoshenko and Woinowsky-Krieger (1959) in Section 30 and Table 8, p. 120. In the simply supported solution, uplift is prevented, and concentrated reactions at each corner result. The magnitude of these corner reactions R are bracketed

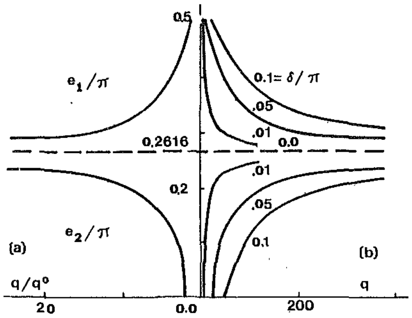


Fig. 6 Extent of unilateral contact versus level of loading given different amounts of sag for a uniformly loaded square plate on sagged supports ($\nu = 0.3$): (*a*) e_1 and e_2 versus q/q^0, (*b*) e_1 and e_2 versus q

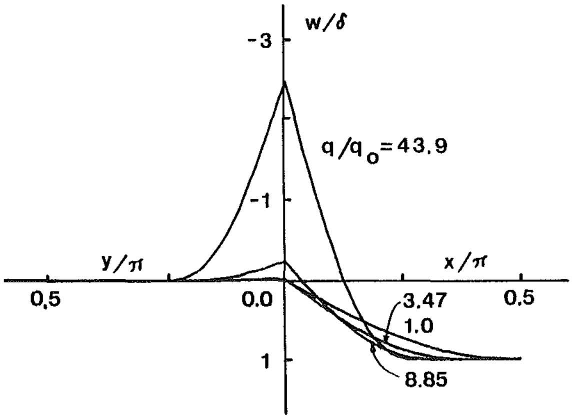


Fig. 7 The edge displacements near the corner for a unilaterally supported uniformly loaded square plate on sagged supports ($\nu = 0.3$)

Table 6 Square plate on unsagged supports loaded by a combination of uniform and concentrated pressure ($e_1 = e_2 = e$)

$\nu = 0.3$

P/qa^2	0	1	3	10	∞
e/π	0.262	0.294	0.326	0.344	0.400

Table 7 Uniformly loaded square plate on unsagged supports: amplification of central moments ($\nu = 0.3$)

$2b/\pi$	k	γ_x	γ_y
1.0	1.36	1.06	1.06
1.5	1.22	1.05	1.03
2.0	1.12	1.03	1.00
3.0	1.05	1.00	1.00

below the values for e_1/π in Table 1. The contact length e_2 is normalized first with respect to π and then $2b$; from the former values it is clear that the extent to which the plate loses contact with the supports remains unchanged for aspect ratios greater than 3.0.

From Table 1 it is obvious that, for aspect ratios $2b/\pi \geq 3$ the maximum deflection, moments, and $(V_x)_{\text{mid}}$ rapidly approach the values calculated for a simply supported rectangular plate. The loss of contact, however, is significant. Also, the support reaction $(V_y)_{\text{mid}}$ differs significantly from the simple support solution. In the case of the concentrated loading, $(M_x)_{\max}$ and $(M_y)_{\max}$ are infinite, for all aspect ratios.

For the particular use of a square plate, the redistribution in deflections, moments and support reactions caused by allowing plate uplift is shown in Figs. 3, 4, and 5 respectively. The variation in contact lengths with Poisson's ratio is shown in Table 2; for a square plate $e_1 = e_2 = e$. The variation in selected deflections, moments and support reactions with Poisson's ratio is given in Table 3 for a uniformly loaded square plate. The behavior of the latter physical quantities for a square plate loaded by a concentrated load with $\nu = 0.3$ is tabulated in Table 4.

Recently, Salamon et al. (1985) investigated, using the finite element method, the behavior of square plates resting on discrete elastic springs at the edges and subjected to uniform and concentrated loadings. The results of the latter paper are rather approximate for the case of infinitely rigid continuous supports, but qualitative agreement is obtained with the results in Table 2 and Fig. 5.

Sagged Supports. A uniformly loaded square plate in unilateral contact with sagged supports is examined in this section. The dependence of the extent of contact on the load intensity for differing amounts of sag is shown in Figs. 6(*a,b*). The contact lengths e_1 and e_2 are plotted versus q/q^0 in Fig. 6(*a*); selected values are tabulated in Table 5. The load intensity at which the plate starts to touch the sagged supports is given by q°, in which case $e_1 = \pi/2$ and $e_2 \approx 0$ (in this problem). The ratio q/q^0 is independent of the amount of sag, δ, as revealed by the expression in (D:71). In Fig. 6(*b*) the contact lengths e_1 and e_2 are plotted versus q for different values of δ. Clearly, as the amount of sag decreases, the loss of contact stabilizes for lower load levels. In the limit as $\delta \to 0$, e_1 and e_2 tend to the values indicated in Table 1 for the square plate on unsagged supports, which do not depend on the load level q. The latter observation is consistent, since the unsagged support problem is a receding contact problem.

The dependence of the edge displacements normalized with respect to the amount of sag, $w(x,0)/\delta$ and $w(0,y)/\delta$, for different values of the ratio q/q^0, is shown in Fig. 7. The variation of $W_c/\delta = w(0,0)/\delta$ with q/q^0 and e_1, e_2 is indicated in Table 5. The numerical results in this study are all for the square plate geometry and Poisson's ratio equal to 0.3.

Combined Loading

To model concentrated loading on plates including the self weight of the plates, combined loading is examined briefly. For a square plate on unsagged supports and Poisson's ratio equal to 0.3, the variation in the extent of contact with different values for the ratio P/qa^2 is indicated in Table 6. From these numbers, it can be seen that the distributed load exerts a major influence on the extent of contact.

Conclusions

The clamping effect in the vicinity of the corners of a uniformly loaded simply supported square plate is plainly illustrated by the distribution of bending moments $M_1(x,x)$ and $M_2(x,x)$ (Fig. 4(*a*)). If the corners of the plate are only partially secured against lifting, the clamping becomes ineffective and the bending moments in the outer portion of the plate increase accordingly. The amplification for unilateral edge supports with no restraint against lifting is evident in Table 1, where the simply supported and unilaterally supported values for $(M_x)_{\max}$ and $(M_y)_{\max}$ are presented. The German Code for Reinforced Concrete (1943) (see Timoshenko and Woinowsky-Krieger, 1959, p. 123) recommended that the simply supported values for $(M_x)^{\text{s.s.}}_{\max}$ and $(M_y)^{\text{s.s}}_{\max}$ be multiplied by the factor $k > 1$, where (using the current notation)

$$k = \frac{3\pi^4 - 5\pi^2 b^2 + 48b^4}{3\pi^4 - 10\pi^2 b^2 + 48b^4} \tag{13}$$

In this paper, let

$$\gamma_x = (M_x)^{\text{u.s.}}_{\max} / (M_x)^{\text{s.s.}}_{\max}, \gamma_y = (M_y)^{\text{u.s.}}_{\max} (M_y)^{\text{s.s.}}_{\max}. \tag{14}$$

For different aspect ratios, the values for k and γ_x, γ_y are presented in Table 7. Clearly, the approximate expression in (13) is always conservative; for the case of a square plate, the bending moments are overestimated by roughly 25 percent.

It is important to recognize that (D:15) can be expressed in the form

$$w(x,y) = \frac{Qa^4}{D\pi^4} \sum_{m=1,3,\ldots}^{\infty} \sum_{n=1,3\ldots}^{\infty} a_{mn}$$

$$\sin(mx)\sin(n\pi y/2b)/[m^2 + (n\pi/2b)^2]^2$$

$$+ \frac{Qa^4}{2D} \sum_{m=1,3\ldots}^{\infty} [Y^{(1)}_m(y)\sin(mx) + Y^{(2)}_m(x)\sin(m\pi y/2b)]$$

$$+ W_c \tag{15}$$

The first term corresponds to the deflection of a simply supported rectangular plate. The remaining terms give rise to the expressions involving the auxiliary $\phi_k(t)$ functions; they occur solely because the restraint on upward deflection has been removed.

The unilateral contact behavior associated with uniform pressure and concentrated loading, and combinations of the latter loadings, have been examined. The influence of the level of loading, the aspect ratio, the amoung of sag, and Poisson's ratio has been determined. As a general guideline, the simply supported deflections and moments portray the unilaterally supported behavior reasonably accurately except near the corners of the plate.

Acknowledgment

The authors would like to thank Nicholas J. Salamon and co-workers for an early draft of their paper.

References

Dempsey, J. P., Keer, L. M., Patel, N. B., and Glasser, M. L., 1984, "Contact Between Plates and Unilateral Supports," ASME Journal of Applied Mechanics, Vol. 51, pp. 324-328.

Dempsey, J. P., and Li, H., 1986, "Rectangular Plates on Unilateral Supports: Part 1—Theory and Numerical Analysis," ASME Journal of Applied Mechanics, Vol. xx, pp. ?-?.

Hansen, E. R., 1975, *A Table of Series and Products*, Prentice Hall, Englewood Cliffs, N.J.

Salamon, N. J., Pawlak, T. P., and Mahmoud, F. F., 1985, "Plates in Unilateral Contact with Simple Supports," ASME Journal of Applied Mechanics, Vol. xx, pp. ?-?.

Timoshenko, S. P., and Woinasky-Kreiger, S., *Theory of Plates and Shells*, McGraw-Hill, New York, 1959.

Transient Gas or Liquid Flow Along a Preexisting or Hydraulically-Induced Fracture in a Permeable Medium

R. H. Nilson
S-CUBED,
La Jolla, CA 92038

F. A. Morrison, Jr.
Lawrence Livermore National Laboratory,
Livermore, CA 94550

Similarity solutions are derived for the transient two-dimensional flow of a gas or liquid along an isolated fracture in a permeable medium. The driving pressure at the fracture inlet is constant, and the confining stress is uniform. Two different cases are considered, preexisting fractures with uniform aperture as well as hydraulic fractures with a variable aperture proportional to the local overpressure (fluid pressure less confining stress). The evolution of the pressure distribution is described by a set of four asymptotic solutions, each having a self-similar form. At early times the flow in the fracture is turbulent, and Darcian seepage losses into the porous surroundings are negligible. At late times the flow in the fracture is laminar, and seepage losses become a dominant consideration. At intermediate times there are two alternative asymptotes, depending upon the physical parameters. The mathematical model also describes the flow along a fracture which is filled with high-permeability porous material as well as the flow in an assemblage of porous blocks.

I Introduction

Flow along narrow channels or fractures having permeable or impermeable walls is important in a number of engineering technologies. Most notable are the geologic energy-extraction applications involving gas or liquid flows in naturally fractured [1] or hydraulically fractured media [2]. The particular application which motivates the present study is the containment evaluation of underground nuclear tests where radioactive gasses may flow outward from the cavity along a number of possible paths, including: preexisting or explosion-induced fractures, hydraulically-driven fractures, rubblized fracture/block media, permeable stemming column, and grouted or bundled electrical cables. In each instance there is a preferred channel or pathway through a permeable surrounding.

In the two-dimensional configuration of Fig. 1, a preexisting fracture (or high-permeability porous layer) penetrates from the boundary into the interior of a permeable medium. Transient fluid motion is induced by an abrupt change in the pressure at the entrance to the channel. Such flows have been previously studied in petroleum applications [3, 4] where it is usually presumed that the fluid has a small and constant compressibility, as appropriate for a liquid, but not for a gas. Also, it is usually presumed that the flow along the fracture is laminar or Darcian, although the non-Darcy flow of a liquid has been previously addressed by Guppy, Cinco-Ley, and Ramey [5] using Laplace transforms and numerical techniques.

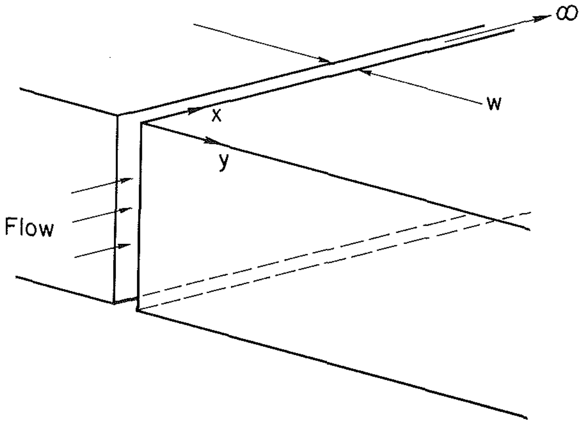


Fig. 1 Preexisting fracture with uniform aperture, w. Domain extends to infinity in the x and y directions. Height in the z direction is large compared to w.

Creation of a flow channel by the action of internal pressure is commonly referred to as hydraulic fracturing. The most familiar applications are stimulation of oil and gas wells [2] and blasting of rock formations [6]. The fracture geometry is often presumed to be planar, as illustrated in Fig. 2, where the fracture height is fixed and the length increases with time. If the fracture length is shorter than the height, the local aperture at any cross section depends upon the pressure distribution along the entire fracture, as assumed in the analyses of Geertsma and DeKlerk [7], Daneshy [8], and Nilson [9]. If,

---

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April 17, 1984; final revision, May 29, 1985.

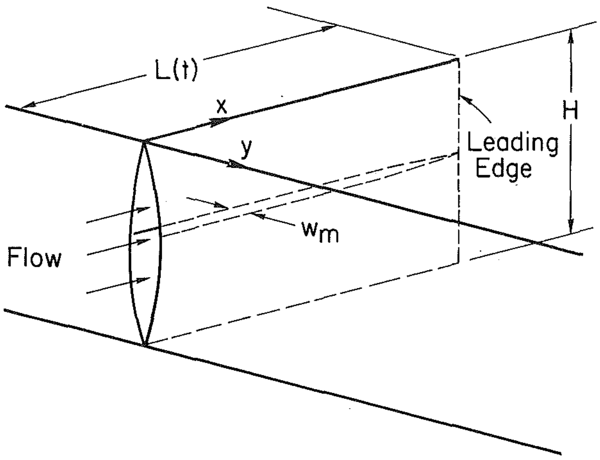


Fig. 2 Hydraulically-driven fracture with fixed height, *H*, and increasing length, *L(t)*. Cross section is elliptical with maximum opening displacement, w_m, proportional to local excess of internal fluid pressure over confining stress.

however, the fracture length is considerably greater than the height, the local aperture depends mainly on the local pressure at any cross section, as assumed in the model of Perkins and Kern [10] and Nordgren [11] and in the present paper.

Although the geometries considered here have been previously investigated by petroleum engineers, the present nuclear containment application requires a different class of solutions. Here, the driving pressure is prescribed, rather than the flow rate. The fluid may be either an ideal gas or a liquid, the friction may be laminar or turbulent, the fracture may be either preexisting or hydraulically induced, and the Darcian leakage into the walls of the fracture may be either a dominant or a negligible consideration. Of this very broad class of problems, only a small subset has been previously explored. Moreover, the comprehensive results reported here allow direct comparison among solutions which exercise a diversity of physical mechanisms. Also, the present solution procedure represents a somewhat different and relatively general approach which is currently being extended to more complex problems, such as evaporating and condensing flows of steam/water/air in fractured, permeable media.

The evolution of the transient flow is described as a sequence of distinct asymptotic domains in which different physical mechanisms are dominant.

1 At early times the seepage losses from the fracture into the porous medium are negligible; at late times these losses are dominant.

2 At early times the fracture flow has a large Reynolds number and inertial or turbulent resistance is therefore dominant; at late times the Reynolds number is small and molecular or laminar friction is dominant.

Within each of the four possible flow regimes (lossless/turbulent, lossless/laminar, loss-dominated/turbulent, loss-dominated/laminar) the two-dimensional time-dependent solutions $P(x,y,t)$ possesses a self-similar form $P(\theta(x,t), \eta(y,t))$ which simplifies the computation and facilitates a comprehensive presentation of results. The present nonlinear solutions with density variation and non-Darcy friction reduce to the known linear solutions [5, 12] when the pressure ratio is near unity and the Reynolds number is small.

II Formulation

The two related geometries illustrated in Figs. 1 and 2 will be analyzed in a parallel fashion. For either geometry the flow along the fracture is assumed to be one-dimensional in the x-direction, and the Darcian seepage into the surroundings is assumed to be two-dimensional in the xy-plane, as appropriate when the fracture height, H, is large compared to the aperture, w.

The preexisting fracture of Fig. 1 has a prescribed aperture which is uniform along the fracture and invarient in time. The aperture could, however, vary in the z-direction (as in the hydrofracture geometry of Fig. 2) in which case an effective aperture or hydraulic diameter should be used in the one-dimensional flow analysis. The preexisting fracture may be either an open channel or, alternatively, it may be filled with a high-permeability porous material.

The hydraulically-driven fracture of Fig. 2 has a variable aperture which depends linearly on the local overpressure, $P(x) - \sigma$, as in the Perkins/Kern/Nordgren model where the maximum displacement at the center of the channel, the cross-sectional area, and the effective aperture are, respectively, given by [10, 11]

$$w_m(x) = \frac{(1-\nu)H}{G}[P(x) - \sigma], \quad A = \frac{\pi}{4} w_m H, \quad w = \frac{\pi}{4} w_m \tag{1}$$

in which G and ν are the shear modulus and Poisson's ratio, σ is the confining stress acting normal to the fracture plane, and H is the height of the fracture which is presumed to be known and constant. Thus the normalized aperture and area, w^* and A^*, are each proportional to the normalized overpressure P^*, such that

$$w^* = A^* = P^* \tag{2}$$

where

$$w^* = \frac{w}{w_o}, \quad A^* = \frac{A}{A_o}, \quad P^* = \frac{P - \sigma}{P_o - \sigma} \tag{3}$$

in which the subscript zero refers to the inlet of the channel ($x = 0$) where the pressure is known, and hence the local channel dimensions, w_o and A_o, can be readily calculated from the formulas given in (1). The normalized geometric relationships (2, 3) are generally applicable in any situation where a flow channel is created by displacements, either compressions or tensions, which are proportional to local overpressure.

The one-dimensional transient flow within a permeable-walled channel is governed by the conservation of mass and momentum [13]

$$\phi_c \frac{\partial}{\partial t}(A\rho) + \frac{\partial}{\partial x}(A\rho u) = -C\rho v_w \tag{4}$$

$$\frac{\partial}{\partial t}(A\rho u) + \frac{\partial}{\partial x}(A\rho u^2) = -A\left(\frac{\partial P}{\partial x} + \rho F\right) \tag{5}$$

in which A and C are the cross-sectional area and circumference of the channel, P is pressure, ρ is density, u is longitudinal velocity, v_w is transverse velocity into the permeable walls of the channel, and F embodies the frictional effects in the fluid. Note that $A \sim Hw$ may vary with x, while $C \sim 2H$ is always constant.

The frictional forces can be expressed by the following linear combination which has the correct asymptotic behavior both for small and for large values of the Reynolds number ($\mathrm{Re} = \rho u d/\mu$)

$$\rho F = C_l \frac{\mu u}{d^2} + C_t \frac{\rho u^2}{d} \tag{6}$$

These equations (4-6) apply equally well in either of the following circumstances, provided that the channel porosity, ϕ_c, the length scale, d, and the dimensionless coefficients, C_l and C_t, are appropriately defined.

1 For flow in an open channel

$$\phi_c = 1 \tag{7a}$$

$$d = w = \text{effective channel aperture} \tag{7b}$$

$$C_l = 12 \text{ for preexisting fracture} \tag{7c}$$

$$= 4\pi \text{ for hydraulic fracture [7]}$$

$$C_t = a(\epsilon/w)^b \tag{7d}$$

in which $C_l = 12$ is based on the laminar Poiseuille flow, C_t is based on experimental results for turbulent flow in geologic fractures ([14]; $a \simeq 0.1$, $b \simeq 0.5$), and ϵ is roughness height of the channel wall.

2 For flow in a channel which is packed with a permeable material (e.g., propped hydrofrac or high permeability stratigraphic layer)

$$\phi_c = \text{porosity of channel material} \tag{8a}$$

$$d = \text{characteristic microscale dimension (e.g., pore dimaeter or particle size)} \tag{8b}$$

$$C_l = d^2/K_c \simeq 180(1-\phi_c)^2/\phi_c^3 \tag{8c}$$

$$C_t = \lambda d/K_c \simeq 2(1-\phi_c)/\phi_c^3 \tag{8d}$$

in which C_l is based on Darcy's law, K_c is the permeability of the channel, and λ is Ergun's constant which is a measurable macroscopic property of the pore structure. Roughly, $K_c \simeq d^2 \phi_c^3/180(1-\phi_c)^2$ and $\lambda \simeq 0.012\ d/(1-\phi_c)$ [15] as noted in (8*c*) and (8*d*). In many geological applications, the conventional friction models described above can only be viewed as crude approximations, particularly when the roughness height is comparable to the fracture aperture and the expected behavior lies somewhere between the extremes of channel flow and porous flow.

The lateral loss-velocity, v_w in (4), accounts for the seepage losses from the channel. It can be determined by consideration of the two-dimensional transient flow field within the surrounding porous medium, as governed by conservation of mass and by Darcy's law, respectively,

$$\phi_m \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{v}) = 0 \tag{9}$$

$$\bar{v} = -\frac{K_m}{\mu} \nabla P \tag{10}$$

in which ϕ_m and K_m are porosity and permeability of the surrounding medium. Since $K_m << K_c$, the superficial velocity, v, is presumed small enough to neglect the high-Re non-Darcy effects which were retained in the channel flow. Equation (10) implies that the seepage-loss velocity v_w in (4) is related to the lateral pressure gradient at the fracture wall, i.e., $v_w = -(K_m/\mu)\partial P/\partial y$ on $y = 0$.

The barotropic equation of state, $\rho = \rho(P)$ is often appropriate, both in gas flows and in liquid flows. Ideal gas flows ($\rho = P/RT$) are often nearly isothermal because the ambient temperature of the matrix cannot be substantially perturbed by the through-flow of a gas which has a relatively small specific heat. Liquid flows are usually characterized as having a small (and constant) compressibility, α, even in nonisothermal circumstances. In either instance, liquid or gas, the time derivative of the density can be replaced as follows

$$\frac{\partial \rho}{\partial t} = \rho \alpha \frac{\partial P}{\partial t} \tag{11}$$

in which the product $\rho\alpha$ is essentially a constant (for the gas, $\rho\alpha = \rho/P = 1/RT$).

The flow is induced by an abrupt change in the pressure along the boundary at $x = 0$, as described by the initial/boundary conditions.

$$P(x,y,0) = P_\infty;\ P(0,y,t) = P_o \tag{12}$$

where P_∞ is the ambient pore pressure in the permeable medium. For a preexisting fracture, the internal pressure within the fracture approaches P_∞ as $x \to \infty$, and the fluid velocity may be either positive ($u > 0$) or negative ($u < 0$) depending on whether $P_o > P_\infty$ or $P_o < P_\infty$. For a hydraulic fracture, however, the internal pressure approaches σ at the leading edge, and it is necessary that $P_o > \sigma$ for the fracture to be open, in which case $u > 0$. In general, P_∞ might be greater or less than σ, but the former case ($P_\infty > \sigma$) is unstable in the sense that fracture propagation can occur spontaneously, without the imposition of the disturbance pressure, P_o, at the boundary. So, here we restrict to the latter case ($P_\infty < \sigma$) which is far more common in the applications.

At very early times the inertial terms on the left side of the *momentum equation* (5) *are dominant, and the considered* channel flow resembles the flow in a shock tube. As the time and length of run increase, the frictional, F, terms in (5) smear out the shockwave and the inertial terms in equation (5) become negligible (since $u\ \partial u/\partial x << u^2/d$ in (5) and (6)). After that, the friction dominated channel-flow gradually slows down as the pressure gradient diminishes, and eventually the lateral seepage losses become a dominant consideration, as described by the present analysis.

III Similarity Transformations

The transient solution of the stated problem can be described by a set of four self-similar asymptotic solutions in each of which the independent variables are

$$\theta = \frac{x}{L_r f(\tau)},\ \eta = \frac{y}{\sqrt{t}} \left(\frac{2N}{N+1} \frac{\phi_m \mu \alpha_m}{K_m} \right)^{1/2},\ \tau = \frac{t}{t_r} \tag{13}$$

where the pore-fluid compressibility parameter, α_m, is either the compressibility of a liquid or $1/P_o$ for a gas. The dependent variables are normalized as follows

$$P^* = \frac{P - P_r}{P_o - P_r},\ \rho^* = \frac{\rho}{\rho_r},\ u^* = \frac{u}{u_r g(\tau)} \tag{14}$$

The time-functions $f(\tau)$ and $g(\tau)$, which presently remain arbitrary, will later be chosen such that P^*, ρ^* and u^* depend only upon θ and η within the four distinct, self-similar time regimes.

The reference scales of length, velocity, time and density are chosen such that P^*, ρ^*, and u^* are always positive, regardless of whether the flow is inward or outward (± below for $u \gtreqless 1$)

$$\rho_r = \frac{\rho(P_o) + \rho(P_r)}{2} \qquad u_r = \pm \frac{\mu}{\rho_r d_o} \frac{C_l}{C_t}$$

$$L_r = \frac{\Delta P d_o^2}{u_r \mu C_l} \qquad t_r = \frac{L_r}{u_r} \frac{2N}{N+1} \phi_c \alpha_c \Delta P \tag{15}$$

in which $\rho(P_o)$ and $\rho(P_r)$ are the fluid densities at pressures P_o and P_r, respectively, and the quantities d_o and C_t are evaluated at the inlet of a hydrofracture. The reference pressure, P_r, is taken as the pressure at the leading edge of the channel flow, so that

$$P_r = P_\infty \text{ for a preexisting fracture} \tag{16a}$$

or

$$P_r = \sigma \text{ for a hydraulic fracture} \tag{16b}$$

In either case, the characteristic pressure difference along the fracture and the corresponding pressure ratio are defined as

$$\Delta P = |P_o - P_r| \text{ and } N = P_o/P_r \tag{17}$$

The compressibility parameter of the channel, α_c, appearing in the timie scale, t_r, is taken as either

$$\alpha_c = \alpha_m \text{ for a preexisting fracture} \tag{18a}$$

or

$$\alpha_c = 1/\Delta P \text{ for a hydraulic fracture} \tag{18b}$$

The lateral seepage flow is driven by the difference between the fracture pressure and the ambient pore pressure, as measured by

$$\Delta P_s = P_o - P_\infty \text{ or } \Delta P_s^* = \frac{\Delta P_s}{\Delta P} = \frac{P_o - P_\infty}{P_o - P_r} \tag{19}$$

For a preexisting fracture $\Delta P_s^* = 1$, since $P_r = P_\infty$ in that circumstance. However, for a hydraulic fracture ΔP_s^* might be greater than unity if the pore pressure, P_∞, were less than the confining stress, $\sigma = P_r$.

Under the transformations stated just above, the space and time derivatives can be replaced by

$$\frac{\partial}{\partial x} = \frac{1}{L_r f} \frac{\partial}{\partial \theta}, \quad \frac{\partial}{\partial t} = \frac{1}{t_r}\left(-\theta \frac{\dot{f}}{f}\frac{\partial}{\partial \theta} - \eta \frac{1}{2\tau}\frac{\partial}{\partial \eta}\right) \tag{20}$$

whereupon the problem statement (4–6, 9–12) transforms into the following format in which ($\dot{}$) represents differentiation with respect to τ

$$-\frac{\dot{f}}{g}\theta\left[\frac{N+1}{2N}\rho^* \frac{\partial A^*}{\partial \theta} + \left(\frac{N-1}{N}\right)^n A^* \frac{\partial P^*}{\partial \theta}\right] + \frac{\partial}{\partial \theta}(A^* \rho^* u^*) = \frac{R^{1/2}}{\Delta P_s^*} \frac{f}{g\sqrt{\tau}} \rho^* \frac{\partial P^*}{\partial \eta} \tag{21}$$

$$-\frac{\partial P^*}{\partial \theta} = fg \frac{u^*}{w^{*2}} + fg^2 \frac{\rho^*}{w^{*(1+b)}} u^* |u^*| \tag{22}$$

$$-\tau \frac{\dot{f}}{f}\theta \frac{\partial P^*}{\partial \theta} - \frac{1}{2}\eta \frac{\partial P^*}{\partial \eta} = \frac{\partial}{\partial \eta}\left(\rho^* \frac{\partial P^*}{\partial \eta}\right) + \frac{\tau}{f^2} M \frac{\partial}{\partial \theta}\left(\rho^* \frac{\partial P^*}{\partial \theta}\right) \tag{23}$$

$$\rho^* = \frac{2N}{N+1}\left[P^*\left(1 - \frac{1}{N}\right) + \frac{1}{N}\right] \tag{24}$$

with boundary conditions

$$P^*(0,0,\tau) = 1; \; P^*(\infty,\eta,\tau) = P^*(\theta,\infty,\tau) = 1 - \Delta P_s^* \tag{25}$$

Equations (21) and (22) describe the flow along the fracture, while the nonlinear diffusion equation (23) describes the two-dimensional seepage flow in the permeable surroundings. The main difference between preexisting and hydraulic fractures is that $w^* = A^* = 1$ in the former case, whereas $w^* = A^* = P^*$ in the latter. Also the exponent n in (21) is either 0 or 1 for the respective cases of preexisting and hydraulic fractures, because of the scaling chosen in equation (18).

The lateral seepage-loss parameter which appears on the right side of equation (21) is defined as follows:

$$R = K_m C_l \frac{\phi_m \alpha_m}{\phi_c \alpha_c}\left[\left(\frac{C}{A}\right)_o \frac{L_r}{d_o} \frac{\Delta P_s}{\Delta P}\right]^2 \tag{26}$$

in which $(C/A)_o \sim 2/w_o$ is the ratio of circumference to cross sectional area at the fracture inlet. Note that the inertial terms have been deleted from the momentum equation (22) as discussed earlier. Furthermore, the longitudinal θ-diffusion term in (23) is negligible for all times of practical interest, since $M = C_l\, \phi_c \alpha_c K_m / (\phi_m \alpha_m d_o^2) << 1$ [12]. Finally, note that in the limit as $N \rightarrow 1$ the ideal gas equation of state (24) also describes a slightly compressible liquid (for which $\rho^* \simeq 1$), so that the liquid-flow results for all pressure ratios are recoverable as a degenerate case of the gas-flow results [16].

The self-similar regimes will now be identified by choosing $f(\tau)$ and $g(\tau)$ such that the problem statement (21–25) becomes independent of τ within specified ranges (early, intermediate, late) of τ. The results for each self-similar regime will first be presented, thereafter giving consideration to the transition behavior.

IV Lossless Regimes

At sufficiently early times ($\tau << 1/R$), the lateral fluid loss from the fracture is insignificant, and the pressure wave penetrates into the fracture almost as though its walls were impermeable. The appropriate transformations for these quasi-lossless regimes are listed below for the separate cases of laminar and turbulent flow.

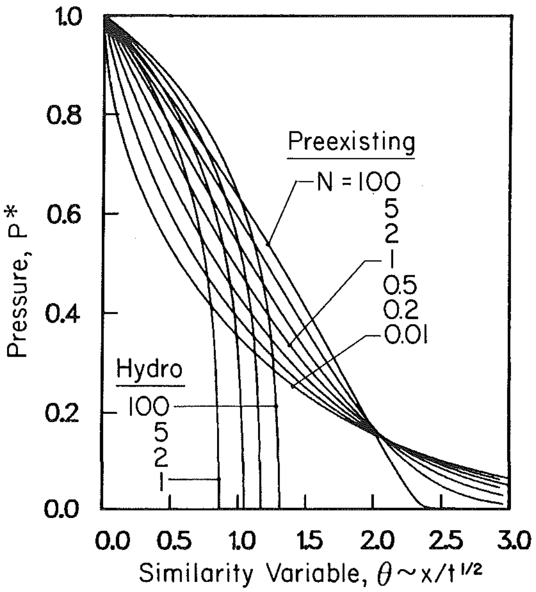


Fig. 3 Pressure distribution along preexisting and hydraulically-driven fractures for lossless flow with laminar friction

1 For laminar flow, let $\dot{f} = g/2$ and $fg = 1$ such that

$$f = \tau^{1/2}, \quad g = \tau^{-1/2} \tag{27}$$

$$\theta = \frac{x}{t^{1/2}}\left(\frac{2N}{N+1}\frac{\phi_c \alpha_c \mu C_l}{d_o^2}\right)^{1/2} \tag{28}$$

and the momentum equation (22) becomes

$$-\frac{\partial P^*}{\partial \theta} = \frac{u^*}{w^{*2}}\left(1 + \frac{\rho^* u^* w^{*(1-b)}}{\tau^{1/2}}\right) \simeq \frac{u^*}{w^{*2}} \text{ for } \tau >> 1 \tag{29}$$

2 For turbulent flow, let $\dot{f} = 2g/3$ and $fg^2 = 1$ such that

$$f = \tau^{2/3}, \; g = \tau^{-1/3} \tag{30}$$

$$\theta = \frac{x}{t^{2/3}}\left[\left(\frac{2N}{N+1}\right)^2 \frac{\phi_c^{\,2} \alpha_c^{\,2} \Delta P\, C_t \rho_r}{d_o}\right]^{1/3} \tag{31}$$

and the momentum equation (22) becomes

$$-\frac{\partial P^*}{\partial \theta} = \frac{\rho^* u^* |u^*|}{w^{*(1+b)}}\left(1 + \frac{\tau^{1/3}}{\rho^* u^* w^{*(1-b)}}\right) \simeq \frac{\rho^* u^* |u^*|}{w^{*(1+b)}} \text{ for } \tau << 1 \tag{32}$$

Thus, the momentum equation becomes independent of τ, both for $\tau >> 1$ and for $\tau << 1$. In either case, the continuity equation (21) takes the following form in which $\beta = 1/2$ or $2/3$ for laminar and turbulent cases, respectively.

$$-\beta\theta\left[\frac{N+1}{2N}\rho^* \frac{\partial A^*}{\partial \theta} + \left(\frac{N-1}{N}\right)^n A^* \frac{\partial P^*}{\partial \theta}\right] + \frac{\partial}{\partial \theta}(A^* \rho^* u^*) = (R\tau)^{1/2} \frac{\rho^*}{\Delta P_s^*} \frac{\partial P^*}{\partial \eta} \simeq 0 \text{ for } R\tau << 1 \tag{33}$$

It is seen that the lateral loss term is negligible and the continuity equation becomes independent of τ so long as $\tau << 1/R$, and for those early times the channel flow is essentially uncoupled from the flow in the surrounding medium.

To facilitate a numerical solution, the expressions for $w^*(P^*)$ $\rho^*(P^*)$, and $u^*(dP^*/d\theta, w^*)$ from equations (2, 24) and (29) or (32) are all substituted into the continuity equation (33) to obtain a second-order ordinary differential equation for $P^*(\theta)$ subject to the boundary conditions $P^*(0) = 1$ and $P^*(\infty) = 0$. The derivative operators are then replaced by central-differences on a discrete grid, and the resulting equations for $P_i^*(\theta_i)$ are linearlized to obtain a tridiagonal

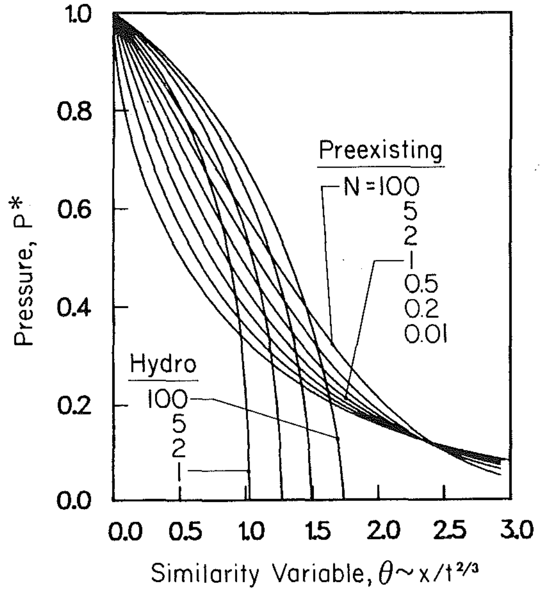


Fig. 4 **Pressure distribution along preexisting and hydraulically-driven fractures for lossless flow with turbulent friction**

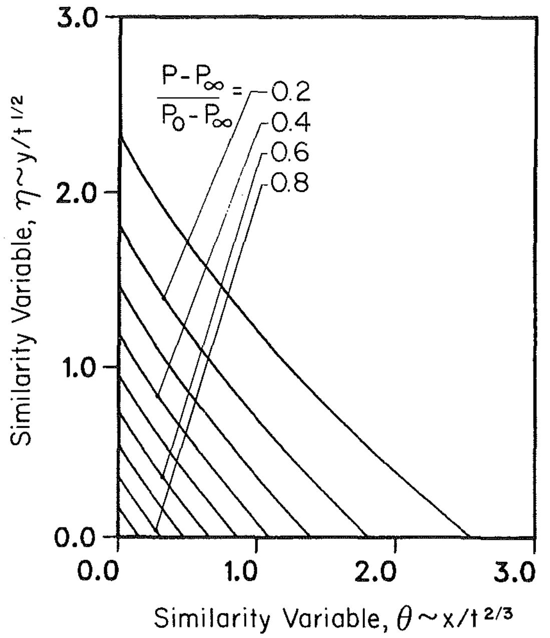


Fig. 5 ***Isobars for lossless flow along a preexisting fracture with turbulent friction. Pressure ratio is near unity ($N \rightarrow 1$).***

algebraic system which is solved by the algorithm of Thomas. Iterations are continued until the $P_i{}^*$ used in the linearization are in close agreement with the $P_i{}^*$ which satisfy the linearized equations.

Laminar solutions with negligible seepage losses are illustrated in Fig. 3 for various values of the pressure ratio, N, which is the only parameter appearing in the problem statement. The lossless flow along a preexisting fracture with $w^*=A^*=1$ is mathematically equivalent to one-dimensional flow in a permeable medium [16, 17, 18]; these solutions are reproduced here as a check on the present finite-difference method and for comparison with the hydrofracture results. The pressure profiles for preexisting fractures ($N=P_o/P_\infty$) generally exhibit a gradual exponential approach to the far field pressure as θ tends to infinity, whereas the profiles for hydraulic fractures ($N=P_o/\sigma$) drive abruptly to $P^*=0$ ($P=\sigma$) at a finite θ near unity. The curves in Fig. 3 also describe the shape of a hydrofracture, since $w^*=P^*$ when local displacement is proportional to local over-pressure. Hydrofracture solutions exist only for $N>1$, since P_o must be greater than σ for the fracture to be open, whereas rarefaction or "drawdown" solutions with $N<1$ are obtainable for a preexisting fracture with fixed aperture.

The leading edge of the hydrofracture occurs at a finite location, θ^*, where P^* goes to zero. Behind that point $P^*>0$, or equivalently $P>\sigma$, and the channel is open. Further ahead, the channel is closed. Right at the leading edge, the fluid velocity must be the same as the propagation velocity of the hydrofracture, requiring that

$$u=\left.\frac{dx}{dt}\right|_\theta \text{ or } u^*=\beta\theta^*\left(\frac{N+1}{2N}\right) \text{ at } \theta=\theta^* \qquad (34)$$

Otherwise there would be fluid passing through the front of the fracture. Recalling that the fluid speed depends upon the aperature and the pressure gradient, the local pressure gradient is theoretically infinite at the leading edge. For a laminar flow,

$$\frac{dP^*}{d\theta}=-\frac{\theta^*}{2w^{*2}}\left(\frac{N+1}{2N}\right) \text{ at } \theta^* \qquad (35)$$

where w^* ($=P^*$) goes to zero. This equation implies that the pressure has a cube root singularity at the leading edge of a laminar flow. Care was therefore taken to accommodate the appropriate singular behavior in the numerical algorithm.

Turbulent flow solutions with negligible seepage losses are illustrated in Fig. 4. As in laminar flow, the disturbance extends to infinity along a preexisting fracture, whereas a hydrofracture disturbance is limited to a finite domain. As before, there is a theoretically-infinite pressure gradient at the leading edge of a hydrofracture, but the singularity is now somewhat weaker, particularly when N is large.

The lossless solutions for preexisting versus hydraulic fractures are very nearly the same, provided that the fluid is a gas and the pressure ratio is large. When N is large, the similarity variables for the two types of fractures are virtually identical, because in that limit the compressibility parameter, $\alpha_c=1/\Delta P$, used in scaling hydrofractures is equivalent to the $\alpha_c=1/P_o$ used in scaling preexisting fractures. Of course, the penetration depth is always somewhat shorter for a hydraulic fracture, compared to a preexisting fracture having the same inlet aperture. But the difference in penetration depth and fluid speed is only a factor of two, or so, when N is large. The two families of solutions are very different, however, when the pressure ratio is small or the fluid is a liquid, because their scaling is then quite different. The penetration depth along preexisting fractures is proportional to $(1/\alpha_c)^m \sim P_o^m$ where $m=1/3$ or $2/3$ for laminar or turbulent flows; whereas for hydraulic fractures the penetration is only proportional to $(1/\alpha_c)^m \sim \Delta P^m$. So, when $\Delta P/P_o << 1$ (i.e., $N \sim 1$), the fluid penetrates into a hydraulic fracture much more slowly than a preexisting fracture, both having the same aperture at the inlet.

This qualitative difference is a consequence of the very different capacitance mechanisms which are operative in the two types of fractures. For a hydrofracture, there are two capacitance mechanisms, fluid compression ($\sim w^*\ \Delta\rho^*$) and solid deformation ($\sim \rho^*\ \Delta w^*$); the latter mechanism always remains of order one, even in the limit as $N \rightarrow 1$ and $\Delta\rho^* \rightarrow 0$. Conversely, for a preexisting fracture the only capacitive mechanism is fluid compression ($\sim w^*\ \Delta\rho^*$) which is of the order $\Delta\rho^*=(N-1)/N$, and hence very small for N near unity or for an "incompressible" liquid. That is the primary difference in the physical processes and the reasoning behind the difference in scaling. At large N, this distinction is of no importance because $\Delta\rho^*$ is then comparable to ΔP^*, and hence fluid compressibility effects are of order unity for both types of fractures.

Although seepage losses do not significantly influence the early-time flow along the channel, the pressure disturbance does penetrate into the surrounding porous medium, in the manner described by the following partial differential equation (from (23) with $\beta = 1/2$ or $2/3$ for laminar and turbulent flow, respectively)

$$-\beta\theta\frac{\partial P^*}{\partial\theta}-\frac{1}{2}\eta\frac{\partial P^*}{\partial\eta}=\frac{\partial}{\partial\eta}\left(\rho^*\frac{\partial P^*}{\partial\eta}\right) \tag{36}$$

For laminar flow in a preexisting fracture, the isobars are simply straight lines which intersect the θ-axis at a $\pi/4$ angle. This can be verified by introducing the independent variable $\xi = \theta + \eta$ which reduces all of the equations (29, 33, 36) to a single ordinary differential equation, $\xi P^{*\prime} + 2(\rho^* P^{*\prime})' = 0$; hence, P^* is a function of ξ alone. Under more general circumstances the isobars are determined by numerically solving the field equation (36), as outlined in the next section, using the channel flow solutions as boundary data on the fracture (i.e., on $\eta = 0$). In the typical results of Fig. 5 the lateral η-penetration and the longitudinal θ-penetration are both of order one.

IV Loss-Dominated Regimes

As the penetration depth increases, the lateral seepage losses become progressively more important. At sufficiently late times ($\tau >> 1/R$), the longitudinal through-flow in the channel is almost entirely consumed by the seepage losses, and the fluid capacitance within the channel becomes negligible. The appropriate transformations for these loss-dominated regimes are listed below for the separate cases of laminar ($\tau >> 1$) and turbulent ($\tau << 1$) flow along the fracture.

1 For laminar flow, let $f = g(\tau/R)^{1/2}$ and $fg = 1$ such that

$$f = (\tau/R)^{1/4},\ g = (\tau/R)^{-1/4}, \tag{37}$$

$$\theta = \frac{x}{t^{1/4}}\left[\frac{2N}{N+1}\,\frac{\phi_m\alpha_m\mu K_m C_l^2}{d_o^{\,4}}\left(\frac{\Delta P_s}{\Delta P}\right)^2\left(\frac{C}{A}\right)_o^2\right]^{1/4} \tag{38}$$

and the momentum equation (22) becomes

$$-\frac{\partial P^*}{\partial\theta}\simeq\frac{u^*}{w^{*2}}\ \text{for}\ \tau>>1 \tag{39}$$

2 For turbulent flow, let $f = g(\tau/R)^{1/2}$ and $fg^2 = 1$ such that

$$f = (\tau/R)^{1/3},\ \ g = (\tau/R)^{-1/6} \tag{40}$$

$$\theta = \frac{x}{t^{1/3}}\left[\frac{2N}{N+1}\,\phi_m\alpha_m\Delta P\,\frac{K_m C_t\rho_r}{d_o\mu}\left(\frac{\Delta P_s}{\Delta P}\right)^2\left(\frac{C}{A}\right)_o^2\right]^{1/3} \tag{41}$$

and the momentum equation (22) becomes

$$-\frac{\partial P^*}{\partial\theta}\simeq\rho^* u^*\,|u^*|\,w^{*-(1+b)}\ \text{for}\ \tau<<1 \tag{42}$$

In either case, the continuity equation (21) takes the following form

$$\frac{\partial}{\partial\theta}(A^*\rho^*u^*)-\frac{\rho^*}{\Delta P_s^*}\frac{\partial P^*}{\partial\eta}=\beta\theta(R\tau)^{-1/2}\left[\frac{N+1}{2N}\rho^*\frac{\partial A^*}{\partial\theta}\right.$$
$$\left.+\left(\frac{N-1}{N}\right)^n A^*\frac{\partial P^*}{\partial\theta}\right]\simeq 0\ \text{for}\ R\tau>>1 \tag{43}$$

It is seen that the capacitance term on the right becomes negligible when $\tau >> 1/R$, and in that limit the equation becomes independent of τ. But, in contrast to the early lossless regime, $\partial P^*/\partial\eta$ now appears in the channel-flow equations, so they must now be solved in conjunction with the porous-flow equation, (36), in which $\beta = 1/4$ or $1/3$ for the laminar and turbulent cases, respectively.

Finite-difference solutions are obtained by successive iterations of a two step procedure. First, the pressure variation along the fracture is calculated by solving the continuity equation (43) together with the friction equation, (39) or (42), and the usual expressions for $\rho^*(P^*)$, $w^*(P^*)$, and $A^*(P^*)$. The numerical approach is basically the same as in the lossless case described previously, except that the lateral seepage term, $\partial P^*/\partial\eta$ in (43), is evaluated using most recent estimates of the pore-pressure distribution in the fracture wall. In the second step, the pore-pressure field in the permeable surroundings is calculated from the parabolic partial differential equation, (36). The integration algorithm marches inward toward the origin in the time-like negative-θ direction. Using backward differences for $\partial P^*/\partial\theta$ and central differences for the η-derivatives, we obtain a tri-diagonal algebraic system for the pressures on each successive line of constant θ. The sweep is initiated by letting $P^* = 1 - \Delta P_s^*$ along a remote line, $\theta = \theta_{\max}$. The most recent calculation of the fracture flow provides boundary data on $\eta = 0$, and it is also required that $P^* = 1 - \Delta P_s^*$ along a remote line, $\eta = \eta_{\max}$. As $\theta \to 0$, the partial differential equation degenerates into an ordinary differential equation in the η-variable. As explained in [12], the present solutions are "outer solutions" which do not satisfy the boundary conditions along $\theta = 0$, except at the point (0,0). There is a thin boundary-layer along $\theta = 0$, wherein $\partial^2 P^*/\partial\theta^2$ must be

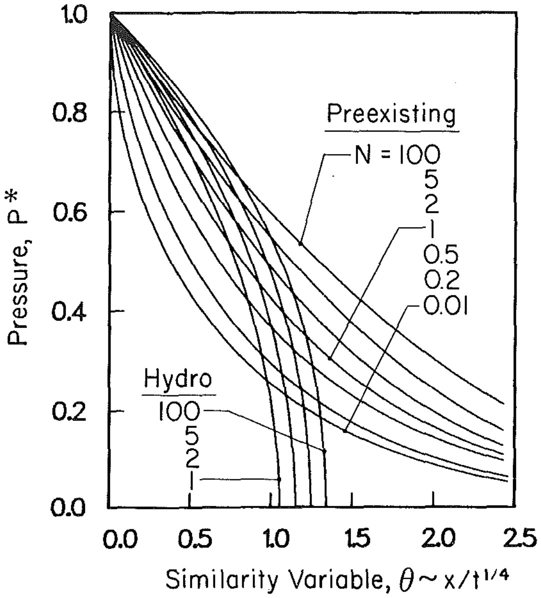


Fig. 6 Pressure distribution along preexisting and hydraulically-driven fractures for loss-dominated flow with laminar friction

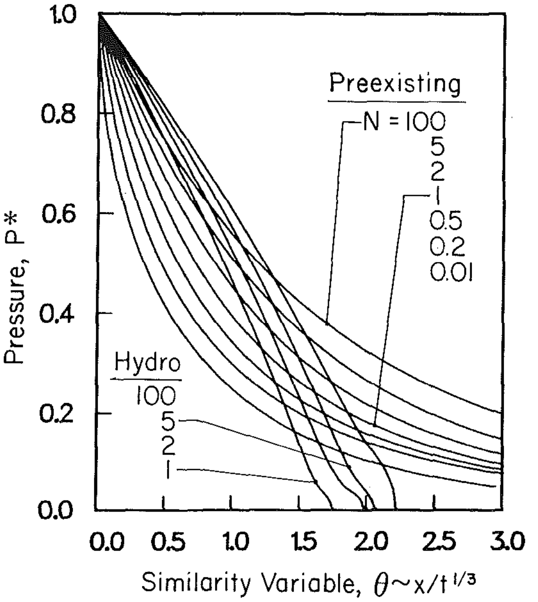


Fig. 7 Pressure distribution along preexisting and hydraulically-driven fractures for loss-dominated flow with turbulent friction

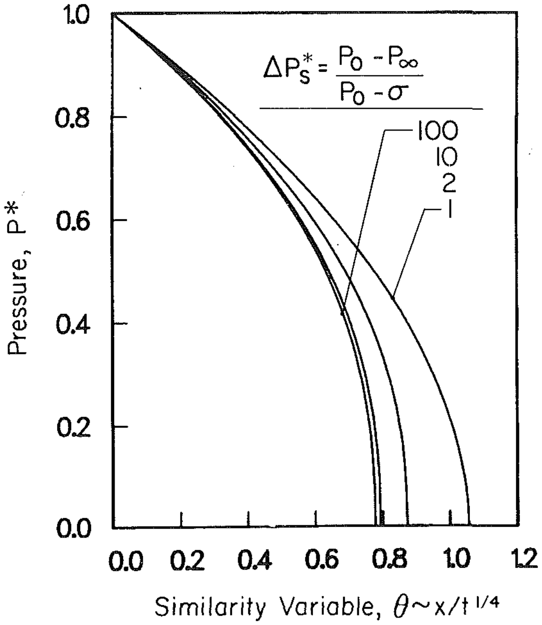


Fig. 8 Pressure distribution along hydraulically-driven fractures for loss-dominated flow with laminar friction. Pore pressure P_∞ is equal to or less than confining stress, σ. Pressure ratio is near unity ($N \to 1$).

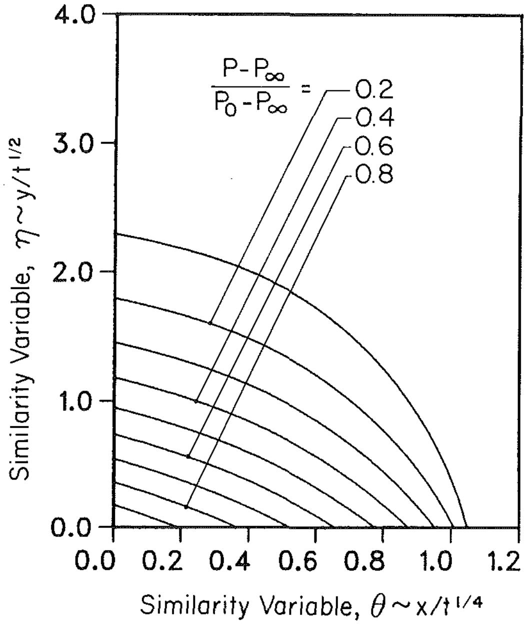


Fig. 9 Isobars for loss-dominated flow along a hydraulically-driven fracture with laminar friction. Pore pressure is the same as confining stress ($\Delta P_s^* = 1$), and pressure ratio is near unity ($N \to 1$).

retained and the equation becomes elliptic. However, the present outer solutions remain valid until the very late times noted in the summary of this paper.

Laminar and turbulent solutions are illustrated in Figs. 6 and 7, respectively. Each figure shows the pressure variation along the channel for various values of the longitudinal pressure ratio, $N = P_o/P_\infty$ or $N = P_o/\sigma$, for preexisting or hydraulic fractures, respectively. Recall, from equations (16) and (19), that the pore pressure parameter ΔP_s^* is always unity for preexisting fractures. Thus, in the comparative hydrofracture solutions of Figs. 6 and 7, the pore pressure P_∞ is chosen to be the same as the confining stress σ, such that $\Delta P_s^* = 1$. The disturbance extends to infinity along a preexisting fracture; whereas the length of a hydrofracture is finite, and the pressure gradient is very steep near the leading edge. It is no longer true that the fluid velocity must match the propagation velocity near the tip of the hydrofracture, but rather that the longitudinal flow be in balance with the seepage losses, which still requires that the pressure gradient become infinite as the width of the channel goes to zero. The turbulent hydrofracture solutions in Fig. 7 have a curious appearance near the leading edge, so particular care was taken to ensure adequate resolution of the detailed structure at the toe of the profile.

In the loss-dominated regimes, the fluid penetration depth along preexisting versus hydraulic fractures is quite comparable, regardless of the pressure ratio N, since the only capacitance mechanism is the fluid storage in the surrounding porous medium. Thus, the similarity variable now depends on the pore-fluid compressibility, α_m, which is the same for preexisting and hydrofractures, rather than α_c, which differs between the two.

The ambient pore pressure might, in general, be less than the confining stress and this would increase the seepage losses from the fracture. However, this consideration has been accounted for, in a first order fashion, by the scaling of the equations. Thus, as seen in Fig. 8, the pressure profile and the normalized length of the fracture are almost insensitive to the pore-pressure parameter $\Delta P_s^* = (P_o - P_\infty)/(P_o - \sigma)$. The illustrative results are for laminar flow and a worst-case condition in which $N \sim 1$. In reality, the pressure ratio N must always be greater than unity, and the maximum possible value of the pore-pressure is $\Delta P_s^* = P_o/(P_o - \sigma) = N/(N-1)$, corresponding to a pore pressure of zero.

The pressure field within the porous medium is illustrated in Figs. 9 and 10 for $\Delta P_s^* = 1$ and $\Delta P_s^* = 10$, respectively. Both plots apply to the incompressible case with $N \sim 1$. In Fig. 9, the longitudinal pressure difference ΔP is the same as the lateral pressure difference ΔP_s, and the isobars are inclined at about 45 deg, as in the comparable pressure plot of Fig. 5. In Fig. 10, however, the pressure variation ΔP along the channel is almost negligible compared to the lateral pressure difference. So, in essence, the channel appears to be a high pressure isobar, and all of the other isobars are nearly parallel to it, except at the leading edge of the channel where all of the isobars converge to a square-root singularity, as in the related heat-transfer problem of a hot isothermal plate moving through a cold, conducting medium.

The infinite gradients near the tip of a hydrofracture are more of a computational annoyance than a physical reality. There will always be some longitudinal diffusion near the tip of the fracture, and this will weaken the gradients predicted by our parabolized equations. Also, the driving fluid will probably stand back a slight distance behind the tip of the fracture, because of the difficulties of moving into the narrowest part of the channel. In reality, the fluid need not penetrate all the way to the tip in order to drive the fracture ahead, because shear forces within the solid transmit pressure loads into the region ahead of the fluid front [9]. Despite these local departures from reality, the overall character of the solution seems quite reasonable, and there is every expectation that the adjustments required to remove the singularities should have only a moderate influence on the quantities of engineering interest, such as fracture length and flow rate at the inlet.

V Summary

Transient isothermal fluid flow along an isolated fracture in a porous medium has been analyzed. Two different cases were considered, a preexisting fracture with uniform aperture (Fig. 1) as well as a hydraulically-induced fracture with local aperture proportional to local overpressure (Fig. 2). The preexisting fracture could be either open or, alternatively, filled with a high-permeability material. A one-dimensional laminar/turbulent model of the flow within the fracture was

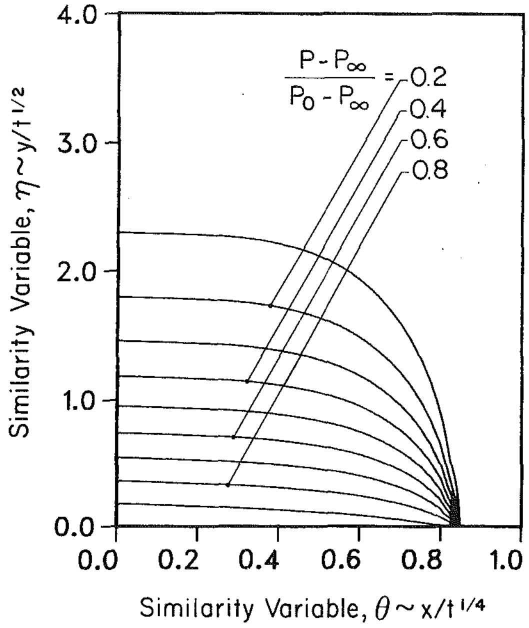


Fig. 10 Isobars for loss-dominated flow along a hydraulically-driven fracture with laminar friction. Pore pressure is much less than confining stress ($\Delta P_S^* = 10$), and pressure ratio is near unity ($N \to 1$).

coupled with a two-dimensional Darcian model of the seepage into the surrounding permeable medium. The problem was formulated such that the pressure ratio N was the only important parameter: $N>1$ for ideal gas injection; $N<1$ for ideal gas withdrawal; and $N \to 1$ describes the flow of a constant-compressibility liquid. Reported results span the full range from $N \to 0$ through $N \to \infty$.

Four different similarity solutions were derived, each having asymptotic validity within a particular range of time. The flow along the fracture becomes either fully turbulent or fully laminar in the limits as $\tau << 1$ and $\tau >> 1$, respectively. Turbulent/laminar transition occurs at roughly $\tau = 1$. Lateral seepage into the permeable fracture walls becomes either a negligible or a dominant consideration in the limits as $R\tau <<$ 1 and $R\tau >> 1$, respectively. The transition occurs at roughly $R\tau = 1$.

The propagation of a pressure disturbance along a fracture can be viewed as a sequence of asymptotic flow regimes.

1 At very early times a shock wave propagates down the fracture; frictional effects rapidly smear the wave front.

2 At early times ($\tau < 1$ and $R\tau < 1$) the flow along the fracture is turbulent and the lateral losses are negligible. The results of Fig. 4 are then applicable.

3 At intermediate times there are three possible scenarios, depending upon the magnitude of the seepage-interaction parameter, R.

a If $R<1$, the turbulent/laminar transition occurs first. The results of Fig. 3 are then applicable in the intermediate time period $1<\tau<1/R$.

b If $R>1$, the negligible-seepage/dominant-seepage transition occurs first. The results of Fig. 7 are then applicable in the intermediate time period $1/R<\tau<1$.

c If $R \simeq 1$, the transitions occur simultaneously and there is no intermediate asymptotic flow. Transition passes directly from stage (2) to stage (4).

4 At late times ($\tau > 1$ and $R\tau > 1$) the channel flow is laminar and losses are dominant. The results of Fig. 6 are then applicable.

5 At very late times ($t >> \phi_m \mu \alpha_c w^6/K_m^3$) longitudinal diffusion within the porous medium eventually becomes important, as described in [12] for the case $N \to 1$.

Transitions from one asymptotic regime to another can be computed by two-dimensional time-dependent numerical procedures, as already demonstrated for some special cases [5, 12, 16].

The reported similarity solutions are useful in several respects. First, they allow a convenient and comprehensive presentation of the results, for all values of the parameters. Secondly, computation is simpler and more reliable with two instead of three independent variables, so the similarity solutions may be viewed as benchmark solutions for verification of generalized computer codes. This computational advantage becomes more important in the multiphase (evaporating or condensing) extensions of the considered problem. Finally, a knowledge of the self-similar asymptotes is usually a sufficient basis for engineering calculations, particularly in the geologic applications where property data is only known within order-of-magnitude estimates. A number of example calculations are reported in reference [19] which concerns the containment analysis of underground nuclear tests.

Acknowledgments

This work was performed under the auspices of the US Department of Energy by the Lawrence Livermore National Laboratory under contract number W-7405-ENG-48. This work was started while the first author was at Sandia National Laboratories; we are indebted to S. K. Griffiths of that Laboratory for many helpful comments and suggestions.

References

1 Van Golf-Racht, T. D., *Fundamentals of Fractured Reservoir Engineering*, Elsevier Scientific Publishing Co., New York, 1982.

2 Veatch, R. W., Jr., "Overview of Current Hydraulic Fracturing Design and Treatment," *J. Petroleum Tech.*, Apr. 1983, pp. 677-687.

3 Matthews, C. S., and Russell, D. G., *Pressure Buildup and Flow Tests in Wells*, Soc. Pet. Eng., Dallas, 1967.

4 Gringarten, A. C., "Interpretation of Tests in Fissured and Multi-Layered Reservoirs with Double-Porosity Behavior: Theory and Practice," *J. Petroleum Tech.*, Apr. 1984, pp. 549-564.

5 Guppy, K. H., Cinco-Ley, H., and Ramey, H. J., Jr., "Effect of Non-Darcy Flow on Constant Pressure Production of Fractured Wells," *Soc. Pet. Eng. J.*, Vol. 21, No. 3, June 1981, pp. 390-400.

6 Kutter, H. K., and Fairhurst, C., "On the Fracture Process in Blasting," *Int. J. Rock Mech. Min. Sci.*, Vol. 8, 1971, pp. 181-202.

7 Geertsma, J., and Haafkens, R., " A Comparison of the Theories for Predicting Width and Extent of Vertical Hydraulically Induced Fractures," *J. Energy Res. Tech.*, Vol. 101, 1979, pp. 8-19.

8 Daneshy, A. A., "On the Design of Vertical Hydraulic Fractures," *J. Petroleum Tech.*, Jan. 1973, pp. 83-93.

9 Nilson, R. H., "Gas-Driven Fracture Propagation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 757-762.

10 Perkins, T. K., and Kern, L. R., "Widths of Hydraulic Fractures," *J. Petroleum Tech.*, Sept. 1961, pp. 937-949.

11 Nordgren, R. P., "Propagation of a Vertical Hydraulic Fracture," *Soc. Pet. Eng. J.*, Aug. 1972, pp. 306-314.

12 Nilson, R. H., "Transient Two-Dimensional Diffusion Along a High-Diffusivity Lamina Which Bisects a Half-Space," *J. Eng. Math.*, Vol. 14, No. 4, 1980, pp. 263-282.

13 Shapiro, A. H., *The Dynamics and Thermodynamics of Compressible Fluid Flow*, II, Ronald Press, New York, 1954.

14 Huitt, J. L., "Fluid Flow in Simulated Hydraulic Fractures," *AIChE Journal*, Vol. 2, No. 2, 1956, pp. 259-264.

15 Ergun, S., "Fluid Flow Through a Packed Column," *Chemical Eng. Prog.*, Vol. 48, 1952, pp. 89-94.

16 Nilson, R. H., "Transient Fluid Flow in Porous Media, Inertia-Dominated to Viscous-Dominated Transition," *J. Fluids Eng.*, Vol. 103, No. 2, 1981, pp. 339-344.

17 Morrison, F. A., Jr., "Transient Gas Flow in a Porous Column," *Ind. Eng. Chem, Fundam.*, Vol. 11, No. 2, 1972, pp. 191-197.

18 Morrison, F. A., Jr., "Similarity in Transient High Speed Gas Flow Through Porous Media," *J. Fluids Eng.*, Vol. 99, 1976, pp. 567-568.

19 Nilson, R. H., and Morrison, F. A., Jr., "Flow of Cavity Gas Along Fractures, Cable Bundles, and Grouted Cables," presented at the Third Symposium on Containment of Underground Nuclear Tests, Idaho Falls, Idaho, Sept. 1985.

APPENDIX

Comparison With Nordgren's Hydrofracture Solutions

The present results for laminar liquid flow along a hydraulic fracture driven by a constant inlet pressure can be compared with Nordgren's results [11] for the same configuration driven by a constant flow rate, Q. In the early-time lossless regime, Nordgren's solution is [7]

$$L(t) = d\left[\frac{GQ^3}{(1-\nu)\mu H^4}\right]^{1/5} t^{4/5} \qquad (A)$$

$$w_m(0,t) = e\left[\frac{(1-\nu)Q^2\mu}{GH}\right]^{1/5} t^{1/5} \qquad (A2)$$

$$\Delta P(t) = \frac{a}{H}\left[\frac{G^3 Q\mu L}{(1-\nu)^3}\right]^{1/4} \qquad (A3)$$

which can be algebraically combined to obtain

$$L = \left(\frac{d^{7/8}}{ea^{1/2}}\,\frac{4}{\pi}\right)\left[\frac{\Delta P {w_o}^2}{\mu}\right]^{1/2} t^{1/2} \qquad (A4)$$

in which $d = 0.68$, $e = 2.50$, $a = 2.75$, and $4/\pi$ is the ratio between midheight aperture, w_m, and effective aperture, w_o at the fracture inlet (see equation (1)). The corresponding expression from equation (28) of the present paper is (with $N \rightarrow 1$, $\phi_c = 1$, $\alpha_c = 1/\Delta P$, $C_l = 4\pi$, $d_o = w_o$)

$$L = \left(\frac{\theta^*}{2\pi^{1/2}}\right)\left[\frac{\Delta P {w_o}^2}{\mu}\right]^{1/2} t^{1/2} \qquad (A5)$$

in which $\theta^* \sim 0.87$ at the leading edge of the fracture (see Fig. 3). Thus, the dependence on physical parameters is identical. The only difference is that the premultiplier constant in curved brackets has the values 0.22 versus 0.25, respectively, for Nordgren versus the present.

In the late-time loss-dominated regime, Nordgren's solution is [7]

$$L = \frac{1}{\pi}\frac{Q}{HC} t^{1/2} \qquad (A6)$$

$$w_m = 4\left(\frac{2}{\pi^3}\right)^{1/4} [\mu(1-\nu)Q^2]^{1/4} t^{1/8} \qquad (A7)$$

$$\Delta P = 4\left(\frac{2}{\pi^3}\right)^{1/4}\left[\frac{\mu G^3 Q^2}{H^5(1-\nu)^3 C}\right]^{1/4} t^{1/8} \qquad (A8)$$

in which the fluid seepage velocity into the fracture walls is assumed to obey the Carter equation

$$v = C/(t-\tau)^{1/2} \qquad (A9)$$

in which $t-\tau$ is the elapsed time since the exposure of the surface. From the one-dimensional error-function solution for transient liquid flow driven by constant pressure $v = \Delta P_s (K_m \phi_m \alpha_m / \mu\pi(t-\tau))^{1/2}$, so that $C = \Delta P_s (K_m \mu \phi_m / \alpha_m \pi)^{1/2}$. Combination of the above relationships yields

$$L = \left(\frac{1}{4\pi}\right)^{3/4}\left[\frac{{w_o}^6}{K_m \phi_m \alpha_m \mu}\left(\frac{\Delta P}{\Delta P_s}\right)^2\right]^{1/4} t^{1/4} \qquad (A10)$$

which is comparable to the present equation (38) for $N \rightarrow 1$, $C_l = 4\pi$, $d_o = w_o$, $(C/A)_o = 2/w_o$

$$L = \theta^*\left(\frac{1}{8\pi}\right)^{1/2}\left[\frac{{w_o}^6}{K_m \phi_m \alpha_m \mu}\left(\frac{\Delta P}{\Delta P_s}\right)^2\right]^{1/4} t^{1/4} \qquad (A11)$$

in which $\theta^* \sim 1.06$ from Fig. 6. Again, the dependence on physical parameters is identical, but the premultipliers are now 0.15 versus 0.21, respectively, for Nordgren versus present.

It is somewhat surprising but reassuring to find reasonably good agreement (withing 10 percent and 40 percent, respectively, for the lossless and loss-dominated cases) between the solutions for constant driving pressure versus constant flow rate. At any given instant they compare fairly well, despite the difference in their prior histories of flow rate and pressure. The fracture driven by constant pressure is understandably longer since it has been previously subjected to greater flow rates and higher early-time pressures than its constant-flow counterpart. To this same level of approximation, the present constant-pressure solutions could presumably be applied to gas flows and/or turbulent flows along preexisting or hydraulically-driven fractures subjected to injection at a constant or variable flow rate.

S. L. Hendricks
Associate Professor,
Department of Engineering Science
and Mechanics,
Virginia Polytechnic Institute
and State University,
Blacksburg, Virginia 24061

Stability of a Clamped-Free Rotor Partially Filled With Liquid

When a flexible rotor is partially filled with liquid, the motion is unstable over some operating range. The extent of this operating range depends on various system parameters such as rotor damping, fluid viscosity, the amount of fluid present, etc. If the rotor is arranged so that it must tilt as it vibrates (as in the clamped-free configuration) then the tilt of the rotor greatly complicates the stability analysis. The source of the complication is that the fluid motion becomes three-dimensional. A three-dimensional stability theory is developed here and applied to a simple clamped-free rotor. The results show that the stability boundaries are influenced by both rotor and fluid "gyroscopic stiffening" effects. Brief experimental results are also reported.

Introduction

When a flexible rotor is partially filled with liquid, the motion is unstable over a certain range of operating speeds. This unstable resonance occurs when the empty rotor vibration frequency approaches the frequency of surface waves in the liquid.

This instability was noted experimentally by Kollmann (1962) and explained analytically by Kuipers (1964) and Wolfe (1968) who analyzed an undamped rotor partially filled with an inviscid liquid. Kuipers (1964) added an external damper to the rotor and was surprised by analytical results that predicted the instability of the rotor-liquid system over all spin speeds (contrary to experimental evidence [Kollman 1962 and later Wolfe 1968]). He attributed this anomaly to the linearization process and hoped that a full nonlinear analysis would alleviate the discrepancy. Hendricks (1981) proved that a damped rotor-inviscid liquid system was indeed unstable over all operating regions, however by including the viscosity of the entrapped fluid in the analysis, physically realistic results are predicted by the linear theory (Hendricks and Morton 1979). Hendricks and Morton (1979) showed how the unstable region around the system resonance was affected by the rotor damping, fluid viscosity, mass ratio, fill ratio, and rotor spin speed. A more recent paper (Hendricks and Klauber 1984) shows how an optimal control algorithm can be used to calculate a control force which will stabilize a two-mass rotor containing liquid.

All analyses mentioned so far have analyzed a rigid hollow circular cylinder which is forced to rotate around its axis of symmetry at a constant spin rate. The cylinder is mounted in the middle of an elastic shaft so that the spin axis can be displaced parallel to itself without tilting. Such a system exhibits no gyroscopic effects and does not couple with any axial motion in the fluid.

This paper will analyze a rigid hollow circular cylinder (cup) mounted on the free end of a massless clamped-free elastic shaft. The cup is partially filled with a viscous incompressible fluid and the resulting system is investigated for stability. The aim of this research is to understand how rotor gyroscopics and axial waves in the fluid affect the stability boundaries.

Rotor Dynamics

Consider a hollow circular cup (mass m, radius a, height L) mounted on the free end of a massless clamped shaft (length l). A linear damper (damping coefficient $\tilde{C}$) is attached to the bottom of the cup. The center of gravity (G) of the empty cup is located a distance d above the bottom of the cup (Fig. 1).

The equations of motion for the rotor are developed using Lagranges Equations. Define three coordinate systems: ($\hat{\mathbf{I}}$, $\hat{\mathbf{J}}$, $\hat{\mathbf{K}}$) inertially fixed; ($\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$) spinning with angular speed $\tilde{\Omega}$; ($\hat{\mathbf{d}}_1$, $\hat{\mathbf{d}}_2$, $\hat{\mathbf{d}}_3$) fixed in the cup. The following relationships exist.

$$\begin{Bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \tilde{\mathbf{k}} \end{Bmatrix} = \begin{bmatrix} \cos\tilde{\Omega}t & \sin\tilde{\Omega}t & 0 \\ -\sin\tilde{\Omega}t & \cos\tilde{\Omega}t & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} \hat{\mathbf{I}} \\ \tilde{\mathbf{J}} \\ \tilde{\mathbf{K}} \end{Bmatrix} \tag{1}$$

$$\begin{Bmatrix} \hat{\mathbf{d}}_1 \\ \hat{\mathbf{d}}_2 \\ \hat{\mathbf{d}}_3 \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_1 & \sin\phi_1 \\ 0 & -\sin\phi_1 & \cos\phi_1 \end{bmatrix} \begin{bmatrix} \cos\phi_2 & 0 & -\sin\phi_2 \\ 0 & 1 & 0 \\ \sin\phi_2 & 0 & \cos\phi_2 \end{bmatrix} \begin{Bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{Bmatrix} \tag{2}$$

The Euler Angles ϕ_1 and ϕ_2 are used to orient the cup-fixed coordinate system. The motion of the shaft is described using the two coordinates $\hat{x}(t,\tilde{z})$, $\hat{y}(t,\tilde{z})$ where $\tilde{z}$ is measured from the point B (bottom of the cup, see Fig. 1). The deflection of a point on the shaft is thus given by

$$\mathbf{R}(\tilde{z}) = \hat{x}(t,\tilde{z})\hat{\mathbf{i}} + \hat{y}(t,\tilde{z})\hat{\mathbf{j}} \tag{3}$$

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 two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 10, 1985; final revision, September 2, 1985.

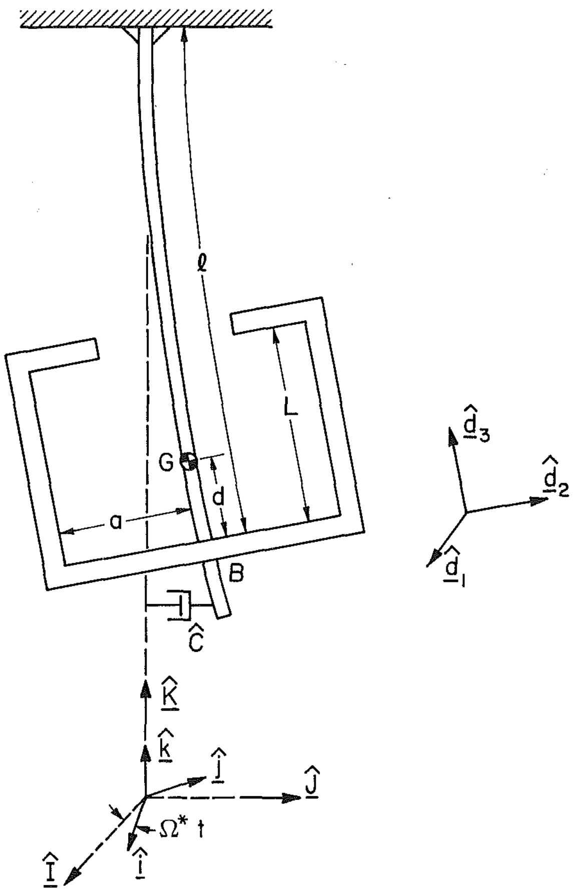


Fig. 1 Clamped-free single critical rotor and coordinate systems

The temporal and spatial variables can be separated

$$\hat{x}(t,\tilde{z})=\tilde{x}(t)\psi(\tilde{z}) \tag{4a}$$

$$\tilde{y}(t,\tilde{z})=\tilde{y}(t)\psi(\tilde{z}) \tag{4b}$$

where $\psi(\tilde{z})$ is the mode shape of the shaft (to be determined from a simple beam theory) which will be normalized so that $\psi(\tilde{z}=0)=1$ and thus the position vector that locates point B is

$$\mathbf{R}_B=\mathbf{R}(\tilde{z}=0)=\tilde{x}(t)\hat{\mathbf{i}}+\tilde{y}(t)\hat{\mathbf{j}} \tag{5}$$

The cup can be oriented by examining the unit vector $\hat{\mathbf{d}}_3$ which can be found from equation (3) by differentiating.

$$\hat{\mathbf{d}}_3\approx\frac{\partial[\mathbf{R}(\tilde{z})]}{\partial\tilde{z}}\bigg|_{\tilde{z}=0}=\tilde{\epsilon}\tilde{x}\hat{\mathbf{i}}+\tilde{\epsilon}\tilde{y}\hat{\mathbf{j}}+\hat{\mathbf{k}} \tag{6}$$

The parameter $\tilde{\epsilon}=d\psi(\tilde{z})/d\tilde{z}\;|_{\tilde{z}=0}$ is a slope constant that measures how much the cup tilts as it runs out. Since this paper specifically examines the effect of tilt on the dynamics of the system, $\tilde{\epsilon}$ is a key parameter. From equation (2)

$$\hat{\mathbf{d}}_3=\sin\phi_2\cos\phi_1\hat{\mathbf{i}}-\sin\phi_1\hat{\mathbf{j}}+\cos\phi_1\cos\phi_2\hat{\mathbf{k}} \tag{7}$$

For small angles

$$\hat{\mathbf{d}}_3=\phi_2\hat{\mathbf{i}}-\phi_1\hat{\mathbf{j}}+\hat{\mathbf{k}} \tag{8}$$

and the Euler angles are identified as

$$\phi_1=\tilde{\epsilon}\tilde{x} \tag{9a}$$

$$\phi_2=\tilde{\epsilon}\tilde{y} \tag{9b}$$

The angular velocity of the cup is

$$\omega=\tilde{\Omega}\hat{\mathbf{k}}+\dot{\phi}_2\hat{\mathbf{j}}+\dot{\phi}_1\hat{\mathbf{d}}_1 \tag{10}$$

Keeping terms to second order the angular velocity can be written in the cup fixed coordinate system as

$$\omega=-\tilde{\epsilon}(\dot{\tilde{y}}+\tilde{\Omega}\tilde{x})\hat{\mathbf{d}}_1+\tilde{\epsilon}(\dot{\tilde{x}}-\tilde{\Omega}\tilde{y})\hat{\mathbf{d}}_2$$

$$+\left[\tilde{\Omega}\left(1-\frac{\tilde{\epsilon}^2}{2}[\tilde{x}^2+\tilde{y}^2]\right)+\tilde{\epsilon}^2\dot{\tilde{x}}\tilde{y}\right]\hat{\mathbf{d}}_3 \tag{11}$$

The position and velocity of the center of mass (point G, Fig. 1) are

$$\mathbf{R}_G=\tilde{x}\hat{\mathbf{i}}+\tilde{y}\hat{\mathbf{j}}+d\hat{\mathbf{d}}_3\approx\tilde{x}(1+\tilde{\epsilon}d)\hat{\mathbf{i}}+\tilde{y}(1+\tilde{\epsilon}d)\hat{\mathbf{j}}+d\hat{\mathbf{k}} \tag{12a}$$

$$\dot{\mathbf{R}}_G=[(\dot{\tilde{x}}-\tilde{\Omega}\tilde{y})\hat{\mathbf{i}}+(\dot{\tilde{y}}+\tilde{\Omega}\tilde{x})\hat{\mathbf{j}}](1+\tilde{\epsilon}d) \tag{12b}$$

For future reference the velocity and acceleration of point B are now listed.

$$\dot{\mathbf{R}}_B=(\dot{\tilde{x}}-\tilde{\Omega}y)\hat{\mathbf{i}}+(\dot{\tilde{y}}+\tilde{\Omega}\tilde{x})\hat{\mathbf{j}} \tag{13a}$$

$$\ddot{\mathbf{R}}_B=(\ddot{\tilde{x}}-2\tilde{\Omega}\dot{\tilde{y}}-\tilde{\Omega}^2\tilde{x})\hat{\mathbf{i}}+(\ddot{\tilde{y}}+2\tilde{\Omega}\dot{\tilde{x}}-\tilde{\Omega}^2\tilde{y})\hat{\mathbf{j}} \tag{13b}$$

The kinetic energy, dissipation function, and potential energy are

$$T=\frac{1}{2}\,m\,\dot{\mathbf{R}}_G\bullet\dot{\mathbf{R}}_G+\frac{1}{2}\,\omega\bullet\underline{\underline{\mathrm{I}}}\bullet\omega \tag{14a}$$

$$D=\frac{1}{2}\,\tilde{C}\dot{\mathbf{R}}_B\bullet\dot{\mathbf{R}}_B \tag{14b}$$

$$V=\frac{1}{2}\,k\mathbf{R}_B\bullet\mathbf{R}_B \tag{14c}$$

where $\underline{\underline{\mathrm{I}}}$ is the inertia dyadic of the cup about the center of gravity and k is the spring constant which results from the bending of the shaft. The unit vectors $(\hat{\mathbf{d}}_1,\hat{\mathbf{d}}_2,\hat{\mathbf{d}}_3)$ are oriented along the principal axis of the cup, and the cup is symmetrical about $\hat{\mathbf{d}}_3$. The inertial dyadic can be written in matrix notation

$$\underline{\underline{\mathrm{I}}}=\begin{bmatrix}\tilde{I}&0&0\\0&\tilde{I}&0\\0&0&\tilde{J}\end{bmatrix} \tag{15}$$

The rotor equations of motion are

$$\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{\tilde{x}}}\right)-\frac{\partial T}{\partial\tilde{x}}+\frac{\partial D}{\partial\dot{\tilde{x}}}+\frac{\partial V}{\partial\tilde{x}}=Q_x \tag{16a}$$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{\tilde{y}}}\right)-\frac{\partial T}{\partial\tilde{y}}+\frac{\partial D}{\partial\dot{\tilde{y}}}+\frac{\partial V}{\partial\tilde{y}}=Q_y \tag{16b}$$

where Q_x and Q_y are the forces on the rotor due to the fluid. These forces will be developed shortly, after the empty rotor equations (i.e., $Q_x=Q_y=0$) are examined. In order to highlight the important terms, the following nondimensional quantities are introduced.

$$\Omega=\tilde{\Omega}/\omega_o \tag{17a}$$

$$C=\tilde{C}/2M_t\omega_o \tag{17b}$$

$$J=\tilde{J}\tilde{\epsilon}^2/M_t \tag{17c}$$

where

$$M_t=M(1+\tilde{\epsilon}d)^2+\tilde{\mathbf{I}}\tilde{\epsilon}^2 \tag{18a}$$

$$\omega_o=(k/M_t)^{1/2} \tag{18b}$$

the nondimensional equivalents of equations (16*a, 16b*) for the empty rotor are

$$\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{Bmatrix}\ddot{x}\\\ddot{y}\end{Bmatrix}+\begin{bmatrix}2C&(J-2)\Omega\\(2-J)\Omega&2C\end{bmatrix}\begin{Bmatrix}\dot{x}\\\dot{y}\end{Bmatrix}$$

$$+\begin{bmatrix}1+(J-1)\Omega^2&-2C\Omega\\2C\Omega&1+(J-1)\Omega^2\end{bmatrix}\begin{Bmatrix}x\\y\end{Bmatrix}=0 \tag{19}$$

where $x=\tilde{x}/a$ and $y=\tilde{y}/a$ are the nondimensional rotor variables.

The fact that the cup tilts as it vibrates is incorporated in the nondimensional term J which measures the polar moment of inertia of the cup weighted by the slope parameter . To see how J affects the undamped empty rotor, the eigenvalue problem associated with equation (19) (setting $C = 0$) is formulated. This leads to the following characteristic equation for the vibration frequency S.

$$S^2+(2-J)\Omega S-[1+(J-1)\Omega^2]=0 \tag{20}$$

Equation (20) can be solved for the two vibration frequencies (given in the rotating reference frame)

$$S=-\Omega+\frac{J\Omega}{2}\pm\left[1+\left(\frac{J\Omega}{2}\right)^2\right]^{1/2} \tag{21}$$

or if the gyroscopic term (J) is small (the usual case)

$$S\approx-\Omega\pm 1+\frac{J\Omega}{2}+\frac{1}{2}\left(\frac{J\Omega}{2}\right)^2 \tag{22}$$

The first term in equation (22) arises because we are working in a rotating reference frame, the next term represents the natural frequency of the system without the gyroscopic term (i.e., $J = 0$), and the last two terms represents the gyroscopic "stiffening". These last terms are present because the spinning cup acts as a gyroscope in that the spinning system resists any change in its angular momentum. This resistance to tilt manifests itself by stiffening the effective spring constant and thereby raising the natural frequency of the system. As we shall see later this gyroscopic stiffening has a predictable influence on the liquid-rotor stability problem.

When the rotor contains a fluid, Q_x and Q_y are not zero since the fluid pushes on the rotor wall. Here we examine the effect that an entrapped inviscid fluid has on the motion of the cup. Later we will need to add the viscosity of the fluid in order to complete the stability analysis. For a thin fluid layer, the forces on the top and bottom of the cup may be neglected and only those on the side wall will be considered. Thus Q_x, Q_y are given by

$$Q_x=\int_0^L\int_0^{2\pi}\frac{\partial}{\partial\tilde{x}}\,\mathbf{R}_w\cdot\tilde{P}(\tilde{r}=a,\theta,\tilde{z},t)\hat{\mathbf{r}}a d\theta d\tilde{z} \tag{23a}$$

$$Q_y=\int_0^L\int_0^{2\pi}\frac{\partial}{\partial\tilde{y}}\,\mathbf{R}_w\cdot\tilde{P}(\tilde{r}=a,\theta,\tilde{z},t)\hat{\mathbf{r}}a d\theta d\tilde{z} \tag{23b}$$

where $\tilde{P}(\tilde{r}, \theta, \tilde{z}, t)$ is the pressure in the fluid,

$$\mathbf{R}_w=\tilde{x}(t)\hat{\mathbf{i}}+\tilde{y}(t)\hat{\mathbf{j}}+\tilde{z}\tilde{\mathbf{d}}_3+a\hat{\mathbf{r}} \tag{24}$$

is a vector which locates a point on the wall, and

$$\hat{\mathbf{r}}=\cos\theta\hat{\mathbf{d}}_1+\sin\theta\hat{\mathbf{d}}_2 \tag{25}$$

is a unit vector pointing in the radial direction.

Equations (23a), (23b) represent the net force on the rotor due to the pressure (a normal distributed force) integrated over the surface of the rotor. The formulation of the governing equations is now completed by developing the fluid equations.

Fluid Equations

The governing equations for the fluid are the equations which conserve mass (the continuity equation) and momentum (the Navier-Stokes equations). Written in a reference frame attached to and moving with the cup ($\hat{\mathbf{d}}_1$, $\hat{\mathbf{d}}_2$, $\hat{\mathbf{d}}_3$), these equations are:

$$\nabla\cdot\mathbf{V}=0 \tag{26a}$$

$$\dot{\mathbf{V}}+(\mathbf{V}\cdot\nabla)\mathbf{V}+\ddot{\mathbf{R}}_B+\alpha\; x\;\mathbf{r}+\omega\; x(\omega\; x\;\mathbf{r})$$
$$+2\omega\; x\;\mathbf{V}+\frac{1}{\rho}\nabla\tilde{\mathbf{P}}=0 \tag{26b}$$

where

$\mathbf{V}$ is the fluid velocity vector
$\tilde{P}$ is the fluid pressure
ρ is the fluid density

and to first order

$$\omega\approx-\tilde{\epsilon}(\dot{\tilde{y}}+\tilde{\Omega}\tilde{x})\hat{\mathbf{d}}_1+\tilde{\epsilon}(\dot{\tilde{x}}-\tilde{\Omega}\tilde{y})\hat{\mathbf{d}}_2 \tag{27a}$$

$$\alpha\approx-\tilde{\epsilon}(\ddot{\tilde{y}}+\tilde{\Omega}\dot{\tilde{x}})\hat{\mathbf{d}}_1+\tilde{\epsilon}(\ddot{\tilde{x}}-\tilde{\Omega}\dot{\tilde{y}})\hat{\mathbf{d}}_2 \tag{27b}$$

$$\mathbf{r}=\tilde{r}\hat{\mathbf{r}} \tag{27c}$$

Equation (26b) is just the usual momentum equations with four extra acceleration terms due to the motion of the coordinate system (attached to the cup). To linearize these equations let

$$\tilde{P}=\frac{1}{2}\rho\tilde{\Omega}^2(\tilde{r}^2-b^2)+P(M_t\omega_o^2/a) \tag{28}$$

where b is the radius of the nominal free surface. The first term on the right hand side of equation (28) represents the pressure due to the solid-body rotation of the fluid and the second is a small perturbation due to any waves that develop. P is the nondimensional pressure perturbation. The fluid velocity can be written as

$$\mathbf{V}=(u\mathbf{r}+v\hat{\theta}+w\hat{\mathbf{z}})a\omega_o \tag{29}$$

where (u, v, w) is the nondimensional velocity in the (radial, azimuthal, axial) direction,

$$\hat{\theta}=-\sin\theta\hat{\mathbf{d}}_1+\cos\theta\hat{\mathbf{d}}_2 \tag{30}$$

is a unit vector in the azimuthal direction, and

$$\hat{\mathbf{z}}=\hat{\mathbf{d}}_3 \tag{31}$$

is a unit vector in the axial direction.

The following additional nondimensional terms and variables are now introduced:

$f=b/a$, a fill ratio that measures how much fluid is present
$\mu=\pi\rho a^2L/M_t$, a mass ratio
$z_o=L/a$, an aspect ratio
$\epsilon=\tilde{\epsilon}a$, nondimensional tilt variable
$r=\tilde{r}/a$, nondimensional radial distance
$z=\tilde{z}/a$, nondimensional axial distance

Equations (26a), (26b) written in nondimensional, linearized, scalar form are:

$$\frac{u}{r}+\frac{\partial u}{\partial r}+\frac{1}{r}\frac{\partial v}{\partial\theta}+\frac{\partial w}{\partial z}=0 \tag{32a}$$

$$\dot{u}-2\Omega v+\frac{\pi z_o}{\mu}\frac{\partial P}{\partial r}+\mathbf{A}\cdot\hat{\mathbf{r}}=0 \tag{32b}$$

$$\dot{v}+2\Omega u+\frac{\pi z_o}{\mu r}\frac{\partial P}{\partial\theta}+\mathbf{A}\cdot\hat{\theta}=0 \tag{32c}$$

$$\dot{w}+\frac{\pi z_o}{\mu}\frac{\partial P}{\partial z}+\mathbf{A}\cdot\hat{\mathbf{z}}=0 \tag{32d}$$

where the rotor motion is incorporated through the term

$$\mathbf{A}=(1+\epsilon z)[(\ddot{x}-2\Omega\dot{y}-\Omega^2x)\cos\theta+(\ddot{y}+2\Omega\dot{x}-\Omega^2y)\sin\theta]\hat{\mathbf{r}}$$
$$+(1+\epsilon z)[(\ddot{y}+2\Omega\dot{x}-\Omega^2y)\sin\theta-(\ddot{x}-2\Omega\dot{y}-\Omega^2x)\cos\theta]\hat{\theta} \tag{33}$$
$$-\epsilon r[(\ddot{y}+\Omega^2y)\sin\theta+(\ddot{x}+\Omega^2x)\cos\theta]\hat{\mathbf{z}}$$

It is now easy to see why tilt in the rotor greatly complicates the analysis. If we do not allow the rotor to tilt (i.e., set $\epsilon = 0$) then the last term in equation (32d) is zero and one solution is that the axial velocity (w) is identically zero and the pressure does not depend on the axial coordinate z. The flow then becomes two dimensional (all fluid variables depend only on r and θ but not z). If the rotor tilts ($\epsilon \neq 0$) then the acceleration of the rotor drives the axial velocity through the last term in equation (32d). The rotor acceleration terms in the radial and azimuthal momentum equations (32 b,c) also introduce variations in the axial direction (z) for the case $\epsilon \neq 0$. The flow is no longer two-dimensional but must be analyzed as a three-dimensional flow (i.e., all fluid variables depend on r, θ, and z).

To complete the fluid analysis the boundary conditions must be specified. These are:

1. The fluid velocity must be zero normal to all solid surfaces.
2. The total pressure at the free surface must be zero and the surface radial velocity must match the fluid radial velocity.

The first condition requires that

$$u(r=1,\theta,z,t)=0 \tag{34a}$$

$$w(r,\theta z=0,t)=0 \tag{34b}$$

$$w(r,\theta,z=z_o,t)=0 \tag{34c}$$

The second boundary condition needs to be massaged into a useful form. The nondimensional pressure is (compare equation (28))

$$\text{pressure}=\frac{1}{2}\,\frac{\mu\Omega^2}{\pi z_o}(r^2-f^2)+P \tag{35}$$

The free surface of the fluid wave is located at

$$r=f+\lambda(\theta,z,t) \tag{36}$$

where λ represents the perturbation from the nominal free surface. Using equation (36) in equation (35), dropping the nonlinear term, and setting $u=\dot{\lambda}$, the second boundary condition becomes

$$u(r=f,\theta,z,t)=-\frac{\pi z_o}{\mu\Omega^2 f}\,\dot{P}(r=f,\theta,z,t) \tag{37}$$

The system is now completely described mathematically; on to the solution.

Solution

The dynamics of the liquid-rotor system are incorporated in the rotor equations (16*a*, 16*b*) augmented with the fluid forces (23*a*, 23*b*) together with the fluid equations (32*a*–32*d*) and boundary conditions (34*a*–34*c*, 37). The rotor coordinates (x,y) depend only on time while the fluid variables (u,v,w,P) depend on the three spatial variables (r,θ,z) as well as time. To solve this set of mixed equations a solution of the following form is assumed.

$$x(t)=xe^{-iS\omega_o t} \tag{38a}$$

$$y(t)=ye^{-iS\omega_o t} \tag{38b}$$

$$u(r,\theta,z,t)=\sum_{m=0}^{M}u_m(r)\cos\left(\frac{m\pi z}{z_o}\right)e^{i(\theta-S\omega_o t)} \tag{38c}$$

$$v(r,\theta,z,t)=\sum_{m=0}^{M}v_m(r)\cos\left(\frac{m\pi z}{z_o}\right)e^{i/\theta-S\omega_o t)} \tag{38d}$$

$$w(r,\theta,z,t)=\sum_{m=0}^{M}w_m(r)\sin\left(\frac{m\pi z}{z_o}\right)e^{i(\theta-S\omega_o t)} \tag{38e}$$

$$P(r,\theta,z,t)=\sum_{m=0}^{M}P_m(r)\cos\left(\frac{m\pi z}{z_o}\right)e^{i(\theta-S\omega_o t)} \tag{38f}$$

All variables are given the same time dependence (parameter S). It is the object of the solution process to determine S. By expanding the axial dependence of the axial velocity (w) in sine terms we have implicitly satisfied the boundary conditions at the top and bottom of the cup. The only azimuthal (θ) mode that is considered is the mode that has exactly one node. All other azimuthal modes produce no net force on the rotor and can be left out of the analysis without any loss. The axial expansion will necessarily be truncated at some finite number of terms introducing an approximation into the solution process; however, if enough terms are included in the series, the solution will retain sufficient accuracy. With expansions (38*a*–38*f*), the governing equations can be written in the following nondimensional form.

$$\begin{bmatrix} 1+(J-1)\Omega^2-S^2-2iCS & -2C\Omega-iS\Omega(J-2) \\ 2C\Omega+iS\Omega(J-2) & 1+(J-1)\Omega^2-S^2-2iCS \end{bmatrix}\begin{Bmatrix} x \\ y \end{Bmatrix}$$

$$=\pi z_o\sum_{n=0}^{M}P_n(r=1)Q_n\quad\begin{Bmatrix} 1 \\ i \end{Bmatrix} \tag{39a}$$

$$\frac{u_m+iv_m}{r}+\frac{d}{dr}\,u_m+\frac{m\pi}{z_o}\,w_m=0 \tag{39b}$$

$$-iSu_m-2\Omega v_m+\frac{\pi z_o}{\mu}\,\frac{d}{dr}\,P_m+(S+\Omega)^2(x-iy)D_m=0 \tag{39c}$$

$$-iSv_m+2\Omega u_m+\frac{i\pi z_o}{\mu r}\,P_m+i(S+\Omega)^2(x-iy)D_m=0 \tag{39d}$$

$$-iSw_m-\frac{m\pi^2}{\mu}\,P_m+r(S^2-\Omega^2)(x-iy)E_m=0 \tag{39e}$$

where

$$Q_n=\begin{cases} 1+\epsilon z_o/2 & n=0 \\ -2\epsilon z_o/(n\pi)^2 & n\text{ odd} \\ 0 & n\text{ even}\neq 0 \end{cases} \tag{40a}$$

$$D_m=\begin{cases} -\frac{1}{2}(1+\epsilon z_o/2) & m=0 \\ 2\epsilon z_o/(m\pi)^2 & m\text{ odd} \\ 0 & m\text{ even}\neq 0 \end{cases} \tag{40b}$$

$$E_m=\begin{cases} -2\epsilon/m\pi & m\text{ odd} \\ 0 & m\text{ even} \end{cases} \tag{40c}$$

The boundary conditions are

$$u_m(r=1)=0 \tag{41a}$$

$$u_m(r=f)=\frac{i\pi z_o S}{\mu\Omega^2 f}\,P_m(r=f) \tag{41b}$$

Equations (39*a*–41*b*) must now be manipulated to yield an eigenvalue problem for the parameter S (which determines the frequency and stability of the motion). To accomplish this task x, y, u_m, v_m, and w_m are eliminated from equations (39*a*–*e*). After much algebra the result is a set of $M+1$ second order ordinary differential equations for P_m.

$$\frac{d^2P_m}{dr^2}+\frac{1}{r}\,\frac{dP_m}{dr}+\left(m\kappa^2-\frac{1}{r^2}\right)P_m=m^2\kappa^2 r\sum_{n=0}^{M}\beta_{mn}P_n(r=1) \tag{42}$$

where

$$\kappa=\frac{\pi}{z_o S}(4\Omega^2-S^2)^{1/2} \tag{43}$$

and β_{mn} represents an array containing rotor and kinematic information. This array is defined in the appendix.

Equation (42) is just a set of Bessel's equations which are coupled by the driving term that depends on the pressure at the rotor wall. The solution to equation (42) is

$$
P_m(r) = \begin{cases} A_o r + B_o \dfrac{1}{r} & m = 0 \\ A_m H_1^{(1)}(m\kappa r) + B_m H_1^{(2)}(m\kappa r) & \\ \quad + r \displaystyle\sum_{n=0}^{M} \beta_{mn} P_n(1) & m \neq 0 \end{cases} \tag{44}
$$

where $H_1^{(1)}$ and $H_1^{(2)}$ represent Hankel functions of the first and second kind of order 1. Hankel functions must be used for the calculation instead of the more common Bessel functions since the latter are not numerically independent in the range of interest. Applying the boundary conditions leads to an equation of the form

$$[F]\{Y\} = 0 \tag{45}$$

where $\{Y\}$ is a $2M + 2$ vector that contains the coefficients A_i and B_i and $[F]$ is a matrix (defined in the appendix) that contains all of the system dynamics. Embedded in the matrix $[F]$ is the parameter S. In order to have a nontrivial solution of equation (45) we must have

$$\det[F] = 0 \tag{46}$$

Equation (46) is the characteristic equation which determines S. Using a root solving technique one can now search for a value of S which makes the determinant zero. This will be one of the $2M + 2$ system eigenvalues. The real part of S represents one of the "natural" frequencies of the rotor-liquid system. If all of the system eigenvalues have negative imaginary parts, the system will be stable. If any of the eigenvalues has a positive imaginary part, the system will be unstable. Once an eigenvalue has been found, the corresponding eigenvector of $[F]$ can be used to determine the mode shape of the system.

Stability Analysis

The inviscid analysis developed in the preceding sections can be used to predict frequencies and mode shapes of the system. Unfortunately the analysis presented so far is not complete enough to yield stability information. The inviscid theory predicts the rotor to be unstable at any spin speed. In order to complete the stability analysis it is necessary to include the viscosity of the entrapped fluid. The viscous theory has been completed (Hendricks 1979) but the analysis is only briefly sketched here due to the complexity of the equations.

If the fluid is viscous then in addition to a normal force (pressure) on the rotor wall there will be shear forces (drag). Thus the first update to the preceeding analysis will be the addition of axial and azimuthal drag terms in equations (23*a*, 23*b*). In the fluid equations it will be necessary to add the usual viscous diffusion term to the Navier-Stokes equation (26*b*). This has the effect of raising the order of the differential equations and significantly complicating the solution. The fluid boundary conditions must now include the no-slip condition at the solid surfaces.

The viscous problem may be solved by using a technique common when working with rotating fluids (Greenspan 1965). The procedure involves an asymptotic expansion in terms of a suitable Reynolds' number. In this problem the natural Reynold's number is

$$\mathrm{Re} = a^2 \omega_o / \nu$$

where ν is the kinematic viscosity of the fluid. The problem then separates into three parts: an inviscid core; a boundary layer; and a viscous correction. The inviscid core is just the analysis presented in the earlier part of this paper. The solution to the inviscid problem provides input for the boundary layer analysis. The solution to the boundary layer is used to calculate a "viscous correction". The complete solution is then accurate to $0(\mathrm{Re}^{-1/2})$ which is sufficient to determine stability. The analysis ignores viscous dissipation at the top and bottom of the cup (Ekman layers) and thereby is valid only for thin liquid layers.

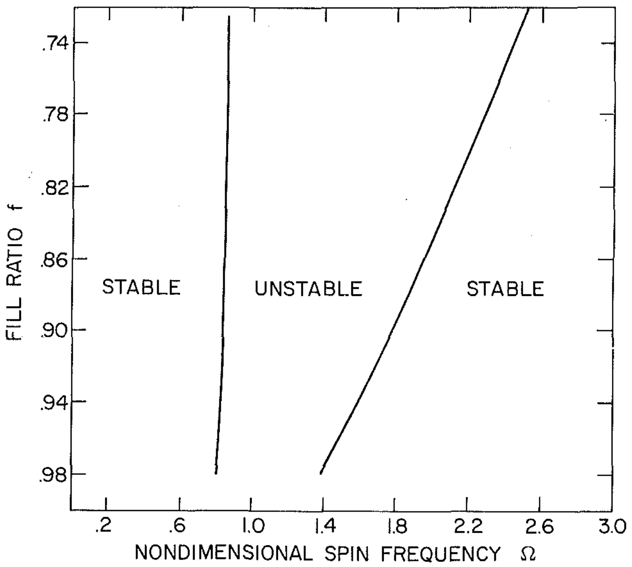


Fig. 2 Typical stability map ($\mu = 0.63$, $C = 0.01$, $\mathrm{Re} = 2.0 \times 10^5$, $z_o = 1.68$, $\epsilon = 0$, $J = 0$)

Results and Experiment

The system depends on the following eight nondimensional parameters

Ω = rotor spin speed

f = fill ratio ($f = 0$ corresponds to a full rotor, $f = 1$ corresponds to an empty rotor)

μ = mass ratio that measures the fluid mass density

C = rotor damping constant ($C = 1$ corresponds to critical damping)

Re = Reynolds number that measures fluid viscosity

z_o = aspect ratio that measure the length of the cup

ϵ = tilt parameter that measures how much the rotor tilts as it vibrates

J = polar moment of inertia of the cup weighted by the tilt parameter

A typical stability calculation consists of defining these eight parameters for the particular system of interest, choosing a starting guess for S, then iterating until equation (46) is satisfied. The viscous correction is then calculated and the complete eigenvalue is checked to see if the eigenvalue has a negative imaginary part. This process is repeated until all eigenvalues have been found. If all of the eigenvalues have negative imaginary parts, the system is stable; otherwise the system is unstable. The results of a stability calculation are conveniently shown on a stabilty map (stability boundaries plotted on the Ω, f plane). Figure 2 shows a typical stability map. For a given amount of fluid (f), the system is stable at low spin speeds. As the spin speed (Ω) increases, the system encounters an unstable region. At higher spin speeds the system is again stable. As more liquid is added, the region of instability expands. Since the analysis is based on a boundary layer theory, it cannot be used to predict stability boundaries if the fluid layer is thinner than the boundary thickness. A correct analysis for a very thin layer would require different boundary conditions than those used here. For this reason the stability

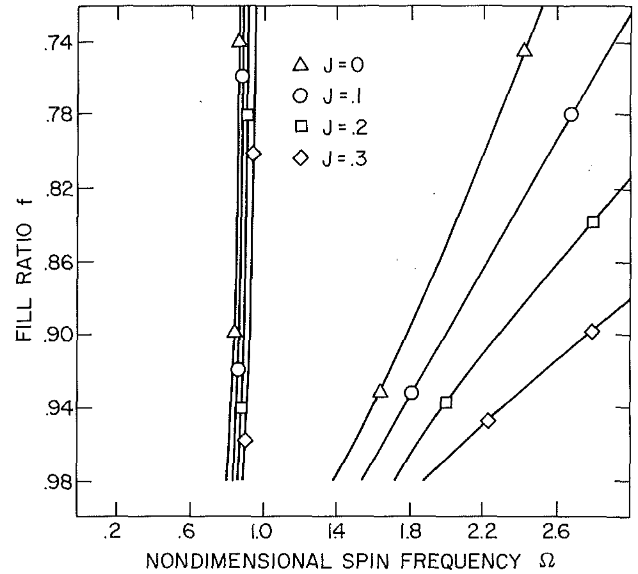


Fig. 3 Sensitivity to rotor gyroscopic constant J (two-dimensional theory, $\mu = 0.63$, $C = 0.01$, $Re = 2.0 \times 10^5$)

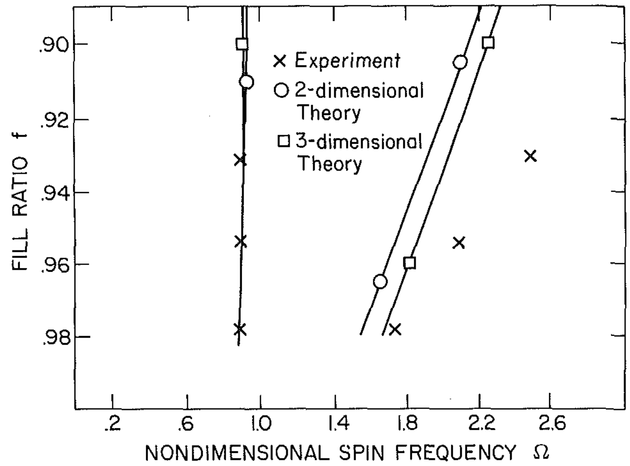


Fig. 4 Stability map ($\mu = 1.1$, $C = 0.013$, $Re = 1.6 \times 10^5$, $J = 0.15$)

boundaries of Fig. 2 cannot be extended directly to the limit $f = 1$.

The effect of the first five parameters on the stability map is reported in Hendricks and Morton (1979). The purpose of this paper is to discover the effect that rotor tilt has on the stability boundaries. The tilt of the rotor is manifested in two ways: first the polar moment of inertia (J) becomes important in the rotor equations (we have already seen how this parameter gyroscopically stiffens the rotor); second the fluid now is driven axially by the motion of the rotor. To separate these two effects, stability maps were generated using only the first term in each of the axial fluid expansions. This has the artificial effect of restricting the fluid motion to two dimensions allowing us to separately examine the effect of the rotor gyroscopic term. Figure 3 shows the effect that changing J has on the stability map. As J increases, the lower stability boundary is slightly raised while the upper stability boundary shows a much greater increase. This is a reasonable result since the gryroscopic "stiffening" effect is more pronounced at higher spin speeds.

Figure 4 shows stability maps using first the two-dimensional theory and then the three-dimensional theory (with ten terms in each axial fluid expansion). The three-dimensional theory allows the fluid to move axially. The axial motion of the fluid has an effect similar to the rotor gyroscopic stiffening. The lower stability boundary is slightly raised while the upper stability boundary shows a much greater increase. Figure 5 is a plot of the fluid mode shape which corresponds to the unstable eigenvalue in the three-dimensional calculation. Note that the fluid has aligned itself to create a restoring torque that resists rotor runout. The fluid thus adds a gyroscopic stiffening effect in addition to the effect produced by the polar moment of inertia of the cup.

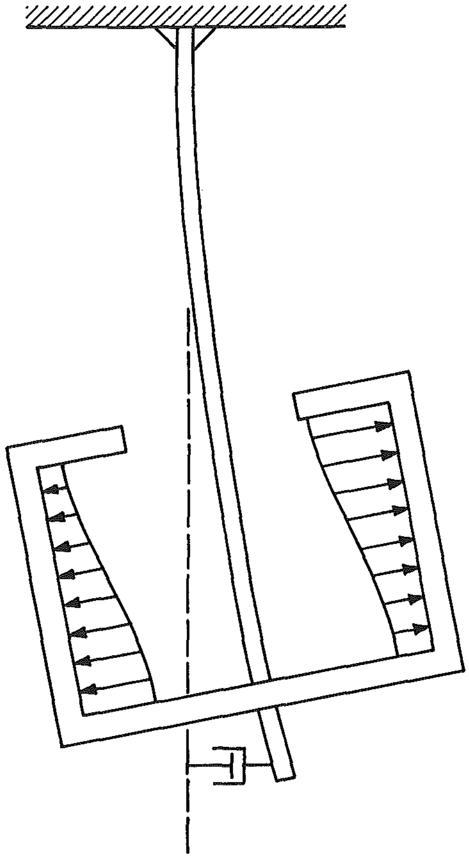


Fig. 5 Fluid pressure distribution

Also shown in Fig. 4 are some experimental points obtained by Dr. M. A. Shaddy, Jr. in an experiment described in Hendricks (1979). The lower stability boundaries show good agreement with the experiment. The three-dimensional upper stability boundary agrees with experiment for thin fluid layers, however, the slope of the upper stability boundaries are different. There are two possible explanations for the discrepancy in the slope. First the simple damper used in the experiment may be providing more damping at higher spin speeds. A more plausible explanation is that the Ekman boundary layers at the top and bottom of the cup may significantly add damping to the system as more fluid is added to this relatively short experimental rotor.

Conclusions

When a viscous liquid is introduced into a flexible rotor there are regions of unstable motion. If the rotor motion is such that the rotor must tilt as it vibrates then there are two gyroscopic stiffening effects that raise both the lower and upper stability boundaries. The first effect is just the usual stiffening due to the polar moment of inertia of the rotor. The second effect arises because the fluid moves so as to create a torque opposing the tilt of the rotor. Both of these effects act to stiffen the rotor and raise the effective natural frequency of the system. The three dimensional theory presented here is necessary in order to accurately predict stability boundaries for a clampled-free rotor (or any rotor where tilt is important).

References

Greenspan, H. P., 1965, "On the General Theory of Contained Rotating Fluid Motions," *Journal of Fluid Mechanics*, Vol. 22, p. 449.

Hendricks, S. L., 1979, "Dynamic of Flexible Rotors Partially Filled With a Viscous Incompressible Fluid," Ph.D. Dissertation, University of Virginia.

Hendricks, S. L., and Morton, J. B., 1979, "Stability of a Rotor Partially Filled With a Viscous Incompressible Fluid," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 913-918.

Hendricks, S. L., 1981, "Instability of a Damped Rotor Partially Filled With an Inviscid Liquid," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, p. 674.

Hendricks, S. L., and Klauber, R. D., 1984, "Optimal Control of a Rotor Partially Filled with an Inviscid Incompressible Fluid," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 863-868.

Kollman, F. G., 1962, Experimentelle und theoretische Untersuchungen uber die Kritischen Drehzahlen flussigkeitsegefulter Hohlkorper," *Forschund auf dem Gebiete des Ingenieurwesens*, Ausgabe B, Vol. 28, pp. 115-123 and 147-153.

Kuipers, M., 1964, "On the Stability of a Flexible Mounted Rotating Cylinder Partially Filled with Liquid," *Applied Scientific Research*, Section A, Vol. 13, pp. 121-137.

Wolfe, J. A., Jr., 1968, "Whirl Dynamics of a Rotor Partially Filled With Liquid," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 676-682.

APPENDIX

The characteristics equation which determines S for the three-dimensional analysis is equation (46). Written out in more detail, this equation is

$$m=1,M\left\{\vphantom{\begin{matrix}a\\a\end{matrix}}\right.\begin{bmatrix} F_{ooa1} & F_{oob1} & F_{ola1} & F_{olb1} \\ F_{ooa2} & F_{oob2} & F_{ola2} & F_{olb2} \\ F_{moa1} & F_{mob1} & F_{mla1} & F_{mlb1} \\ F_{moa2} & F_{mob2} & \underbrace{F_{mla2} \quad F_{mlb2}}_{l=1,M} \end{bmatrix} \begin{Bmatrix} A_o \\ B_o \\ A_l \\ B_l \end{Bmatrix} = 0 \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} l=1,M$$

The following ($[M+1] \times [M+1]$) matrices are defined

$$\alpha_{mn} = \mu(S+\Omega)^2 D_m Q_n / [1-(S+\Omega)^2 - (2iC-J)(S+\Omega)] \quad m,n=(0,M)$$

$$\beta_{mn} = \mu(S^2-\Omega^2) m\pi E_m Q_n / [1-(S+\Omega)^2 - (2iC-J)(S+\Omega)] \quad m,n=(0,M)$$

where Q_n, D_m, and E_m are given in equations (40*a-c*). The following ($M \times M$) matrix is also needed

$$[C_{mn}] = [\delta_{mn} - \beta_{mn}]^{-1} m,n=(1,M)$$

where δ_{mn} is the usual Kronecker delta.

The elements of the $[F]$ matrix are now constructed according to the following definitions:

$$F_{ooa1} = (2\Omega-S)\left(1+\alpha_{oo}+\sum_{n=1}^{M}\sum_{l=1}^{M}\alpha_{on}C_{nl}\beta_{lo}\right)$$

$$F_{oob1} = (2\Omega+S)+(2\Omega-S)\left(\alpha_{oo}+\sum_{n=1}^{M}\sum_{l=1}^{M}\alpha_{on}C_{nl}\beta_{lo}\right)$$

$$F_{ola1} = (2\Omega-S)\sum_{n=1}^{M}\alpha_{on}C_{nl}H_1^{(1)}(l\kappa)$$

$$F_{olb1} = (2\Omega-S)\sum_{n=1}^{M}\alpha_{on}C_{nl}H_1^{(2)}(l\kappa)$$

$$F_{ooa2} = (2\Omega-S)\left(\left[\frac{S+\Omega}{\Omega}\right]^2+\alpha_{oo}+\sum_{n=1}^{M}\sum_{l=1}^{M}\alpha_{on}C_{nl}\beta_{lo}\right)$$

$$F_{oob2} = (2\Omega+S)\left(\frac{\Omega^2+2\Omega S-S^2}{\Omega^2 f^2}\right)+(2\Omega-S)\left(\alpha_{oo}+\sum_{n=1}^{M}\sum_{l=1}^{M}\alpha_{on}C_{nl}\beta_{lo}\right)$$

$$F_{ola2} = F_{ola1}$$

$$F_{olb2} = F_{olb1}$$

$$F_{moa1} = (2\Omega-S)\left(\alpha_{mo}+\beta_{mo}+\sum_{n=1}^{M}\sum_{l=1}^{M}[\alpha_{mn}+\beta_{mn}]C_{nl}\beta_{lo}\right)$$

$$F_{mob1} = F_{moa1}$$

$$F_{mla1} = [(2\Omega+S)H_1^{(1)}(m\kappa)-m\kappa S H_o^{(1)}(m\kappa)]\delta_{ml} + (2\Omega-S)\sum_{n=1}^{M}(\alpha_{mn}+\beta_{mn})C_{nl}H_1^{(1)}(l\kappa)$$

$$F_{mlb1} = [(2\Omega+S)H_1^{(2)}(m\kappa)-m\kappa S H_o^{(2)}(m\kappa)]\delta_{ml} + (2\Omega-S)\sum_{n=1}^{M}(\alpha_{mn}+\beta_{mn})C_{nl}H_1^{(2)}(l\kappa)$$

$$F_{moa2} = (2\Omega-S)\left[\alpha_{mo}+\left(\frac{S+\Omega}{\Omega}\right)^2\beta_{mo}+\sum_{n=1}^{M}\sum_{l=1}^{M}\left(\alpha_{mn}+\left[\frac{S+\Omega}{\Omega}\right]^2\beta_{mn}\right)C_{nl}\beta_{lo}\right]$$

$$F_{mob2} = F_{moa2}$$

$$F_{mla2} = \left[\left(\frac{2\Omega^3+5\Omega^2 S-S^3}{\Omega^2 f}\right)H_1^{(1)}(m\kappa f)-(m\kappa S)H_o^{(1)}(m\kappa f)\right]\delta_{ml} + (2\Omega-s)\sum_{n=1}^{M}\left(\alpha_{mn}+\left[\frac{S+\Omega}{\Omega}\right]^2\beta_{mn}\right)C_{nl}H_1^{(1)}(l\kappa)$$

$$F_{mlb2} = \left[\left(\frac{2\Omega^3+5\Omega^2 S-S^3}{\Omega^2 f}\right)H_1^{(2)}(m\kappa f)-m\kappa S H_o^{(2)}(m\kappa f)\right]\delta_{ml} + (2\Omega-s)\sum_{n=1}^{M}\left(\alpha_{mn}+\left[\frac{S+\Omega}{\Omega}\right]^2\beta_{mn}\right)C_{nl}H_1^{(2)}(l\kappa)$$

Pipe Bend Analysis by Thin Shell Theory

J. F. Whatham
Australian Atomic Energy Commission,
Research Establishment,
Sutherland, NSW, 2232, Australia
Mem. ASME

Thin shell theory is applied to pipe bends terminated by flanges or flange-ended tangent pipes and subjected to any end loading, either in-plane or out-of-plane. Graphs of flexibility factor versus pipe bend characteristic are presented for in-plane bending of a wide range of pipe elbows terminated by flanges or short flange-ended tangents. Experimental results verify the thin shell solutions for in-plane and out-of-plane bending of a flanged pipe elbow. The capabilities of a computer program BENDPAC are also described.

Introduction

The general linear shell equations for analyzing straight pipes or circular cylindrical shells are well known (Sanders, 1983) and an attempt is now made to present equivalent equations for analyzing pipe bends or discontinuous toroidal shells—because of this discontinuity a pipe bend is more akin to part of a spiral than part of a torus. The pipe bend loading is by force or moment in any direction on the ends, the pipe bend being terminated by rigid flanges or tangent pipes.

The equations of Novozhilov (1970) are used because they represent a consistent linear theory and the reciprocal theorem holds; that is, the flange-ended pipe bend flexibility matrix is symmetrical and hence a useful check on calculations (Whatham, 1982). The review of pipe bend analysis of Axelrad and Emmerling (1984) indicates that the application of thin shell theory usually involves some simplifying assumptions; no approximations are made here other than those inherent in the theory, so the solutions are exact.

Governing Equations

A segment of curved pipe is represented in Fig. 1 by its middle surface, acted upon by stress resultants to produce displacements u, v, w and a rotation ψ about axis u; there are rotations about the other axes but these do not enter the analysis.

From the Novozhilov equations we obtain:

(*i*) **Equilibrium Equations**

$$\frac{\partial}{\partial\theta}\left(\delta T_\theta^* + \delta M_\theta^*\right) + \frac{\sin\theta}{\rho}\left(T_\eta^* + M_\eta^*\right) + \frac{\partial P^*}{\partial\eta} = 0$$

$$\frac{\partial}{\partial\eta}\left(T_\eta^* + M_\eta^*\right) - \frac{\sin\theta}{\rho}P^* + \frac{\partial}{\partial\theta}\left(\delta P^*\right) - Q^* = 0 \qquad (1)$$

$$\delta T_\theta^* - \frac{\partial^2}{\partial\theta^2}\left(\delta M_\theta^*\right) + \frac{\cos\theta}{\rho}\left(T_\eta^* - M_\eta^*\right) - \frac{\sin\theta}{\rho}\frac{\partial M_\eta^*}{\partial\theta} - \frac{\partial Q^*}{\partial\eta} = 0$$

where $P^* = T_{\eta\theta}^* + M_{\eta\theta}^*$

$$Q^* = \frac{1}{\delta}\left[\frac{\partial M_\eta^*}{\partial\eta} + \frac{\partial}{\partial\theta}\left(\delta M_{\eta\theta}^* + \delta M_{\theta\eta}^*\right)\right]$$

$$T_\eta^* = T_\eta/Et,\ M_\eta^* = M_\eta/rEt$$

$$\rho = R/r,\ \delta = 1 + \frac{\cos\theta}{\rho}.$$

(*ii*) **Stress-Strain Equations**

$$T_\theta^* = \frac{1}{1-\nu^2}\left(\epsilon_\theta + \nu\epsilon_\eta\right) \qquad M_\theta^* = \frac{\gamma r}{12(1-\nu^2)}\left(\kappa_\theta + \nu\kappa_\eta\right)$$

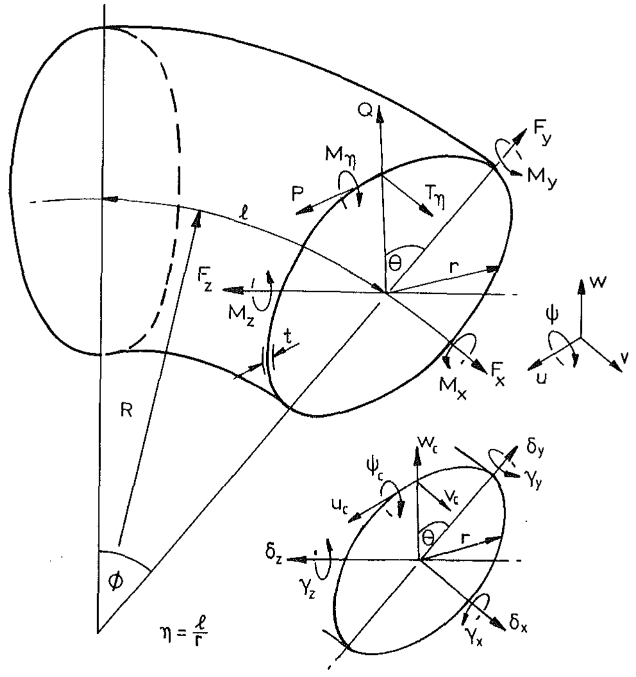


Fig. 1 Pipe bend middle surface

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript recevied by ASME Applied Mechanics Division, November 29, 1984; final revision, July 16, 1985.

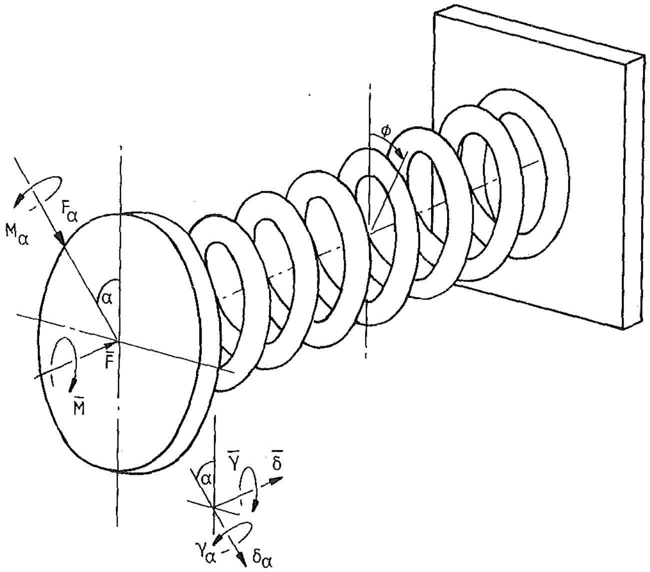


Fig. 2 Spiral pipe configuration

$$T^*_\eta = \frac{1}{1-\nu^2}(\epsilon_\eta + \nu\epsilon_\theta) \qquad M^*_\eta = \frac{\gamma r}{12(1-\nu^2)}(\kappa_\eta + \nu\kappa_\theta)$$

$$T^*_{\eta\theta} = \frac{1}{2(1+\nu)}\left(\omega + \frac{\gamma}{6} r\tau\right) \qquad M^*_{\eta\theta} = M^*_{\theta\eta} = \frac{\gamma}{12(1+\nu)} r\tau \tag{2}$$

where $\gamma = (t/r)^2$.

(*iii*) Strain-Displacement Equations

$$\epsilon_\theta = \frac{1}{r}\left(\frac{\partial u}{\partial\theta} + w\right)$$

$$\epsilon_\eta = \frac{1}{\delta r}\left(\frac{\partial v}{\partial\eta} - \frac{\sin\theta}{\rho}u + \frac{\cos\theta}{\rho}w\right)$$

$$\omega = \frac{1}{\delta r}\left(\frac{\partial u}{\partial\eta} + \frac{\sin\theta}{\rho}v\right) + \frac{1}{r}\frac{\partial v}{\partial\theta}$$

$$\lambda = \frac{1}{r}\frac{\partial v}{\partial\theta} - \frac{\partial\psi}{\partial\theta}$$

$$\kappa_\theta = \frac{1}{r^2}\left(\frac{\partial u}{\partial\theta} - \frac{\partial^2 w}{\partial\theta^2}\right)$$

$$\kappa_\eta = \frac{1}{\delta r}\left(\frac{\partial\psi}{\partial\eta} + \frac{\sin\theta}{\rho r}\left(\frac{\partial w}{\partial\theta} - u\right)\right) \tag{3}$$

$$\tau = \frac{1}{\delta r^2}\left(\frac{\partial u}{\partial\eta} + \frac{\sin\theta}{\rho}v\right) + \frac{1}{r}\frac{\partial\psi}{\partial\theta}$$

$$\beta = \frac{1}{r}\frac{\partial^2 v}{\partial\theta^2} + \psi$$

where $\psi = \frac{1}{\delta r}\left(\frac{\cos\theta}{\rho}v - \frac{\partial w}{\partial\eta}\right)$.

Strains λ and β require some explanation; they represent the twist and out-of-plane curvature given to the pipe wall by warping of the pipe cross-section—ψ should be $\sin\psi$ in β but ψ is small. The four strains ϵ_θ, κ_θ, λ, β completely describe the distortion of the pipe cross-section and therefore must be continuous along a stressed pipe; it is more convenient to equate them than the displacements across the junction of two pipes.

It is interesting that, allowing for the different notation, the Novozhilov equations are identical to earlier equations of Love (1944) except in expressions for τ and $T^*_{\eta\theta}$. Thus, if shear is zero, as in the pure in-plane bending of a curved pipe neglecting end effects, the Love equations give the same solution as those of Novozhilov. Furthermore, in his classic paper, Love (1888) had given the τ equation of Novozhilov, so, overall, the only difference is in the expressions for $T^*_{\eta\theta}$; Love had

$$T^*_{\theta\eta} = T^*_{\eta\theta} = \frac{1}{2(1+\nu)}\omega \tag{4}$$

which causes the flanged pipe bend flexibility matrix to lose symmetry.

Combining equations (1), (2) and (3), we obtain eight governing equations

$$\frac{\partial}{\partial\eta}\xi = A\xi \tag{5}$$

where

$$A = \begin{bmatrix} 0 & A_2 \\ A_1 & 0 \end{bmatrix} = 8 \times 8 \text{ matrix}$$

$$\xi = (u/r, w/r, M^*_\eta, T^*_\eta, v/r, \psi, P^*, Q^*).$$

Submatrices A_1 and A_2, involving differentials in θ only, are:

A₁

$\frac{u}{r}$	$\frac{w}{r}$	M^*_η	T^*_η
$\frac{\sin\theta}{\rho} - \nu\delta\frac{\partial}{\partial\theta}$	$-\frac{\cos\theta}{\rho} - \nu\delta$	0	$(1-\nu^2)\delta$
$\frac{\sin\theta}{\rho} - \nu\delta\frac{\partial}{\partial\theta}$	$-\frac{\sin\theta}{\rho}\frac{\partial}{\partial\theta} + \nu\delta\frac{\partial^2}{\partial\theta^2}$	$\frac{12(1-\nu^2)}{\gamma}\delta$	0
$-\left(1+\frac{\gamma}{12}\right)\frac{\partial}{\partial\theta}\left(\delta\frac{\partial}{\partial\theta}\ \right)$	$-\frac{\partial}{\partial\theta}\left(\delta - \frac{\gamma}{12}\delta\frac{\partial^2}{\partial\theta^2}\ \right)$	$-\frac{\sin\theta}{\rho} - \nu\frac{\partial}{\partial\theta}(\delta\ \ )$	$-\frac{\sin\theta}{\rho} - \nu\frac{\partial}{\partial\theta}(\delta\ \ )$
$\delta\frac{\partial}{\partial\theta} - \frac{\gamma}{12}\frac{\partial^2}{\partial\theta^2}\left(\delta\frac{\partial}{\partial\theta}\ \right)$	$\delta + \frac{\gamma}{12}\frac{\partial^2}{\partial\theta^2}\left(\delta\frac{\partial^2}{\partial\theta^2}\ \right)$	$-\frac{\partial}{\partial\theta}\left(\frac{\sin\theta}{\rho}\ \right) - \nu\frac{\partial^2}{\partial\theta^2}(\delta\ \ )$	$\frac{\cos\theta}{\rho} + \nu\delta$

$\mathbf{A}_2$

$\frac{v}{r}$	ψ	P^*	Q^*
$-\frac{\sin\theta}{\rho}-\frac{3}{3+\gamma}\,\delta\,\frac{\partial}{\partial\theta}$	$-\frac{\gamma}{3+\gamma}\,\delta\,\frac{\partial}{\partial\theta}$	$\frac{6(1+\nu)}{3+\gamma}\,\delta$	0
$\frac{\cos\theta}{\rho}$	$-\delta$	0	0
$\frac{\gamma}{2(1+\nu)(3+\gamma)}\,\frac{\partial}{\partial\theta}\left(\delta\,\frac{\partial}{\partial\theta}\ \right)$	$\frac{-\gamma}{2(1+\nu)(3+\gamma)}\,\frac{\partial}{\partial\theta}\left(\delta\,\frac{\partial}{\partial\theta}\ \right)$	$-\frac{\gamma}{3+\gamma}\,\frac{\partial}{\partial\theta}(\delta\ )$	δ
$\frac{-\gamma}{2(1+\nu)(3+\gamma)}\,\frac{\partial}{\partial\theta}\left(\delta\,\frac{\partial}{\partial\theta}\ \right)$	$\frac{\gamma}{2(1+\nu)(3+\gamma)}\,\frac{\partial}{\partial\theta}\left(\delta\,\frac{\partial}{\partial\theta}\ \right)$	$\frac{\sin\theta}{\rho}-\frac{3}{3+\gamma}\,\frac{\partial}{\partial\theta}(\delta\ )$	$-\frac{\cos\theta}{\rho}$

Kalnins (1969) used equations similar to (5) with the same eight variables to calculate the stresses in a particular pipe elbow terminated by tangent pipes and subjected to in-plane and out-of-plane bending. The stress-strain relations were more elaborate than equations (2) and the govering equations were solved by step-by-step integration along the pipe with finite difference equations in θ. Nevertheless, the solutions obtained compare well with the closed form solutions of the present work.

Spiral Analogy

Analyzing a pipe bend under the minimum number of loadings, to cover every loading case by superposition of solutions and neglecting end effects, is equivalent to analyzing a spiral tube under the four loadings, F_α, M_α, $\bar{F}$, $\bar{M}$ shown in Fig. 2. The spiral is conceived as having zero pitch, the turns being coincident but not interfering with each other.

Each of the four load/displacement combinations is decoupled—for example, force F_α only produces displacement δ_α—so that the four cases are solved independently; a pipe bend taken from a quarter of a turn of the spiral would be loaded as shown in Fig. 3.

Consider the solution to equation (5) as the sum of four vectors:

$$\xi=\xi_a+\xi_b+\xi_c+\xi_d \tag{6}$$

where $\xi_a=(0,0,0,T^*_{\eta a},0,0,P^*_a,0)$

$$\xi_b=(u_b/r,\ w_b/r,\ M^*_{\eta b},\ T^*_{\eta b},\ v_b/r,\ \psi_b,\ P^*_b,\ Q^*_b)$$

$$\xi_c=(u_c/r,\ w_c/r,\ 0,\ 0,\ v_c/r,\ \psi_c,\ 0,\ 0)$$

$$\xi_d=(u_d/r,\ w_d/r,\ M^*_{\eta d},\ T^*_{\eta d},\ v_d/r,\ \psi_d,\ P^*_d,\ Q^*_d)$$

Vector ξ_a consists of the known elements of the problem—two stress resultants $T^*_{\eta a}$, P^*_a—which are in equilibrium with one of the spiral loadings F_α, M_α, $\bar{F}$ or $\bar{M}$, The stress resultant distributions required are the simplest to give the loading.

Vector ξ_b is a complete set of stress resultants and displacements; the stress resultants are self-equilibrating, producing no net forces or moments on the pipe cross-section, while the displacements are repeated each turn of the spiral as in a torus. The vector components are in two groups for their ϕ dependence; u_b/r, w_b/r, $M^*_{\eta b}$, $T^*_{\eta b}$ vary as $T^*_{\eta a}$, and v_b/r, ψ_b, P^*_b, Q^*_b vary as P^*_a.

Vector ξ_c consists of displacement components which increase with each turn of the spiral. They are compatible with the relevant spiral end displacement δ_α, γ_α, $\bar{\delta}$ or $\bar{\gamma}$ but the strains they generate must be repeated in each turn of the spiral. It happens that these displacements do not distort the pipe cross-section nor do they affect the spiral diameter; the deflected shape for δ_α is a series of loops.

Vector ξ_d represents end effects and, like vector ξ_b, consists of self-equilibrating stress resultants and displacements. All components decay with distance from the pipe end, satisfying the equation

$$\frac{\partial}{\partial\eta}\,\xi_d=A\,\xi_d. \tag{7}$$

Solving Without End Effects

Subtracting the end effects from equation (5)

$$\frac{\partial}{\partial\eta}(\xi_a+\xi_b+\xi_c)=A\,(\xi_a+\xi_b+\xi_c).$$

Rearranging

$$\left(A-I\frac{\partial}{\partial\eta}\right)\xi_b+\left(A-I\frac{\partial}{\partial\eta}\right)\xi_c=-\left(A-I\frac{\partial}{\partial\eta}\right)\xi_a \tag{8}$$

where $I=$ unit diagonal matrix.

To derive the ξ_a components, we refer to Fig. 1 and note that the net forces and moments acting on the pipe cross-section, from the spiral end loading, equal the stress resultants integrated around the circumference:

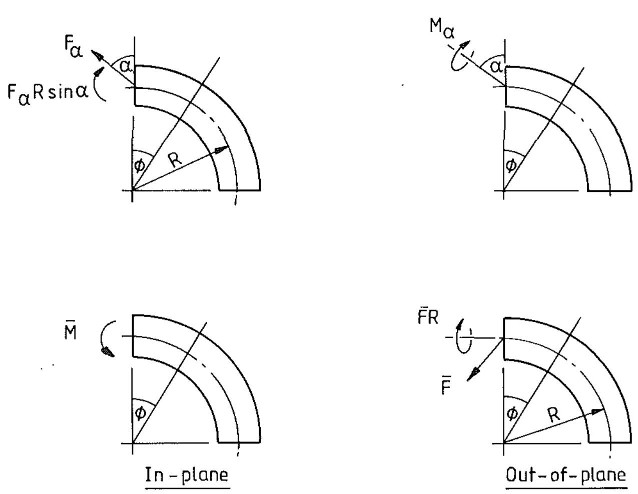


Fig. 3 Spiral loads on a pipe bend

$$\left\{\begin{matrix} F_\alpha r\sin(\alpha+\phi) \\ -F_\alpha r\cos(\alpha+\phi) \\ -\bar{F}r \\ M_\alpha\sin(\alpha+\phi)+\bar{F}R \\ -M_\alpha\cos(\alpha+\phi) \\ F_\alpha R\sin(\alpha+\phi)-\bar{M} \end{matrix}\right\} = \left\{\begin{matrix} F_xr \\ F_yr \\ F_zr \\ M_x \\ M_y \\ M_z \end{matrix}\right\} = r^2Et\int_0^{2\pi}\left\{\begin{matrix} T^*_\eta \\ -P^*\sin\theta+Q^*\cos\theta \\ P^*\cos\theta+Q^*\sin\theta \\ P^* \\ (M^*_\eta+T^*_\eta)\sin\theta \\ -(M^*_\eta+T^*_\eta)\cos\theta \end{matrix}\right\}d\theta. \qquad (9)$$

In matrix form:

$$r^2Et\int_0^{2\pi} B\bullet x\, d\theta = F \qquad (10)$$

where $B = 6\times 4$ matrix operator

$x = (M^*_\eta, T^*_\eta, P^*, Q^*)$

$F = (F_xr, F_yr, F_zr, M_x, M_y, M_z)$.

The dimensionless stress resultants are divided into two parts:

$$x = x_a + x_b \qquad (11)$$

where $\int_0^{2\pi} B\bullet x_a\, d\theta = F/r^2Et$

$$\int_0^{2\pi} B\bullet x_b\, d\theta = 0.$$

The self-equilibrating condition for x_b is easily satisfied in a round pipe if x_b components are expressed as Fourier series. As an aside, if the governing equations had been in terms of variables ϵ_θ, κ_θ, M_η, T_η, λ, β, P, Q (Whatham 1981 b) then the displacement u, v, w, ψ would have been continuous around the pipe if

$$\int_0^{2\pi} B\bullet y\, d\theta = 0 \qquad (12)$$

where $y = (-\epsilon_\theta, \kappa_\theta, \lambda, \beta)$.

As with vector x_b, this condition is easily satisfied in a round pipe.

From equations (9) the simplest x_a components are

$$T^*_{\eta a} = (\tfrac{1}{2}F_x r + M_y\sin\theta - M_z\cos\theta)/\pi r^2Et$$
$$P^*_a = (\tfrac{1}{2}M_x - F_y r\sin\theta + F_z r\cos\theta)/\pi r^2Et$$
$$M^*_{\eta a} = Q^*_a = 0. \qquad (13)$$

The components of ξ_a are then

$$T^*_{\eta a} = F^*_\alpha\sin(\alpha+\phi)(\tfrac{1}{2}-\rho\cos\theta),$$
$$-M^*_\alpha\cos(\alpha+\phi)\sin\theta \quad \text{or} \quad \bar{M}^*\cos\theta$$
$$P^*_a = F^*_\alpha\cos(\alpha+\phi)\sin\theta,$$
$$\tfrac{1}{2}M^*_\alpha\sin(\alpha+\phi) \quad \text{or} \quad \bar{F}^*\left(\frac{\rho}{2}-\cos\theta\right)$$

where F^*_α, $\bar{F}^* = (F_\alpha, \bar{F})/\pi rEt$ and M^*_α, $\bar{M}^* = (M_\alpha, \bar{M})/\pi r^2Et$.

To derive the ξ_c components we again refer to Fig. 1, where the displacements and rotations δ_x, δ_y, δ_z, γ_x, γ_y, γ_z represent rigid body movements of the pipe cross-section; their relationships to the spiral end displacements resemble equations (9)

$$\left\{\begin{matrix} \gamma^*_\alpha r\phi\sin(\alpha+\phi) \\ -\gamma^*_\alpha r\phi\cos(\alpha+\phi) \\ -\bar{\gamma}^* r\phi \\ \delta^*_\alpha r\phi\sin(\alpha+\phi)+\bar{\gamma}^*R\phi \\ -\delta^*_\alpha r\phi\cos(\alpha+\phi) \\ \gamma^*_\alpha R\phi\sin(\alpha+\phi)-\bar{\delta}^* r\phi \end{matrix}\right\} = \left\{\begin{matrix} \gamma_x r \\ \gamma_y r \\ \gamma_z r \\ \delta_x \\ \delta_y \\ \delta_z \end{matrix}\right\} \qquad (15)$$

where δ^*_α, $\bar{\delta}^* = (\delta_\alpha, \bar{\delta})/2\pi rn$, γ^*_α, $\bar{\gamma}^* = (\gamma_\alpha, \bar{\gamma})/2\pi n$, n = number of spiral turns.

From the geometry of Fig. 1

$$u_c = \gamma_x r - \delta_y\sin\theta + \delta_z\cos\theta$$
$$w_c = \delta_y\cos\theta + \delta_z\sin\theta$$
$$v_c = \delta_x + \gamma_y r\sin\theta - \gamma_z r\cos\theta$$
$$\psi_c = \gamma_y\sin\theta - \gamma_z\cos\theta. \qquad (16)$$

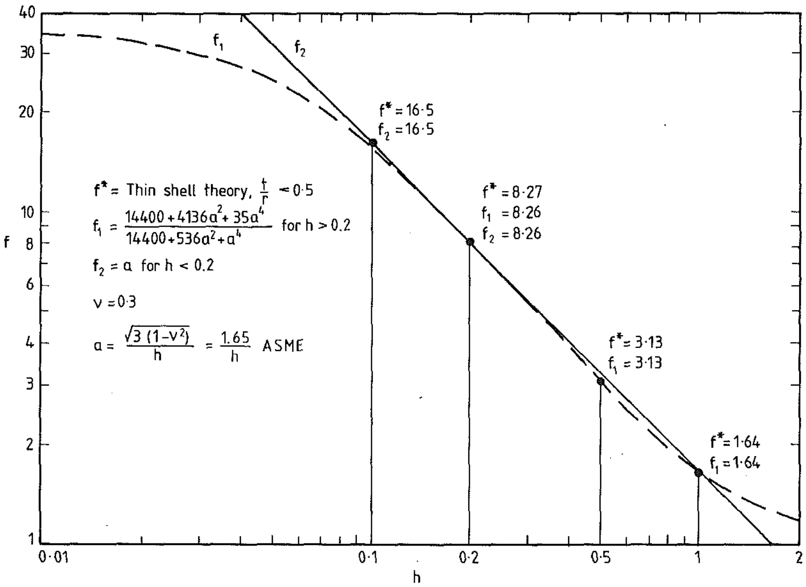
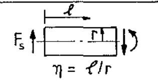
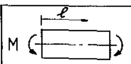
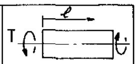


Fig. 4 Flexibility of pipe bends without end effects

Table 1 Flexibility factors
(Refer Fig. 4)

$\nu = 0.3$				Thin shell theory		
h	a^2	f_1	f_2	$\frac{t}{r} = 0.01$	0.05	0.1
0.1	273	-	16.523	16.518	16.491	-
0.2	68.25	8.2623	8.2614	8.2739	8.2627	8.2082
0.5	10.92	3.1287	-	3.1288	3.1310	3.1375
1.0	2.73	1.6352	-	1.6353	1.6359	1.6378

Table 2 Straight pipe solutions

		 $\eta = \ell/r$		
$\frac{u}{r}$		$\left((2+\nu)\eta-\frac{\eta^3}{6}+\nu\eta\right)\frac{r^2F_s}{EI}\sin\theta$	$\frac{(\eta^2-2\nu)rM}{2EI}\sin\theta$	$-\frac{\eta(1+\nu)rT}{EI}$
$\frac{w}{r}$	$\frac{\nu F_a}{EA}$	$-\left((2+\nu)\eta-\frac{\eta^3}{6}\right)\frac{r^2F_s}{EI}\cos\theta$	$-\frac{\eta^2 rM}{2EI}\cos\theta$	
M^*_η		$-\frac{\gamma\eta r^2F_s}{12EI}\cos\theta$	$\frac{\gamma rM}{12EI}\cos\theta$	
T^*_η	$-\frac{F_a}{EA}$	$-\frac{\eta r^2F_s}{EI}\cos\theta$	$\frac{rM}{EI}\cos\theta$	
$\frac{v}{r}$	$-\frac{\eta F_a}{EA}$	$-\frac{\eta^2r^2F_s}{2EI}\cos\theta$	$\frac{\eta rM}{EI}\cos\theta$	
ψ		$\left(2+\nu-\frac{\eta^2}{2}\right)\frac{r^2F_s}{EI}\cos\theta$	$\frac{\eta rM}{EI}\cos\theta$	
P^*		$\left(1+\frac{\gamma\nu}{6(1+\nu)}\right)\frac{r^2F_s}{EI}\sin\theta$		$-\frac{(3+\gamma)rT}{6EI}$
Q^*		$-\frac{\gamma(1-\nu)r^2F_s}{12(1+\nu)EI}\cos\theta$		
I		$\pi r^3t(1+\frac{\gamma}{12})$	$\pi r^3t(1+\frac{\gamma}{12})$	$\pi r^3t(1+\frac{\gamma}{3})$

The components of ξ_c are then

$$\begin{aligned}
u_c/r &= \delta^*_\alpha\phi\cos(\alpha+\phi)\sin\theta, \quad \gamma^*_\alpha\phi\sin(\alpha+\phi)(1+\rho\cos\theta)\\
&\qquad \text{or} \quad -\bar{\delta}^*\phi\cos\theta\\
w_c/r &= -\delta^*_\alpha\phi\cos(\alpha+\phi)\cos\theta, \quad \gamma^*_\alpha\phi\sin(\alpha+\phi)\rho\sin\theta\\
&\qquad \text{or} \quad -\bar{\delta}^*\phi\sin\theta \qquad\qquad (17)\\
v_c/r &= \delta^*_\alpha\phi\sin(\alpha+\phi), \quad -\gamma^*_\alpha\phi\cos(\alpha+\phi)\sin\theta\\
&\qquad \text{or} \quad \bar{\gamma}^*\phi(\rho+\cos\theta)\\
\psi_c &= -\gamma^*_\alpha\phi\cos(\alpha+\phi)\sin\theta \quad \text{or} \quad \bar{\gamma}^*\phi\cos\theta.
\end{aligned}$$

Note that by equations (3), the distortion strains $\epsilon_{\theta c}$, $\kappa_{\theta c}$, λ_c, β_c are zero and all strains from ξ_c are repeated each turn of the spiral.

We are now in a position to solve equation (8). The result of $(A-I(\partial/\partial\eta))\xi_c$ is the vector $-\xi_c/\rho\phi$ whose components are repeated each turn of the spiral; it has only one unknown, namely the spiral end displacement for the particular loading. The ϕ terms cancel out leaving equation (8) in θ only; there are eight equations for F_α or M_α loading and four equations for $\bar{F}$ or $\bar{M}$ loading, with $\bar{F}$ only involving submatrix A_2 and $\bar{M}$ only involving A_1. Expressing the variables as truncated Fourier series with appropriate parity, and equating the coefficients of like terms (Fourier analysis) gives a series of simultaneous equations which then provide the solution.

Pipe Bend Flexibility Without End Effects

Pipe bend flexibility is a major factor in determining the forces transmitted by pipe networks and the ASME (1980) code recommends a formula for in-plane bending which neglects end effects. The formula has, however, proved effective on pipe bends contiguous with tangent pipes.

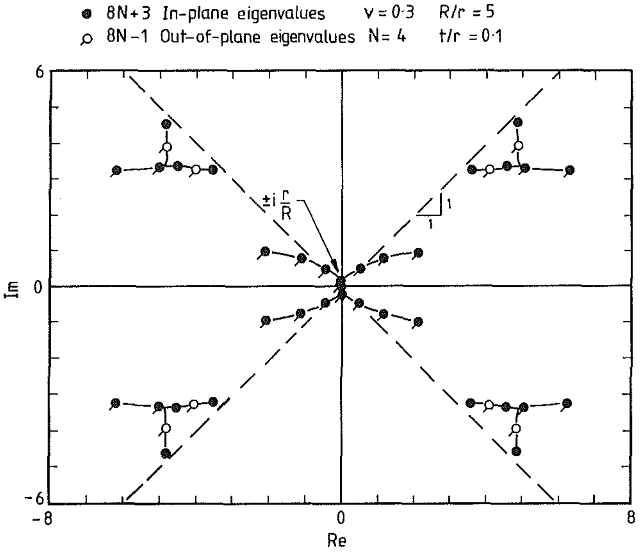


Fig. 5 Typical pipe bend eigenvalues

Consider the spiral loaded by $\bar{M}$, equivalent to pure in-plane bending of a curved pipe neglecting end effects; the end rotation is

$$\bar{\gamma} = f\,2\pi n\frac{\bar{M}R}{EI} \tag{18}$$

where flexibility factor f accounts for increased rotation, compared to that from elastic line theory, because the pipe cross-section flattens. Flexibility factor versus pipe bend characteristic ($h = Rt/r^2$) is plotted in Fig. 4 with some numerical values given in Table 1; Poisson's ratio was assumed to be 0.3.

When $a < 8$ and $t/r < 0.05$, the exact thin shell calculations verify the second approximation formula of von Kármán (1911):

$$f_1 = \frac{14400+4136a^2+35a^4}{14400+536a^2+a^4} \tag{19}$$

except that

$a = \sqrt{3(1-\nu^2)}/h$ (Clark and Reissner, 1951)

instead of

$a = \sqrt{3}/h$ (von Kármán).

For $a > 8$, the formula of Beskin (1945) applies

$$f_2 = a. \tag{20}$$

The ASME code effectively recommends equation (20) if we assume that $\nu = 0.3$.

Most pipe bend research concerns in-plane bending because flattening of the pipe cross-section has such a significant effect on bend flexibility and stress distribution. Now consider the out-of-plane loading of the spiral by force $\bar{F}$; the resulting displacement, by both thin shell theory and elastic line theory neglecting the direct shear component, is

$$\bar{\delta} = 2\pi n\frac{\bar{F}R^3}{GI_p} \tag{21}$$

where $G = E/2(1+\nu)$ and $I_p = 2I$

there being no flattening of the pipe cross-section. For small radius bends ($R < 5r$), the thin shell deflection would be expected to exceed that given by equation (21) because of the increased importance of the direct shear component, but, in fact, the thin shell deflection was less (Whatham, 1981a).

Straight Pipes

The solutions for straight pipes with various end loads are given in Table 2; the pipe cross-section remains circular for all

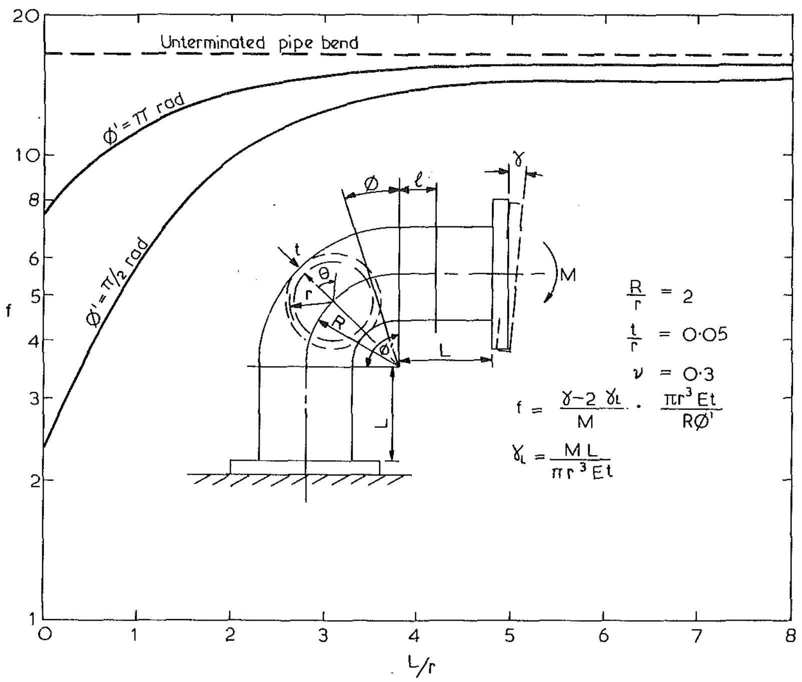


Fig. 6 Flexibility versus tangent length

loadings, despite the fact that F_s and M produce a cos θ distribution of hoop strain, so that beam theory applies. The moment of inertia of a round tube is exactly

$$I = \pi r^3 t(1+\gamma/4) \qquad (22)$$

but solving the problems by thin shell theory gave the slightly different expressions in the bottom row for apparent moments of inertia.

In the case of F_s loading, the deflection has a shear component which thin shell theory gave as

$$\delta_{\text{shear}} = (2+\nu)\eta\frac{r^3 F_s}{EI}. \qquad (23)$$

This compares with a finding by Cowper (1966) from a strength of materials analysis, if $I = \pi r^3 t$

$$\delta_{\text{shear}} = (4+3\nu)\eta\frac{r^3 F_s}{2EI} \qquad (24)$$

Incorporation of End Effects

We now turn our attention to the pipe bend terminations. Having obtained particular solutions to equation (8) which neglect end effects, we add sufficient complementary solutions from the homogeneous equation (7) to satisfy pipe end conditions. Vector ξ_d is a function of θ, η but, in seeking a separable solution, let

$$\xi_d = \xi(\theta)\, e^{\Omega\eta} \qquad (25)$$

Substituting in equation (7)

$$A\,\xi(\theta) = \Omega\,\xi(\theta). \qquad (26)$$

This is an eigenvalue problem; the variables were expressed by Fourier series truncated to N terms and, by Fourier analysis, $8N+3$ equations were obtained for in-plane F_α or $\bar{M}$ loading or $8N-1$ equations for out-of-plane M_α or $\bar{F}$ loading; the resulting equations have been published (Whatham, 1983). Eigenvalues and eigenvectors of equation (26) are all complex and can be obtained by a standard subroutine such as EB06AD (Hopper, 1973); a typical eigenvalue spectrum is shown in Fig. 5 where only those eigenvalues with negative real parts are required since end effects decay with distance from the end.

The three eigenvalues with zero real parts arise from the components of u, v, w, ψ, which displace the pipe cross-section as a rigid body. Had these components been neglected, or the governing equations written in terms of variables ϵ_θ, κ_θ, M_η, T_η, λ, β, P, Q, the three eigenvalues would not have been generated but then the rigid body displacements would have had to be calculated separately for the pipe bend deflection.

Combining eigenvectors for the total end effect

$$\xi_d = \operatorname{Re}\sum_{j=1}^{J} C_j\, \xi_j(\theta)\, e^{\Omega_j\eta} \qquad (27)$$

where

$$J = \begin{cases} 2N \text{ (in-plane loading)} \\ 2N-1 \text{ (out-of-plane loading)} \end{cases}$$

C_j = coefficient for jth eigenvector, determined from the end conditions

Since C_j, $\xi_j(\theta)$ and Ω_j are all complex, the end effect is described along the pipe by Fourier series with exponentially decaying coefficients.

Two curved pipe terminations are considered—a rigid flange and a tangent pipe. To represent a rigid flange, the distortion strains ϵ_θ, κ_θ, λ, β are made zero at the pipe end whereas, for the tangent pipe, the stress resultants and distortion strains are equated across the junction and thus values of C_j are obtained; these end conditions are discussed further by Whatham and Thompson (1979) and Whatham (1982).

Pipe Bend Flexibility With End Effects

Flexibility factors, derived for a typical pipe elbow and **U** bend with equal length flange-ended tangents, are plotted versus tangent length in Fig. 6. Obviously the shorter the tangents, the greater is the restriction to pipe cross-section flattening and the less is the flexibility. For the pipe dimension ratios shown, long tangents reduced the flexibility of the elbow and the **U** bend by 11.5 and 5.5 percent, respectively, compared to their flexibilities if end effects were neglected.

Flexibility factors are plotted versus bend characteristic h in Figs. 7 and 8 for pipe elbows with flanged ends or with short (L/r = 1) tangents with flanged ends under in-plane bending; additional graphs for other tangent lengths on **U** bends as well

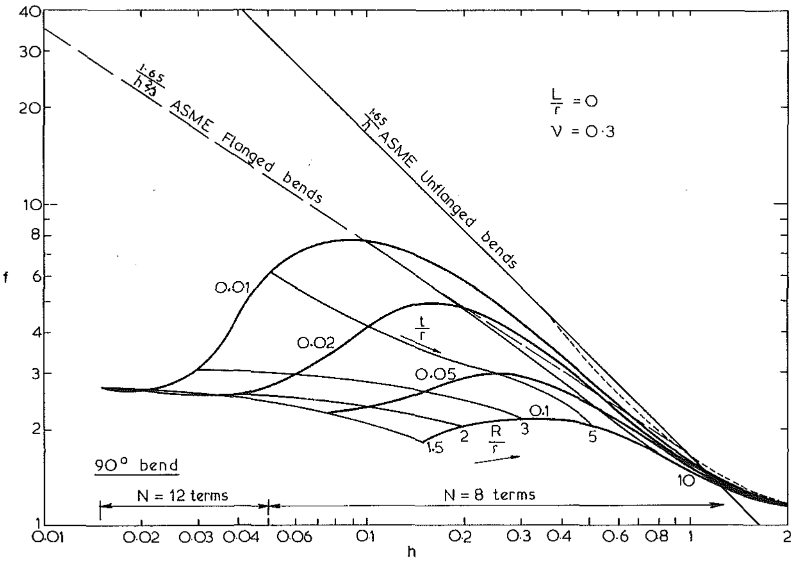


Fig. 7 Flexibilty of flanged elbows

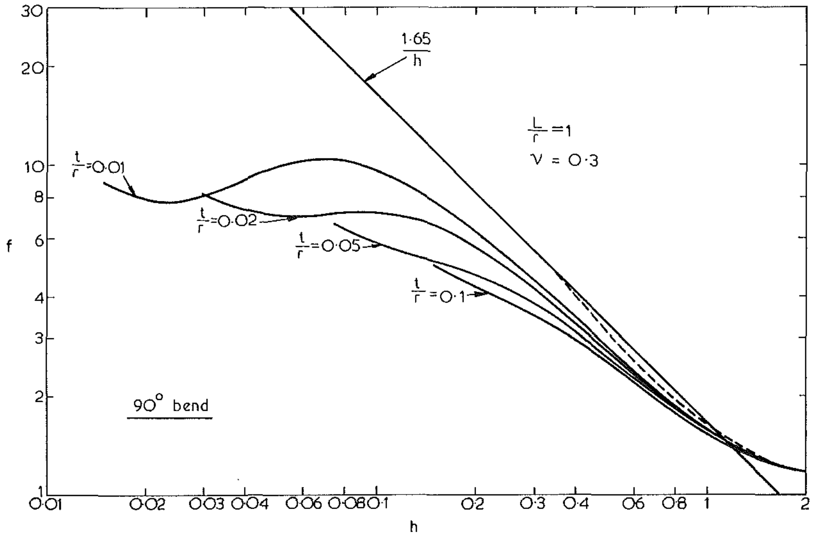


Fig. 8 Flexibilty of short tangent elbows

as pipe elbows have been presented (Whatham and Thompson, 1979) and values tabulated (Whatham, 1982-1985).

Verification by Experiment

In-plane and out-of-plane bending moments were applied to a cast steel pipe elbow of 125 mm bore, 375 mm radius of curvature and 12.5 mm wall thickness, which had flanges of 400 mm diameter and 90 mm thickness to provide sufficient stiffness at the ends.

Deflections matched those calculated by thin shell theory but stress distribution is a more sensitive indicator. Hoop stresses derived from strain gauge measurements taken on the outer surface midway between the elbow ends, and thin shell solutions with and without flanges, are plotted in Figs. 9 and 10; they verify the analysis and show the significant stress reduction effected by the flanges.

The BENDPAC Computer Program

A computer program package, BENDPAC, written in the FORTRAN IV and ASSEMBLER languages for an IBM3033 computer, is available from either the Australian Atomic Energy Commission or the National Energy Software Center, Argonne National Laboratory. Pipe bends with any loading or end condition shown in Fig. 11 may be analyzed for overall flexibility, stresses and cross-section distortion. Wall thickness should be between one and ten percent of the pipe radius, or may be as high as thirty percent for pure in-plane bending (Whatham, 1981a), but there is no restriction on the bend radius, bend angle, or tangent length; the interactions of effects from opposite ends of the curved pipe and tangents are taken into account. As far as computation is concerned, a bend angle may be several hundred degrees; this would be used, for example, if examining the end effect on a curved pipe without interference from the opposite end.

Tables have been published (Whatham, 1982-1985) giving the complete stress state, the cross-section distortion and the flexibility of a range of pipe elbows ($R/r = 2$, 3, 5 and $t/r = 0.01$–0.1) with flanged ends under in-plane or out-of-plane forces or bending moments, also with tangent pipe ends under in-plane or out-of-plane bending moments only.

Conclusions

Equations and solving procedure have been presented for obtaining thin shell solutions in closed form for pipe bends subjected to four particular end loadings, but, by superposition, solutions can then be obtained for any force or moment

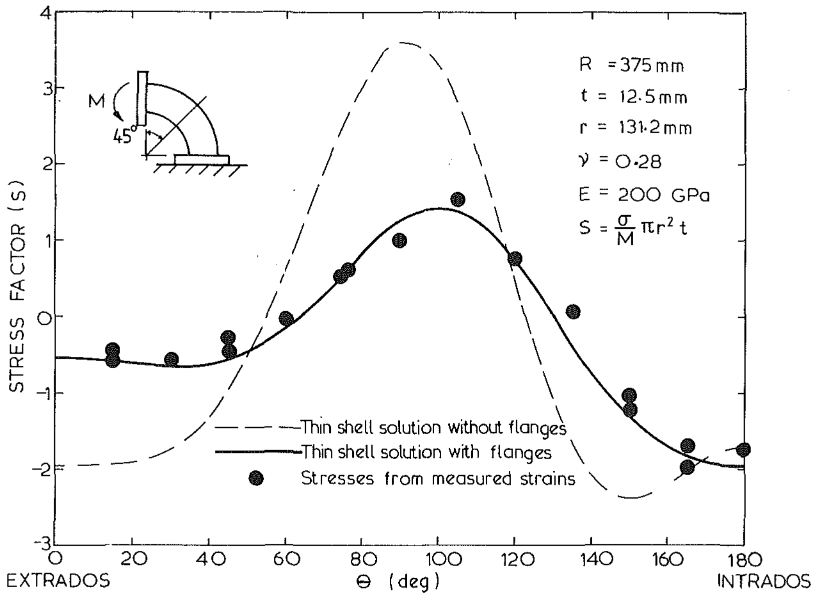


Fig. 9 In-plane bending hoop stress on outside surface

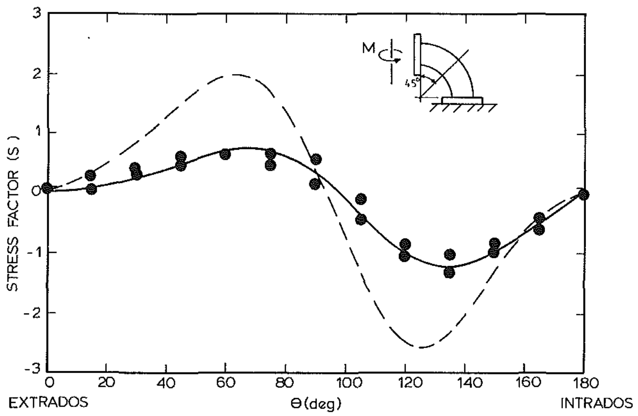


Fig. 10 Out-of-plane bending hoop stress on outside surface

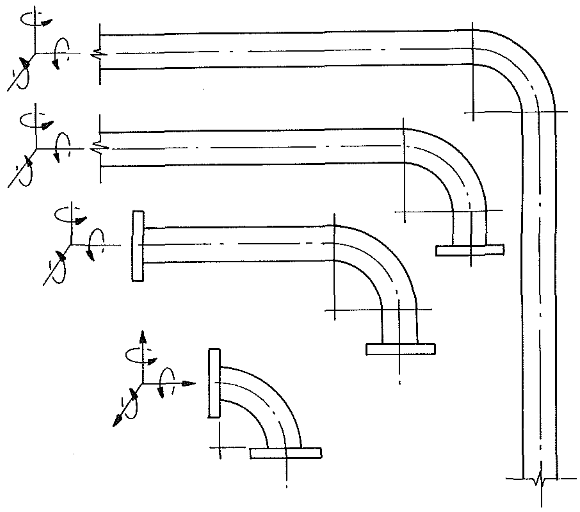


Fig. 11 Loading options of BENDPAC

loading on the ends, either in-plane or out-of-plane. The method of including pipe bend terminations, consisting of rigid flanges or tangent pipes, was then described and the stresses, calculated in a flange ended pipe elbow from in-plane and out-of-plane bending, were checked by experiment.

The work has resulted in a computer program BENDPAC and published tables of stresses, cross-section distortions and flexibilities for a range of pipe elbows. As the analysis is exact, within the limitations of thin shell theory, it provides benchmark solutions for assessing numerical methods of pipe bend analysis as well as providing design data.

Acknowledgment

The author acknowledges the advice and encouragement of Professor J. J. Thompson of the School of Nuclear Engineering, University of New South Wales.

References

ASME, 1980, *Power Piping Standard* ANSI-B31.1, American Society of Mechanical Engineers.

Axelrad, E. L., and Emmerling, F. A., 1984, "Elastic Tubes," ASME *Applied Mechanics Reviews*, Vol. 37, pp. 891-897.

Beskin, L., 1945, "Bending of Curved Thin Tubes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 12, pp. A-1 to A-7.

Clark, R. A., and Reissner, E., 1951, "Bending of Curved Tubes," *Advances in Applied Mechancis*, Vol. 2, Academic Press, New York, pp. 93-122.

Cowper, G. R., 1966, "The Shear Coefficient in Timoshenko's Beam Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 335-340.

Hopper, M. J., 1973, "Harwell Subroutine Library," AERE-R7477.

Kalnins, A., 1969, "Stress Analysis of Curved Tubes," *Proceedings of the First International Conference on Pressure Vessel Technology*, Delft, pp. 223-235.

Kårmån, Th. von, 1911, "Uber die Formanderung dunnwandiger Rohre," *Zerischrift des Vereines Deutscher Ingenieure*, Vol. 55, pp. 1889-1895.

Love, A. E. H., 1888, "On the Small Free Vibrations and Deformations of Thin Elastic Shells," *Philosophical Transactions of the Royal Society*, Vol. 129(A), pp. 491-546.

Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, 4th Ed., Dover Publications, New York.

Novozhilov, V. V., 1970, *Thin Shell Theory*, Wolters-Noordhoff, Groningen.

Sanders, J. L., 1983, "Analysis of Circular Cylindrical Shells," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 1165-1170.

Whatham, J. F., and Thompson, J. J., 1979, "The Bending and Pressurizing of Pipe Bends with Flanged Tangents," *Journal of Nuclear Engineering and Design*, Vo. 54, pp. 17-28.

Whatham, J. F., 1981a, "Thin Shell Equations for Circular Pipe Bends," *Journal of Nuclear Engineering and Design*, Vol. 65, pp. 77-89.

Whatham, J. F., 1981b, "Thin Shell Analysis of Non-Circular Pipe Bends," *Journal of Nuclear Engineering and Design*, Vol. 67, pp. 287-296.

Whatham, J. F., 1982, "Analysis of Circular Pipe Bends with Flanged Ends," *Journal of Nuclear Engineering and Design*, Vol. 72. pp. 175-187.

Whatham, J. F., 1983, "Thin Shell Analysis of Flanged Pipe Bends," *Transactions of the Institution of Engineers Australia*, Vol. CE25, pp. 1-12.

Whatham, J. F., 1982-1985, "Results of Pipe Bend Analysis," Parts I-XII, Australian Atomic Energy Commission Reports E551 to E555, E576, E577, E585 to E588 and E605.

Flow-Induced Vibrations Caused by Roughness in Pipes Conveying Fluid

A. Shulemovich
New York, NY
Mem. ASME

This paper presents a theoretical investigation of self-excited vibrations of pipes conveying fluid due to roughness. A model of a laminar friction, considered as the excitation mechanism, is based on Prandtl's universal velocity distribution for the turbulent boundary layer and on Nikuradse's experiments. The analysis has shown that the friction characteristic has a negative slope in a certain range of fluid velocities. This range is a function of pipe roughness and is shifted to lower flow velocities due to roughness growth during pipe operation. It was found that the differential operator of a piping loop motion based on the nonlinear restoring characteristic coincides with the differential operator of Duffing's equation for the hardening system. The energy method was used to obtain the approximate closed-form solution for the amplitude of steady self-excited vibrations. The unstable response with jump phenomena can appear due to interaction of small turbulent disturbances in conveying fluid with a given nonlinear system.

Introduction

Flow-induced vibrations of pipes conveying fluid have been studied by many investigators. In the extensive review papers of Chen [1, 2] there are numerous references pertaining to this problem.

The particular interest of the present study is aimed at the self-excited vibrations of piping loops conveying fluid, since the mechanism of such vibrations is still not well understood. Sustained axial vibrations with perceptible amplitudes (Fig. 1) can be observed when a certain flow velocity threshold is exceeded. A given problem may be classified as parallel-flow-induced vibration of cylindrical, elastically supported structures subjected to internal flow [2].

Three excitation mechanisms are usually adopted for the analysis of flow-induced vibrations, namely: self-excited vibration, parametric and combination resonances, and forced vibrations. The self-excited vibration mechanism for axial flow was considered by Burgreen et al. [3], Benjamin [4], Quinn [5, 6] and by other researchers [1, 2].

The parametric and combination resonances and forced vibrations with periodically disturbed flow were considered by Hopkins [7], Y. N. Chen [8], S. S. Chen [9], Ginsberg [10], Paidoussis et al. [11, 12, 13, etc.], and by many other investigators for pipes with different boundary conditions and different fluid flows: steady flow, pulsating flow and two-phase flow.

The limited interest in the self-excited vibration for pipes subjected to axial flow may apparently be accounted for by the relatively small amplitudes caused by this type of flow-induced vibration. M. P. Paidoussis [14, 15] examined a "tentative" analytical model of Quinn and found that a flow-induced damping was not taken into account and concluded that the forced vibration mechanism is the most appropriate model for the transverse vibration of pipes immersed in axial flow.

However, the self-excited mechanism combined with other excitation can apparently play a substantial role in the appearance of considerable axial vibrations for piping loops conveying fluid (Fig. 1).

The sustained flow-induced vibrations of piping loops and elements of the attached equipment are often accompanied by hysteresis effects (jump phenomenon). The jump phenomenon was observed, for instance, for the upper parts of actuators of regulating valves on feedwater lines (NSP Prairie Island Nuclear Power Plant) and for a main steam line

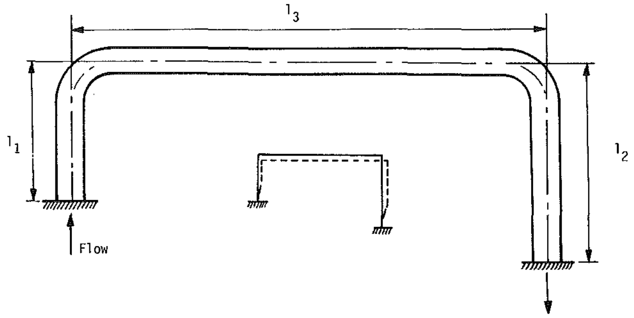


Fig. 1 Model of the piping loop

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. 10077, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May, 1984; final revision, April, 1985.

(Nebraska Cooper Nuclear Station) with average axial amplitudes in the range 4–7 mm.

The jump phenomenon is known for the forced vibrations of lightly damped nonlinear systems. If the amplitudes of excitation are small, the hysteresis effects, owing to intrinsic structural damping, do not appear. However, the hysteresis effects may appear if the self-excited mechanism, neutralizing structural damping, is involved in the excitation.

This paper deals with a model of the self-excited vibration mechanism focused on the fluid-roughness interaction when Reynolds numbers are very large (e.g., for the main steam line of the Nebraska Cooper Nuclear Station, $\mathrm{Re} > 3 \times 10^7$). For a given model the aforementioned flow velocity threshold is apparently a function of the interaction between the pipe surface and turbulent flow with the laminar sublayer.

The first mode of vibration (the long span of length l_3 moves like a rigid body) will dominate in the dynamic motion, and only this mode is considered to be important for engineering purposes.

By analogy with a belt friction system, which exhibits self-excited vibrations, it can be shown that the laminar friction characteristic has a negative slope in a certain range of fluid velocities. In view of this, it has to be noted that all existing empirical and semiempirical theories have generally been developed for determining the head loss value. By Boussinesq's introduction [16] the turbulent resistance is expressed in analogy with the coefficient of viscosity in Stokes law for laminar flow through the so-called "apparent" or "virtual" (also "eddy") viscosity. The head loss consisting of laminar and turbulent frictions is usually calculated using a dimensionless resistance coefficient. The latter one is obtained from the experiments of Nikuradse [17], Bauer and Galavics [18], and other researchers. For instance, a model of self-excited vibrations proposed by Quinn [5] is based on the use of the total head loss. However, since laminar and turbulent resistances are different in physical nature, it is useful, for better understanding of self-excited vibrations caused by roughness, to consider them separately. Really, the laminar resistance is frictional and under certain conditions can act opposite to the dissipative forces of the piping loop. In turn, the turbulent resistance (uniform flow without fluctuations), as it will be shown below, interacts with the restoring elastic forces of the piping loop and, in case of steady vibrations, is proportional to the static displacement of the piping legs.

Self-exciting vibration occurs in an unconservative mechanical system if the damping resistance is balanced or exceeded by forces from some uniform source of power. For the model under consideration, the laminar boundary sublayer created by fluid flow represents such a source and under certain conditions can convert the mechanical system into a conservative one.

A diminution of the laminar friction can apparently be explained as the result of a decrease in the laminar sublayer's thickness with an increase of the flow velocity and subsequent interaction between the conveying fluid and the rough surface of the pipe. As was shown by Nikuradse's measurements, the velocity gradient near a rough wall is less steep than that near a smooth one. Actually, this effect originates in pipes conveying fluid when the thickness of a laminar sublayer becomes less than the height of protrusions. The viscous sublayer finds itself "in the shadow of protrusions" and practically ceases to exist. Thus a growth of the protrusion's height during operation or an increase of the flow velocity may eventually lead to the aforementioned phenomenon. This tentative model of the self-excited vibration mechanism based on the fluid-roughness interaction is believed to be new.

After the onset of self-excited vibrations, the amplitude increases until some nonlinear effects limit any further increase. The nonlinear stiffness of piping loops originated by geometric distortions of the elastic model is considered in the capacity of the nonlinear effect. Asymptotic methods were used for the analysis. Although for the pipes in operating the data about damping and roughness values are questionable, the theoretical results can be derived in a form which allows one to determine all parameters of self-excited vibrations caused by roughness with an accuracy sufficient for engineering practice.

Laminar Friction in the Rough Pipe

If Reynolds numbers are large, the flow in pipes conveying fluid may be treated as a turbulent flow with a laminar sublayer. The shearing stress consists of a laminar (viscous) contribution τ_v and a turbulent contribution τ_t.

$$\tau_o = \tau_v + \tau_t$$

τ_o denotes the shearing stress at the wall.

The viscous contribution in analogy with Stokes's law for laminar flow is introduced as

$$\tau_v = \mu \partial u / \partial z$$

Here μ, u, and z denote viscosity, fluid velocity and distance from the wall, respectively.

For the explanation and analysis of the self-excited vibration mechanism, it is important to conjecture that feasible axial motions of the pipe with relatively small velocities do not have any influence on the velocity profile in the turbulent layer. All interactions occur within the laminar sublayer, and an analysis of the laminar stress behavior during transition from a hydraulically smooth regime to a completely rough regime [19] may ascertain the assumed analogy with the belt friction system.

The velocity distribution in the laminar sublayer is linear to a good degree of approximation [19] and, therefore

$$\tau_v(z=0) \cong \tau_v(z=\delta) = \mu u(z=\delta)/\delta \qquad (1)$$

Here δ denotes a thickness of the laminar boundary sublayer.

The universal logarithmic velocity distribution from Prandtl's theory for large Reynolds numbers agrees well with experimental data within the turbulent boundary layer and

Nomenclature

τ_v = laminar (viscous) contribution
τ_t = turbulent contribution
τ_0 = shearing stress at the wall
μ = viscosity
δ = thickness of the laminar boundary sublayer
ν = kinematic viscosity of the conveying fluid
ρ = fluid density
ψ = $2.5 \ln \delta/k_s - B$
η = $v_* k_s/\nu$
π = 3.14. . .
ϕ = angle
ϵ = P/EI
ξ = s/l
λ = $12 v_B EI/l^3$
ζ = $l^4/7.984(EI)^3$
Σ = symbol of summation
χ = n_Σ/m
ω_0^2 = K_Σ/m
ω = frequency
Θ = phase angle
$\sigma(A)$ = $\int_o^{2\pi} f(\dot{x}_I)\dot{x}_I dt$
α = small, positive parameter
β = frequency of the stationary random excitation in the neighborhood of
γ = phase of excitation

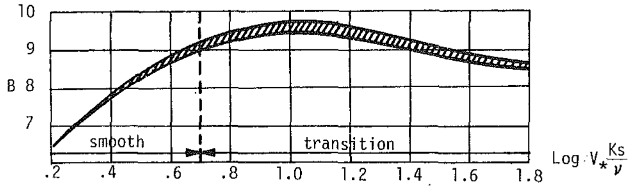


Fig. 2 Experimental spread for the roughness function B in terms of $V_* K_s/\nu$ for Nikuradse's sand roughness (from Schlichting [19])

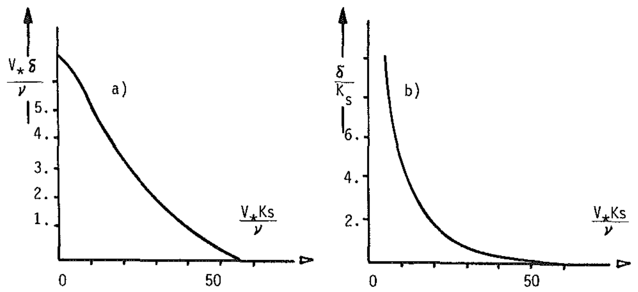


Fig. 3 Effect of wall roughness on δ according to Nikuradse's experiments: (*a*) From Rotta [20]; (*b*) curve rebuilt from Rotta's plot

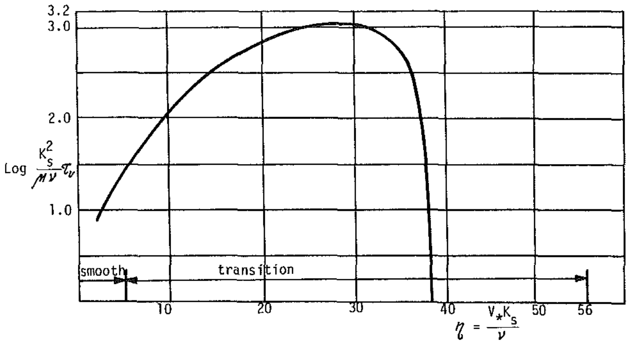


Fig. 4 Diagram of the viscous shearing stress as a function of the roughness Reynolds number

can be used to compute the velocity $u(z=\delta)$ at the top of the laminar boundary sublayer. Schlichting [19] developed, depending on the nature of the particular roughness, the logarithmic velocity distribution which is represented in the following form

$$u = v_*(2.5 \ln z/k_s + B) \tag{2}$$

where

k_s = sand roughness height (average height of protrusions)
$B = B(\eta)$ = function of the roughness Reynolds number, $\eta = v_* k_s/\nu$
ν = kinematic viscosity of the conveying fluid
v_* = friction velocity at the wall, $v_* = (\tau_o/\rho)^{1/2}$, ($\rho$- fluid density).

Equation (1) combined with equation (2) gives

$$\tau_v = \frac{\mu v_*}{\delta}(2.5 \ln\delta/k_s + B) \tag{3}$$

Equation (3) can also be written in the following form

$$k_s^2\tau_v/\mu\nu = k_s\eta\psi/\delta \tag{4}$$

where

$$\psi = 2.5 \ln\delta/k_s + B$$

B and δ/k_s are functions of the roughness Reynolds numbers. The function B was plotted by Schlichting [19] in accordance with Nikuradse's experiments (Fig. 2).

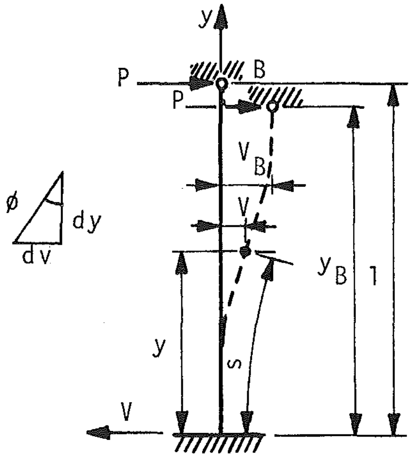


Fig. 5 Model of the piping loop's leg

The effect of wall roughness on the thickness of the boundary sublayer was also deduced and plotted by Rotta [20] from Nikuradse's experiments on flow through circular pipes (Fig. 3(*a*)). For convenient computation this curve was rebuilt as shown in Fig. 3(*b*). Equation (4) was plotted as a function of the Reynolds number η in semilogarithmic scale (Fig. 4). As can be seen, the "crisis of the laminar friction" occurs when $\eta \cong 29$, i.e., in the transition region. The slope of the viscous shearing stress takes varying negative values in the range $\eta = 29$ to 38, and the onset of the self-excited vibration may be expected in this range. The viscous shearing force F_s induced by the fluid flow may be defined as

$$F_s = \tau_v S_p$$

where

S_p = active internal pipe surface, $S_p = 2\pi R l_3$ (Fig. 1)
R = internal pipe radius

For the laminar boundary sublayer, the friction velocity can be expressed in the form of [19]

$$v_* = 2(\nu\bar{u}/R)^{1/2} \tag{5}$$

where $\bar{u}$ denotes a mean velocity of the flow.

Using equation (5) and $\eta = 29$, the threshold flow velocity $\bar{u}_t$ corresponding to "crisis of laminar friction" can be found as

$$\bar{u}_t = 210 R\nu/k_s^2 \tag{6}$$

The conformable critical Reynolds number is $\text{Re} = 420R^2/k_s^2$.

All of the aforementioned equations and relationships are based on the roughness used by Nikuradse with sand of maximum density. There is a correlation between the real roughness of commercial pipes and the equivalent sand roughness, for instance, that was plotted by Moody [21] and other researchers. Apparently, these experimental results are related to pipes used for industrial processes without intensive corrosion and deposition of chemical components at the pipe walls.

As can be seen from Fig. 4, within the region of the decaying laminar friction, equation (3) can be approximated for engineering applications by the following

$$\tau_v = N\cos w \tag{7}$$

where

$$N = 1{,}100\mu\nu/k_s^2, \quad w = 0.1745(\eta - 29)$$

The approximation in form of equation (7) is in tolerable agreement with equation (3) in the range $\eta = 29$ to 38, and is convenient for further analysis.

Equation (7) will be used as a characterization of self-excited vibrations. In view of the fact that apparently neither the logarithmic velocity distribution nor any other simple relationship can adequately describe velocity in the immediate neighborhood of the wall ($\delta < k_s$, $z \geq \delta$) some deviations from equations (3) and (7) may be expected and further experimental and theoretical work for more accurate description of the laminar friction behavior is required. However, it seems questionable that reliable quantitative results can be achieved. The general obstacle is that due to irregularities of the wall the viscous flow in the hollows between protrusions is not predictable. Nevertheless, for the complete rough regime, " . . . the viscous sublayer practically ceases to exist, while the flow in immediate proximity to the wall consists entirely of eddies, generated by the flow around the individual protrusions" [22]. Thus, equation (7) reflects qualitatively the existence of decaying laminar friction and can be used for the explanation and understanding of the axial self-excited vibration mechanism due to roughness in pipes conveying fluid.

Nonlinear Stiffness of the Piping Loop

The piping loop contains two legs elastically deformed during axial motion of the long span (Fig. 1). The stiffness of the legs can be considered using the static model depicted in Fig. 5. A similar problem was solved in [23].

The general differential equation of the beam elastic curve is

$$\frac{d^2v}{dy^2}\left[1+(dv/dy)^2\right]^{-3/2}=M/EI$$

M and EI denote a bending moment and a flexural rigidity of the leg, respectively.

It is seen from Fig. 5 that $M/EI = d\phi/ds$, $dv/ds = \sin\phi$, and one can show that

$$d^2v/ds^2=\left[1-(dv/ds)^2\right]^{1/2} M/EI$$

Taking into account the shift of the point B during deformation of the leg, the bending moment is given by

$$M=-P(y_B-y)+M_B.$$

M_B = constant moment, providing zero angle at point B. The differential of the value $y_B - y$ is

$$ds\cos\phi=ds\left[1-(dv/ds)^2\right]^{1/2}$$

Integrating from s to l

$$y_B-y=\int_s^l\left[1-(dv/ds)^2\right]^{-1/2}ds$$

Finally, the differential equation of the elastic curve may be written as

$$\frac{d^2v}{ds^2}\left[1-(dv/ds)^2\right]^{-1/2}=$$

$$-\frac{P}{EI}\int_s^l\left[1-(dv/ds)^2\right]^{1/2}ds+M_B/EI \tag{8}$$

The nonlinear equation (8) may be solved by a perturbation method. Representing the binomials of the equation (8) as expansion series

$$v''\left[1+0.5(v')^2+0.375(v')^4+\ . \ . \ .\ \right]=$$

$$=-\frac{P}{EI}\int_s^l\left[1-0.5(v')^2-0.125(v')^4-\ . \ . \ .\ \right]ds+M_B/EI \tag{9}$$

Here, $v' = dv/ds$, $v'' = d^2v/ds^2$. The solution of equation (9) may be found in a power series of the parameter ϵ

$$v=\epsilon v_o+\epsilon^2 v_1+\epsilon^3 v_2+\ . \ . \ . \tag{10}$$

The parameter $\epsilon = P/EI$ corresponds to the dimensionless parameter $Pl^2/EI < 1$.

Introducing solution (10) into equation (9) and collecting all terms of not higher than the fourth order in ϵ gives the equation

$$\epsilon v_o''+\epsilon^2 v_1''+\epsilon^3 v_1''+\epsilon^4 v_2''+0.5\epsilon^3 v_o''(v_o')^2+0.5\epsilon^4(v_o')^2 v_1''+$$

$$+\epsilon^4 v_o' v_1' v_o''=-\epsilon\int_s^1\left[1-0.5\epsilon^2(v_o')^2-\epsilon^3 v_o' v_o'\right]ds+\epsilon M_B/EI$$

The constant moment M_B can be represented as

$$M_B=\sum_{i=0}^{\infty}\epsilon^{2i}M_{2i}$$

Equating the terms of equal order in ϵ it can be found that

$$\begin{aligned}
&v_o''=s-l+M_o/P\\
&v_1=0\\
&v_2''+0.5v_o''(v_o')^2=0.5\int_s^l(v_o')^2ds+M_2/P\\
&v_3''+v_o'v_1'v_o''=\int_s^1 v_o'v_1'ds
\end{aligned} \tag{11}$$

Since $v_1 = 0$, it is obvious that $v_3 = 0$. The first equation of the system of equations (11) gives after integration

$$v_o=0.5s^2-ls+M_os/P+C_1 \tag{12}$$

$$v_o=s^3/6-0.5ls^2+0.5M_os^2/P+C_1s+C_2$$

where C_1, C_2 = integration constants.

For boundary conditions: $s=0$, $v_o=0$, $v_o'=0$; $s=1$, $v_o'=0$, the solution (12) may be written in the form of

$$v_o=l^3\xi(0.25-\xi/6),\ v_o'=0.5l^2\xi(\xi-1) \tag{13}$$

$$\xi=s/l$$

The third equation of the system (11) with regard to (13) is as follows:

$$v_2''=-0.125(s-1.5l)\,(s^4-6ls^3+9l^2s^2)$$

$$+0.125\int_s^l(s^4-6ls^3+9l^2s^2)ds+M_2/P$$

After analogous operations the second approximation gives

$$\begin{aligned}
&v_2'=-0.125\xi l^5(0.85-4.5\xi^2+5.25\xi^3-1.8\xi^4+0.2\xi^5)\\
&v_2=-0.0625\xi^2l^7(0.85-2.25\xi^2+2.1\xi^3-0.6\xi^4+0.067\xi^5)
\end{aligned} \tag{14}$$

Thus the total displacement at the point $\xi = 1$ may be obtained by combining the solutions (13) and (14)

$$v_B=Pl^3/12EI-P^3l^7/98.51(EI)^3 \tag{15}$$

The first part of the solution (15) represents a well-known linear solution of the given problem and the second one represents a small nonlinearity of the displacement.

In order to analyze self-excited vibration, the solution (15) has to be expressed as $P = P(v_B)$. It can be done again by means of the perturbation method.

The solution (15) may be presented as

$$P-\zeta P^3=\lambda \tag{16}$$

Here

$$\lambda=12v_BEI/l^3;\ \zeta=l^4/7.984(EI)^3$$

The solution of equation (16) may be found in form of the series

$$P=P_o+\zeta P_1+\zeta^2P_2+\ . \ . \ . \tag{17}$$

Substituting the solution (17) into equation (16)

$$P_o+\zeta P_1+\zeta^2P_2+\ . \ . \ .\ -\zeta(P_o+\zeta P_1+\zeta^2P_2+\ . \ . \ .\ )=\lambda$$

and equating terms with equal power of ζ

$$P_o=\lambda,\ P_1=P_o^3=\lambda^3,\ P_2=3P_o^2P_1=3\lambda^5 \tag{18}$$

Therefore

$$P\cong\lambda+\zeta\lambda^3 \tag{19}$$

It has been noted that λ may be represented as

$$\lambda = Kv_B$$

where K = stiffness of the linear system. Hence solution (19) is

$$P = Kv_B + \zeta K^3 v_B^3$$

and represents the resistance of the single leg corresponding to the displacement v_B at point B. For a given piping loop with two legs

$$P = K_\Sigma + n_\Sigma v_B^3, \tag{20}$$

$$K_\Sigma = K_1 + K_2,\ n_\Sigma = \zeta_1 K_1^3 + \zeta_2 K_2^3$$

where K_1, K_2 = stiffness of the first and second legs, respectively. The coefficient n_Σ in equation (20) is a small quantity.

Self-Excited Vibration of the Piping Loop

The equation of the piping loop motion with respect to the solution (20) may be written in the following form

$$m\ddot{x} + c\dot{x} + K_\Sigma x + n_\Sigma x^3 = F(\tilde{u},\tilde{u}) \tag{21}$$

Here:

c = damping coefficient
m = mass of the piping loop's longer run of length l_3 (Fig. 1)
$\tilde{u}$ = relative sliding velocity, $\tilde{u} = v_* - \dot{x}$
$F(\tilde{u},\tilde{u})$ = induced force

and the dots above a letter indicate the derivatives with respect to time.

At rest, the induced force $F(\tilde{u},\tilde{u})$ equals the head loss and the static displacement x_0 may be derived from the equation

$$K_\Sigma x_o + n_\Sigma x_o^3 = F(\tilde{u}) = S_p \tau_v(v_*) + F_t(\tilde{u}) \tag{22}$$

where

$F(\tilde{u})$ = head loss, calculated, for instance, with the use of the Darcy equation
$F_t(\tilde{u})$ = turbulent component of the head loss

The induced force $F(\tilde{u},\tilde{u})$ may be presented as

$$F(\tilde{u},\tilde{u}) = S_p \tau_v(\tilde{u}) + F_t(\tilde{u}) \tag{23}$$

For small oscillations of the relative sliding velocity $\tilde{u}$ the friction force $S_p\tau_v(\tilde{u})$, induced into the pipe, may be expanded in Taylor's series as

$$S_p\tau_v(\tilde{u}) = S_p\left[\tau_v(v_*) - \frac{d\tau_v(v_*)}{d\tilde{u}_{(\tilde{u}=v_*)}}\dot{x} + \frac{1}{2}\cdot\frac{d^2\tau_v(v_*)}{d\tilde{u}^2_{(\tilde{u}=v_*)}}\dot{x}^2 - \frac{1}{6}\cdot\frac{d_v^3(v_*)}{d\tilde{u}^3_{(\tilde{u}=v_*)}}\dot{x}^3 + \ldots\right] \tag{24}$$

Substituting the expression (24) into equality (23)

$$F(\tilde{u},\tilde{u}) = F(\tilde{u}) - S_p\left[\frac{d\tau_v(v_*)}{d\tilde{u}_{(\tilde{u}=v_*)}}\dot{x} - \frac{1}{2}\cdot\frac{d^2\tau_v(v_*)}{d\tilde{u}^2_{(\tilde{u}=v_*)}}\dot{x}^2 + \frac{1}{6}\cdot\frac{d^3\tau_v(v_*)}{d\tilde{u}^3_{(\tilde{u}=v_*)}}\dot{x}^3 - \ldots\right]$$

Introducing a new variable $x_1 = x - x_0$, taking into account equations (1) and (7) and assuming that $3n_\Sigma x_I x_o(x_I + x_o)$ is a negligible quantity ($x_o < \max x_1$)

$$\ddot{x}_1 + \omega_o^2 x_1 + \chi x_1^3 = Gf(\dot{x}_1) \tag{25}$$

where

$$\omega_o^2 = K_\Sigma/m,\ \chi = n_\Sigma/m,\ G = 0.1745 S_p N k_s/m\nu,$$

$$f(\dot{x}_1) = \dot{x}_1 \sin w - \dot{x}_1 c/mG - 0.0875\dot{x}_1^2 k_s \cos w/\nu - 0.0051\dot{x}_1^3\ k_s^2 \sin w/\nu^2 + \ldots \tag{26}$$

Equation (25) may be solved by different methods: Ritz method, perturbation method, Kryloff and Bogoliuboff's method, etc. In seeking a steady-state solution, the energy method can be used with the same efficiency, but with a simpler procedure. The energy method is based on an assumption that friction forces have no influence on the shape of the mode of vibration and, therefore, the solution of equation (25) may be found in the form of the solution to the Duffing's equation for the hardening system.

$$\ddot{x}_1 + \omega_o^2 x_1 + \chi\, x_1^3 = 0 \tag{27}$$

The solution to the Duffing's equation (27) may be derived by Struble's method [24] in the following form

$$x_1(t) = A\cos(\omega t - \Theta_o) + \frac{\chi A^3}{32\omega_o^2}(1 - 21\chi A^2/32\omega_o^2)\cos 3(\omega t - \Theta_o) - {} - \chi^2 A^5 \cos 5(\omega t - \Theta_o)/1{,}024\omega_o^4 + 0(\epsilon^3) \tag{28}$$

$$\omega = \omega_o(1 + 3\chi A^2/8\omega_o^2 - 15\chi^2 A^4/256\omega_o^4) + 0(\epsilon^3) \tag{29}$$

Θ_o = phase angle.

For a whole period of vibration, the net work done by inertial and elastic forces equals zero and, hence, the net work done by the friction force also equals zero. Therefore

$$\sigma(A) = \int_0^{2\pi/\omega} f(\dot{x}_1)\dot{x}_1\,dt = 0 \tag{30}$$

Substitution of the function $f(\dot{x}_1)$, i.e., equation (26), and $x_1(t) \cong A\cos(\omega t - \Theta_o)$ into equation (30) and subsequent integration give a first approximation for the amplitude of self-excited vibrations

$$A \cong \frac{16.2\nu}{k_s\omega}\left[\frac{\sin w - c/mG}{\sin w}\right]^{1/2} \tag{31}$$

It is self-evident that equation (31) is valid only in the range $\sin w \geq c/mG$.

Since the parameter χ is assumed to be small, the relationship (31) represents the amplitude of vibrations with accuracy sufficient for engineering applications.

The highest feasible amplitude is ($\omega \approx \omega_o$)

$$A \cong 16.2\nu/k_s\omega_o \tag{32}$$

The numerical example can be presented, for instance, for the main steam line (Nebraska Cooper Nuclear Station).

steam: temperature 295°C (586K), pressure $0.816\ .10^6\,\mathrm{kg/m^2}$, corresponding kinematic viscosity $\nu \cong 0.485 \cdot 10^{-6}\ \mathrm{m^2/s}$
pipe: inside diameter 0.61m, $k_s \cong 0.001$m, $\omega_o \cong 0.3$ Hz, $\mathrm{Re}_t \geq 3.9 \cdot 10^7$

Using (32), max $A \cong 4.2$mm. It can be shown that

$$\partial\sigma(A)/\partial A < 0$$

and, hence, in accordance with Liapunov's theorem [25], the regime $x_1(t) = A\cos\omega t$ with amplitude of vibration given by (31) is stable.

Hitherto the flow was considered to be uniform without disturbances whereas the real turbulent motion contains very irregular fluctuations. A concise survey of experimental data pertaining to the measurements of turbulent fluctuations is given in [19]. Although the fluctuations of the pressure and velocity do not exceed several percent, they may have a substantial influence on the behavior of the system.

The irregular fluctuations presented in the form of random, wide-band input are to be superimposed on a given nonlinear self-oscillating system.

Since a system under consideration with single degree of freedom acts as a narrow bandpass filter, the governing equation may be written in the form

$$\ddot{x}_1 + \omega_o^2 x_1 = \alpha q(\dot{x}_1) + Q\cos(\beta t + \gamma) \tag{33}$$

where

α = small, positive parameter
Q = random amplitude of a given harmonic excitation
β = frequency of the stationary random excitation in the neighborhood of ω_o
γ = phase of excitation
q = function specifying the nonlinearity of the system and stipulating the existence of self-excited regimes. In the given case $q = Gf(x_1) - x_1^3$.

The deterministic analysis of equation (33), using Kryloff and Bogoliuboff's procedure, has been made, for example, in [26]. The response curve of this system is typical for a lightly damped, nonlinear system with peculiar hysteresis effects (jump phenomena) which appear with deviations of the frequency of the excitation.

The survey of applications to random excitation of this system is available, for instance, in [27].

The families of response curves may be calculated for different levels of the excitation.

Discussion and Conclusion

A tentative mathematical model of flow-roughness interactions, based on Prandtl's universal velocity distribution and on Nikuradse's experimental data, has been employed for the explanation of self-excited vibrations of piping loops conveying fluid. The existence of the negative slope for the friction characteristic gives every reason to assume an analogy with a belt friction system. A mechanism of self-excited vibrations of the belt friction system is well understood, and a theoretical investigation has been carried out in a similar manner.

Derived approximate closed-form solutions allow us to predict the onset of self-excited vibrations and the steady-state amplitude.

This study shows that the height of protrusions and kinematic viscosity of the conveying fluid play the leading part in the proposed mechanism of self-excited vibrations for a piping loop. In general, the effect of increasing the conveying fluid temperature is to decrease kinematic viscosity and, hence, to decrease the threshold flow velocity $\bar{u}_t$. The effect of the roughness growth during pipe operation is also to decrease $\bar{u}_t$.

Analytical results for steady-state regimes are in good agreement with the observed average amplitudes of oscillations.

An analysis of the system geometric nonlinearity has been undertaken for the explanation of attendant hysteresis effects. An additional excitation caused by turbulent disturbances brings a given nonlinear system to the irregular transition motions (jump phenomena). Due to complexity of the interpretation of experimental data, the nonlinear response of systems such as this is often treated as a linear one contaminated with noise. However, the importance of nonlinear components for the proper mathematical modeling of real systems is realized, and the method recently proposed [28] can apparently be useful for recognizing the structural nonlinearities.

The important role of the self-exciting mechanism is, also, decreasing structural damping, to extend the resonant peak further up the free vibration backbone, equation (29).

Thus the anticipated responses of the system under consideration are: steady-state, self-excited vibrations with relatively small stable amplitudes for the flow without considerable fluctuations, and unstable forced vibrations accompanied by jump phenomena due to interaction of the self-oscillating system with random excitation for the flow with severe turbulent disturbances. The peak amplitudes of the latter ones can be much greater than amplitudes of self-excited vibrations.

In spite of relatively small amplitudes, the steady-state self-excited vibrations can cause malfunctions and fatigue failure of piping systems and attached equipment.

A given tentative model is not apparently the only one which describes the dynamic behavior of piping loops conveying fluid. Similar responses can also be simulated based on other suitable models.

References

1 Chen, S. S., "Parallel Flow-Induced Vibrations and Instabilities of Cylindrical Structures," *Shock and Vibration Digest*, Vol. 6, No. 10, 1974, pp. 2-12.

2 Chen, S. S., "Flow-Induced Vibrations of Circular Cylindrical Structures, Part I: Stationary Fluids and Parallel Flows," *Shock and Vibration Digest*, Vol. 9, No. 10, 1977, pp. 25-38.

3 Burgreen, D., Byrnes, J. J. and Benforado, D. M., "Vibration of Rods Induced by Water in Parallel Flow," ASME *Transactions*, Vol. 80, No. 5, 1958, pp. 991-1003.

4 Benjamin, T. B., "Dynamics of a System of Articulated Pipes Conveying Fluid: 1. Theory; 2. Experiment," *Proc. Royal Soc.* (London), 261 (Ser. A), 1961, pp. 457-499.

5 Quinn, E. P., "Vibration of Fuel Rods in Parallel Flow," U.S. AEC Rept. GEAP-4966, General Electric Co., 1965.

6 Quinn, E. P., "Vibration of SEFOR Fuel Rods in Parallel Flow," U.S. AEC Rept. GEAP-4966, General Electric Co., 1965.

7 Hopkins, G. R., "Stability of Fluid Conveying Tubes with Periodic Perturbations," Ph.D. thesis, University of Alabama, 1969.

8 Chen, Y. N., "Flow-Induced Vibrations in Tube Bundle Heat Exchangers with Cross and Parallel Flow: Part 1, Parallel Flow," *Proceedings of Flow-Induced Vibration in Heat Exchangers*, ASME, 1970, pp. 57-66.

9 Chen, S. S., "Dynamic Stability of Tubes Conveying Fluid," ASCE *J. Engr. Mech. Div.*, Vol. 97 (EMS), Oct. 1971, pp. 1469-1485.

10 Ginsberg, J. H., "The Dynamic Stability of a Pipe Conveying Pulsatile Flow," *Intl. J. Engr. Sci.*, Vol. 11, 1973, pp. 1013-1024.

11 Paidoussis, M. P., and Sundararajan, C., "Parametric and Combination Resonances of a Pipe Conveying Pulsating Fluid," ASME Journal of Applied Mechanics, Vol. 42, No. 4, 1975, pp. 780-784.

12 Paidoussis, M. P., "Stability of Flexible Slender Cylinders in Pulsatile Axial Flow," *J. Sound and Vib.*, Vol. 42, No. 1, 1975, pp. 1-11.

13 Paidoussis, M. P., and Issid, N. T., "Experiments in Parametric Resonance of Pipes Containing Pulsatile Flow," ASME Journal of Applied Mechanics, Vol. 43, No. 2, June, 1976, pp. 198-204.

14 Paidoussis, M. P., "Dynamics of Flexible Slender Cylinders in Axial Flow. Part 1: Theory," *Journ. of Fluid Mech.*, Vol. 26, 1966, pp. 717-751.

15 Paidoussis, M. P., "Dynamics of Cylindrical Structures Subjected to Axial Flow," *Journ. of Sound and Vib.*, Vol. 29, No. 3, 1973, pp. 365-385.

16 Boussinesq, J., *Theorie de l'ecoulement tourbillant*, Mem. pres. Acad. Sci. XXIII, 46, Paris, 1877.

17 Nikuradse, J., "Stromungsgesetze in rauhen Rohren," *Forsch. Arb. Ing.-Wes*, No. 361, 1933, pp. 306-332.

18 Bauer, B., and Galavics, G., *Expermentelle und theoretische Untersuchungen uber die Rohrreibung von Heizwasserleitungen*, Mitt. d. Fernheizkraftwerkes d. ETH Zurich 1936.

19 Schlichting, H., *Boundary-Layer Theory* (6th ed.), McGraw-Hill, New York, 1968.

20 Rotta, J., "Das in Wandnähegültige Geschwindigkeitsgesetz turbulenter Strömungen," *Ingr. Arch.*, Vol. 18, 1950, pp. 277-280.

21 Moody, L. F., "Friction Factors for Pipe Flow," ASME *Transactions*, Vol. 66, 1944, pp. 671-678.

22 Monin, A. S., and Yaglom, A. M., *Statistical Fluid Mechanics: Mechanics of Turbulence*, Vol. 1, ch. 3, p. 289. The MIT Press, Cambridge, Mass., 1971.

23 Malikov, G. F., Shneyderman, A. L., and Shulemovich, A. M., *Calculations of the Elastic Tensiometric Elements*, Publishing House "Machine Building," USSR, Moscow, 1964.

24 Nayfen Ali Hasan, *Perturbation Methods*, John Wiley & Sons, New York, 1973.

25 Liapunov, A., *Probleme general de la stabilite du movement*, Annals Math. Studies No. 17, Princeton, 1947.

26 Bogoliuboff, N. N., and Mitropolsky, U. A., *Asymptotic Methods in Theory of Nonlinear Vibrations*, Publishing House "Nauka," USSR, Moscow, 1974.

27 Crandall, S. H., *Random Excitation of Nonlinear Systems*, Random Vibration, Vol. 2, The MIT Press, Cambridge, Mass., 1963.

28 Kirshenboim, J., and Ewins, D. J., "A Method for Recognizing Structural Nonlinearities in Steady-State Harmonic Testing," ASME *Journal of Vib., Acoust., Stress and Reliab. in Design*, Vol. 106, 1984, pp. 49-52.

J. T. Tozzi
Commander—USCG Headquarters

C. H. von Kerczek
Adjunct Associate Professor,
Mechanical Engineering Department,
The Catholic University of America,
Washington, D.C. 20064

The Stability of Oscillatory Hagen-Poiseuille Flow

The linear stability theory of the nonzero mean, sinusoidally oscillating flow in a tube of circular cross section is examined. It is found that the relevant axisymmetric disturbances in the oscillatory flow are more stable (i.e., have larger decay rates) than the axisymmetric disturbances of the mean flow alone. This result holds for values of the cross-sectional average oscillation velocity amplitude at least as large as seven-tenths the average mean-flow velocity amplitude. Although the instantaneous velocity profile contains generalized inflection rings for a substantial portion of the oscillation period, the disturbances do not become instantaneously unstable at any time, even for very low frequency oscillations.

1 Introduction

The flow in a tube of circular cross section which is driven by the combination of a steady and a sinusoidally time-varying axial pressure gradient is called oscillatory Hagen-Poiseuille flow. Such a flow has many important applications. Hence it seems important to examine its stability characteristics.

It is well known that steady Hagen-Poiseuille flow is stable to infinitesimal disturbances for all values of the Reynolds number (Salwen and Grosch [1]). It would be interesting to know what the effects of temporal oscillations are on these stability characteristics. The simpler case of plane Poiseuille flow has a critical value of Reynolds number above which the flow is unstable. It was shown by von Kerczek [2] that the sinusoidal oscillation of this flow stabilizes it for a range of values of dimensionless frequency Ω. This range of values of Ω is less than four times and greater than one-tenth the value of the frequency ω_1 of the unstable disturbance mode of the underlying mean flow. However, for very low-frequency oscillations ($\Omega < \omega_1/10$), plane Poiseuille flow is slightly destabilized. In contrast to plane Poiseuille flow, it was shown in reference [3] that at low values of Reynolds number plane Couette flow is rendered less stable, although not destabilized, by superimposed small-amplitude flow oscillations driven by the channel walls. Plane Couette flow is similar to Hagen-Poiseuille flow in the sense that it is also stable to infinitesimal disturbances at all values of the Reynolds number (Davey [4]).

The superimposed temporal variations of the pressure gradient in oscillatory Hagen-Poiseuille flow causes the instantaneous velocity profile to develop inflection rings at some radial location within the tube during most of the oscillation cycle. In axisymmetric flows, generalized inflection rings (for satisfying the Rayleigh criterion) occur when

$$\frac{\partial^2 U}{\partial r^2} - \frac{1}{r}\frac{\partial U}{\partial r} = 0 \qquad (1)$$

where U is the basic flow velocity and r is the radial coordinate. Intuitions developed from the linear stability theory of steady flows suggests that the oscillatory flow is highly unstable at least during the time interval in which the inflection rings are present. In particular, this is expected for the cases of low frequency imposed oscillations.

Four experimental investigations (Gilbrech and Combs [5], Ramaprian and Tu [6], Sarpkaya [7], and Yellin [8]) to measure the effects of low frequency flow oscillations on the transition characteristics of Hagen-Poiseuille flow have shown that these oscillations tend to suppress the development and growth of turbulent spots. The linear stability theory has no known relationship to the mechanics of transition to turbulence because it predicts that steady Hagen-Poiseuille flow is completely stable. However, the experiments do suggest that the oscillations enhance Hagen-Poiseuille flow's stabilty. A measure of this enhancement of stability is defined here as the relative change σ in the decay rates of disturbances due to the imposed flow oscillation.

It is worthwhile to note that von Kerczek and Davis [9] show that the Stokes layer by itself is not unstable. Similarly, the linear stability theory shows purely oscillatory pipe flow to be completely stable (Yang and Yih [10]), although such flows do exhibit in experiments a kind of intermittent transition between laminar and turbulent flow (Hino et al. [11]).

The research described in this paper deals with the stability of low-frequency, sinusoidally modulated Hagen-Poiseuille flow. Only axisymmetric disturbances are analyzed. The full time-dependent linear stability theory is analyzed by a combination of numerical and high order perturbation methods. The quasi-steady approximation to the full stability theory is also examined numerically. Sections 2–5 contain, respectively, the formulation of this stability problem; the methods of solution of this problem; a discussion of the computational results with a view toward interpreting 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 two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jaunary, 1984; final revision, December, 1984.

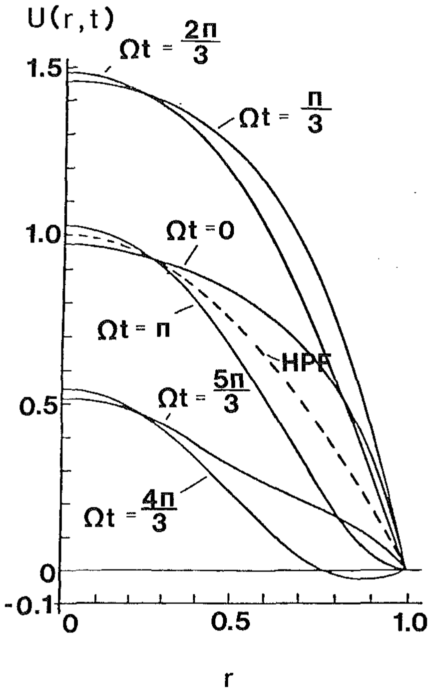


Fig. 1 Basic flow velocity profiles. HPF denotes the steady-flow profile. For the unsteady profiles, β = 5.8, Δ = 0.5

experimental results of references [5–8] and some concluding remarks.

2 Formulation

The flow of a homogeneous incompressible fluid through an infinitely long tube of circular cross section is considered. This flow is forced by the combined steady and unsteady axial pressure gradient

$$\frac{\partial P_0}{\partial z} = -\rho(P_0 + Q_0 \cos\omega t) \tag{2}$$

where ρ is the fluid density, z is the axial coordinate, t is time, and ω is the angular frequency. The flow is described in cylindrical coordinates with r and θ denoting radial and azimuthal coordinates respectively. The radius of the tube is denoted by a and ν is the kinematic viscosity of the fluid. The problem is made dimensionless by the length scale a, the velocity scale $U_0 = P_0\, a^2/4\nu$, and the time scale a/U_0. Henceforth, all variables are assumed to be appropriately scaled.

The exact solution of the Navier-Stokes equation describing oscillatory Hagen-Poiseuille flow is

$$\mathbf{V} = (0, 0, U(r,t)) \tag{3a}$$

$$P = -z(1 + \Lambda\cos\Omega t) \tag{3b}$$

where

$$U(r,t) = U_s(r) + \Delta U_p(r,t) \tag{4a}$$

$$U_s(r) = 1 - r^2 \tag{4b}$$

$$U_p(r,t) = \mathrm{Re}\left\{ i\left[\frac{J_0(i^{3/2}\beta r)}{J_0(i^{3/2}\beta)} - 1\right] e^{i\Omega t}\right\} \tag{4c}$$

$$\Lambda = Q_0/P_0 \tag{4d}$$

$$\Omega = a\omega/U_0 \tag{4e}$$

$$\beta = a(\omega/\nu)^{1/2} \tag{4f}$$

where

$$\Delta = 4\Lambda/\beta^2 \tag{4g}$$

The function $J_0(x)$ is the Bessel function of the zeroth order (Abramowitz and Stegun [12]). The parameter β appearing in the velocity U_p is of primary importance. It can be interpreted as either a dimensionless frequency which is independent of the mean-flow time scale or as the geometric parameter that measures the ratio of tube radius to the Stokes layer thickness, $\delta_s = (\nu/\omega)^{1/2}$. The parameter Δ is the most convenient measure of oscillation amplitude. It is approximately equal to the ratio of the centerline oscillatory velocity amplitude to mean centerline velocity and is almost always numerically small. The relationship between Δ and the ratio, S, (used by Sarpkaya [7]) of the cross-sectional average oscillation velocity amplitude to the average mean-flow velocity is given by

$$S = 2\Delta\left[\left(1 - \frac{2C}{\beta}\right)^2 + \left(\frac{2D}{\beta}\right)^2\right]^{1/2} \tag{5a}$$

$$C = \frac{\mathrm{ber}(\beta)\mathrm{bei}'(\beta) - \mathrm{bei}(\beta)\mathrm{ber}'(\beta)}{\mathrm{ber}^2(\beta) + \mathrm{bei}^2(\beta)} \tag{5b}$$

and

$$D = \frac{\mathrm{bei}(\beta)\mathrm{bei}'(\beta) + \mathrm{ber}(\beta)\mathrm{ber}'(\beta)}{\mathrm{ber}^2(\beta) + \mathrm{bei}^2(\beta)} \tag{5c}$$

where ber(β) and bei(β) are Kelvin functions [12].

Figure 1 shows the distribution of the velocity $U(r,t)$ at various instants of time in one oscillation cycle. In this figure, Δ is 0.5 and β has the value of 5.8 (fairly low frequency). For this value of β, inflection rings occur somewhere within the tube for the time intervals

$$\begin{aligned} &0 \le \Omega t \le 0.55 && \text{(A)} \\ &1.296 \le \Omega t \le 3.691 && \text{(B)} \\ &4.438 \le \Omega t \le 6.283 && \text{(C)} \end{aligned} \tag{6}$$

during one cycle ($0 \le \Omega t \le 2\pi$) of the oscillation. One objective of this investigation is to determine the relevance to the stability characteristics of the existence of inflection rings over such a substantial portion of the oscillation cycle.

The basic oscillatory flow (3) is disturbed by infinitesimal axisymmetric disturbances of axial wave number α. The resulting flow is assumed to be governed by the linearized Navier-Stokes and continuity equations. Such disturbances may be described by the Stokes stream function $\phi(r,t)$ exp($i\alpha z$) in such a way that the radial and axial components of the disturbance velocity are obtained by the formulas

$$u = \mathrm{Re}\left\{-\frac{i\alpha}{r}\phi e^{i\alpha z}\right\} \tag{7a}$$

$$v = \mathrm{Re}\left\{\frac{1}{r}\frac{\partial\phi}{\partial r}e^{i\alpha z}\right\} \tag{7b}$$

Accordingly, the linearized Navaier-Stokes equations can be reduced to the single governing disturbance equation (see Tozzi [13])

$$\mathcal{L}\frac{\partial\phi}{\partial t} = \frac{1}{R}\mathcal{L}^2\phi - i\alpha U\mathcal{L}\phi + i\alpha\{\phi\mathcal{L}U + \alpha^2 U\phi\} \tag{8a}$$

with the boundary conditions

$$\phi(0,t) = \frac{\partial\phi}{\partial r}(0,t) = 0 \tag{8b}$$

$$\phi(1,t) = \frac{\partial\phi}{\partial r}(1,t) = 0 \tag{8c}$$

where

$$\mathcal{L} = \frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} - \alpha^2 \tag{9}$$

and where R = $U_0 a/\nu$ is the mean-flow Reynolds number.

Since the basic flow $U(r,t)$ is $2\pi/\Omega$ periodic in time, the

Floquet Theorem (Coddington and Levinson [14]) is invoked to calculate solutions of the form

$$\phi(r,t)=\eta(r,t)e^{\lambda t} \qquad (10)$$

where, $\eta(r,t+2\pi/\Omega)=\eta(r,t)$. The stability or instability of the basic flow is governed by the sign of the real part of the Floquet exponent λ. It is worthwhile to observe that for steady basic flows the solution (10) is also valid. In such a case, $\eta(r,t)$ is constant in time (hence trivially satisfying the periodicity condition) and λ is the usual stability exponent. The next section deals with the procedure for obtaining the solution (10) of the equations (8).

3 Methods of Solution

The stability problem (8) is solved in two ways. The first way obtains the time-dependent solution (10) by a Chebyshev spectral expansion in terms of the independent variable r and a high order series expansion in terms of the parameter Δ. The second way of solving equations (8) makes use of the quasi-steady approximation to the equations and can be obtained as a byproduct of the full time-dependent problem.

In the full time-dependent problem, the function $\phi(r,t)$ is expanded in a Chebyshev polynomial series as follows:

$$\phi(r,t)=\sum_{n=1}^{N}a_n(t)T_{n-1}(\tilde{r}) \qquad (11)$$

where

$$\tilde{r}=2r-1 \qquad (12a)$$

and

$$T_n(\tilde{r})=\cos(n\cos^{-1}\tilde{r}) \qquad (12b)$$

The τ-method, as described by Gottlieb and Orszag [15], is used to convert equations (8) into the system of ordinary differential equations

$$\mathbf{Q}\cdot\frac{d\mathbf{a}}{dt}=\frac{1}{\mathrm{R}}\mathbf{P}\cdot\mathbf{a}+i\alpha\mathbf{V}\cdot\mathbf{a}+i\alpha\Delta\mathbf{W}(t)\cdot\mathbf{a} \qquad (13)$$

where

$$\mathbf{a}(t)=(a_1,a_2,\ldots,a_{N'})^{+} \qquad (14)$$

is the vector of Chebyshev coefficients and $\mathbf{Q},\mathbf{P},\mathbf{V}$, and $\mathbf{W}$ are $N'\times N'$ matrices, respectively representing the operators $\mathcal{L}$, $\mathcal{L}^2$, $-U_s\mathcal{L}+\mathcal{L}U_s+\alpha^2U_s$ and $-U_p\mathcal{L}+\mathcal{L}U_p+\alpha^2U_p$ together with boundary conditions (8b and c), on the Chebyshev basis T_n, $n=0,\ldots,N'-1$, $N'=N-4$. The matrix $\mathbf{W}(t)$ is $2\pi/\Omega$ time-periodic.

By invoking the Floquet Theorem, the vector $a(t)$ has the form

$$\mathbf{a}(t)=\mathbf{b}(t)e^{\lambda t} \qquad (15)$$

where $\mathbf{b}(t+2\pi/\Omega)=\mathbf{b}(t)$ is the spectral expansion coefficient vector of the function $\eta(r,t)$.

For the value of $\Delta=0$, equation (15) is the solution of the steady Hagen–Poiseuille flow stability problem in which $\mathbf{b}(t)=\mathbf{b}_0$, a constant vector, and $\lambda=\lambda_0$ is the usual steady-flow stability exponent. It is important to note that the steady Hagen-Poiseuille flow stability problem has for each value of α an infinite number of disturbance modes each of which is described by a mode vector $\bar{b}_0$ and characteristic value λ_0. The eigenvalues λ_0 of the steady problem can be ordered according to the value of their real parts. This ordering results in a descending sequence that accumulates on $-\infty$. Only the first few values of the λ_0's are of interest. These values are called the principal modes and have real parts clustered near zero. The finite Chebyshev series expansion (11) produces N' approximations to these characteristic modes, but only a number $N^*<<N'$ are accurate approximations of the principal modes. The unsteady flow stability problem has the same features.

A solution of problem (13) for $\Delta\neq0$ is obtained by a perturbation expansion about one of the steady-flow principal modes as follows:

$$\mathbf{b}(t)=\mathbf{b}_0+\Delta\mathbf{b}_1(t)+\Delta^2\mathbf{b}_2(t)+\ldots\ldots \qquad (16a)$$

$$\lambda=\lambda_0+\Delta\lambda_1+\Delta^2\lambda_2+\ldots \qquad (16b)$$

The method described by von Kerczek [2] (see also Tozzi [13] for a very detailed treatment), is used to calculate as many as 30 to 40 terms of the series (16). The solution $\mathbf{a}(t)$ is the limit of the sequence of partial sums of the series (16). This limit is determined with the aid of the Shanks transformation (Shanks [16]).

The quasi-steady approximation of the solution (15) is valid in those cases in which the frequency ratio $f=\omega_0/\Omega$, $\omega_0=\mathrm{Im}(\lambda_0)$, has a very large value. In such cases, the slow time (parametric time) t_0 corresponds to fixed instants in the oscillation time period $[t_0, t_0+2\pi/\Omega]$. The fast time τ ranges over the characteristic time period $2\pi/\omega_0$ of the disturbance. Hence the quasi-steady solution $\phi(r,\tau,t_0)$ holds in the time interval $[t_0-2\epsilon\pi/\omega_0, t_0-2\epsilon\pi/\omega_0]$ for some positive value of ϵ.

The quasi-steady approximate solution can be obtained from the governing equation (13) by holding time in the matrix $\mathbf{W}$ fixed at the value t_0 and solving the resulting equation as a steady stability problem. This procedure yields

$$\mathbf{a}(\tau,t_0)=\mathbf{b}_0(t_0)e^{\lambda_0(t_0)\tau} \qquad (17)$$

The eigenvector $\mathbf{b}_0(t_0)$ and eigenvalue $\lambda_0(t_0)$ are obtained numerically by solving the algebraic eigenvalue problem

$$\left[-\lambda_0(t_0)\mathbf{Q}+\frac{1}{\mathrm{R}}\mathbf{P}+i\alpha\mathbf{V}+i\alpha\Delta\mathbf{W}(t_0)\right]\cdot\mathbf{a}=0 \qquad (18)$$

for each instant $t_0\epsilon[0, 2\pi/\Omega]$. (EISPACK routines were used. Smith et al. [17].)

4 Computational Results

There is no clear relationship between the disturbance wave number α and the observed laminar to turbulent transition because steady Hagen-Poiseuille flow is stable to infinitesimal disturbances. Furthermore, the decay rates of disturbances decrease monotonically with decreasing values of α for fixed values of the Reynolds number R (Davey and Drazin [18]). Thus, it is not clear which values of α to choose for examining the effects of the flow oscillation on the stability of Hagen-Poiseuille flow. In plane Poiseuille flow, the critical value α_c of α (made dimensionless by the half-channel width) is about 1.0 (reference [2]) and in the Blasius boundary layer α_c is about 0.35 (made dimensionless by the displacement thickness) (Schlichting [19]). Hence, the range of values $0.35<\alpha<1.0$ is probably the most relevant to examine the effects of flow oscillation on stability. In this study, attention is focused on the value $\alpha=0.5$; but other values of α were examined also.

For a fixed value of α the infinite sequence of decaying disturbance modes of steady Hagen-Poiseuille flow was ordered in a descending sequence according to increasing values of the decay rate. The mode with the smallest decay rate is mode one. The first few modes can be classified as wall and center modes. A center mode has a value of phase velocity near 1.0 (the dimensionless centerline velocity of the mean flow). A wall mode has a value of the phase velocity in the range between 0.0 and about 0.5. This classification is based mainly on the ideas of plane Poiseuille flow which does have a critical disturbance mode whose phase velocity is associated with the mean flow velocity at the critial layer (Lin [20]). The mechanism that transfers mean-flow energy into disturbance energy is active at the critical layer where mean-flow and phase velocities are equal.

The wall modes have lower frequency than the center modes although the wall modes are more heavily damped than the

Table 1 The perturbation coefficients $\mathrm{Re}(\lambda_j)$ of series (16*b*). Here R = 3000, $\beta = 5.8$, $\alpha = 0.5$, N = 40, and the number of terms in series (16) = 40.

j	$\mathrm{Re}(\lambda_j)$
0	−0.68519E−1
2	−0.28563E−1
4	−0.27953E−1
6	−0.14584E−1
8	+0.86816E−1
10	+0.50027E00
12	+0.19262E+1
14	+0.21448E+1
16	−0.24835E+2
18	−0.12196E+3
20	−0.30939E+3
22	+0.63828E+2
24	−0.45441E+5
26	+0.47130E+7
28	−0.39114E+9
30	−0.23008E+11
32	−0.30155E+13
34	+0.56562E+15
36	−0.36862E+17
38	−0.79170E+18
40	+0.10745E+21

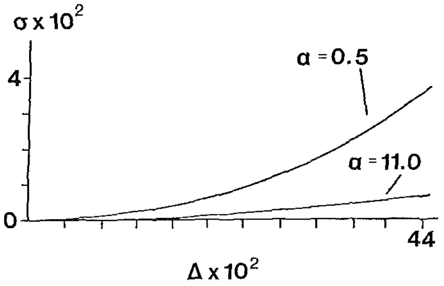


Fig. 2 Relative change σ of least stable center mode versus Δ for R = 3000, β = 5.8

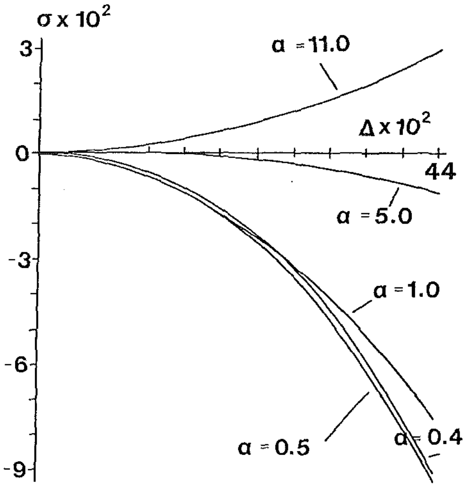


Fig. 3 Relative change σ of least stable wall mode versus Δ for R = 3000, β = 5.8

center modes. However, it is expected that the wall modes are more affected by the flow oscillation because their low frequency places them closer to the oscillatory vorticity field which is confined effectively within the Stokes layer. Hence particular attention was focused on the effects of flow oscillation on the stability characteristics of the wall modes.

It is not possible to examine the stability characteristics over the entire range of values of R and β. The experimental results of Sarpkaya [7] indicate that the most interesting effects of flow oscillation occur for $2000 < \mathrm{R} < 6000$ and for values of β in the range $4 < \beta < 8$. The case of $\mathrm{R} = 3000$ and $\beta = 5.8$ is typical of the behavior in these ranges and, thus, was chosen for detailed investigation.

The primary results of the solution of the full time-dependent equations (8) are the values of the perturbation coefficients λ_j, $j = 0,2,4, \ldots$, of series (16*b*) for the complex growth rate λ of the disturbance. All the odd coefficients of this series are zero because the basic coefficient λ_0 is a simple eigenvalue of the underlying steady flow stability problem about which the perturbation series (16) is calculated and the flow modulation is sinusoidal. A sample computational result for series (16*b*) is given in Table 1 for the values $\mathrm{R} = 3000$, $\beta = 5.8$, and $\alpha = 0.5$. This case was computed by taking $N = 40$ terms of the Chebyshev expansion (11) (see Tozzi [13] for details).

Series (16*b*) is summed using the Shanks transformation (Shanks [16]) on the partial sums. This procedure yields converged values of λ for values of Δ over three times as large as the radius of convergence of series (16). For example the radius of convergence of the series (16) for the case given in Table 1 was experimentally determined to be about $\Delta = 0.12$. However, accurate values of λ for this case were obtained for values of Δ up to 0.44 by use of the Shanks transformation. All calculations for the time-dependent stability problem were made in this fashion.

The complex growth (decay) rate was computed for the least stable center and wall modes for the cases $\mathrm{R} = 3000$, $\beta = 5.8$, and a number of values of α. These results are shown plotted as a function of Δ in Figs. 2 and 3. In these figures, $\lambda_r = \mathrm{Re}(\lambda)$ and $\lambda_{or} = \mathrm{Re}(\lambda_0)$ are the decay rates; and σ is the relative change of the decay rate, defined by

$$\sigma = \frac{\lambda_{or} - \lambda_r}{\lambda_{or}} \tag{19}$$

Figures 2 and 3 show that the relative changes in the decay rates are all extremely small so that the mean effect of oscillation is small. The oscillation tends to make the wall modes more stable and the center modes slighlty less stable. However, the overall effect is surprisingly small. Some large-wave-number results (for $\alpha = 5$ and 11) were included only because these have some relevance to the nonlinear stability studies of Davey and Nguyen [21]. Their nonlinear theory for the steady base flow predicts that disturbances with wave number α about equal to 10 will be nonlinearly unstable. Yet, the linear theory shows that these disturbances are almost completely unaffected by base flow oscillations.

It is very significant that the imposed oscillation frequency Ω is very small compared to the disturbance frequency ω_0 for all the disturbances examined in Figs. 2 and 3. For the cases $0.4 < \alpha < 1.0$ the frequency ratio $f = \omega_0/\Omega$ is in the range 20 to 90. For the case $\alpha = 11$, the value of f is over 150. Thus, the quasi-steady theory should be valid; and one would think that, for the extensive time in which the inflection ring criterion is satisfied, inequalities (6), this flow might become unstable. In this regard the instantaneous growth rate G, defined by

$$G = \frac{1}{\|\mathbf{a}\|} \frac{d\|\mathbf{a}\|}{dt} \tag{20}$$

where $\|\mathbf{a}\|$ is the Euclidean norm of the Chebyshev expansion coefficients (14), was examined. Figures 4 and 5 show the variations of G over one oscillation cycle for the case $\mathrm{R} = 3000$, $\beta = 5.8$, $\alpha = 0.5$, and various values of Δ. These figures show that neither the first wall nor the first center mode ever becomes instantaneously unstable, even though in both cases the frequency ratio f is large. For the wall mode $f = 23$ and for the center mode $f = 41$.

The graphs in Figs. 4 and 5 were computed both by the full time-dependent stability theory solution with equation (20) and by the quasi-steady approximation ($f \rightarrow \infty$) in which

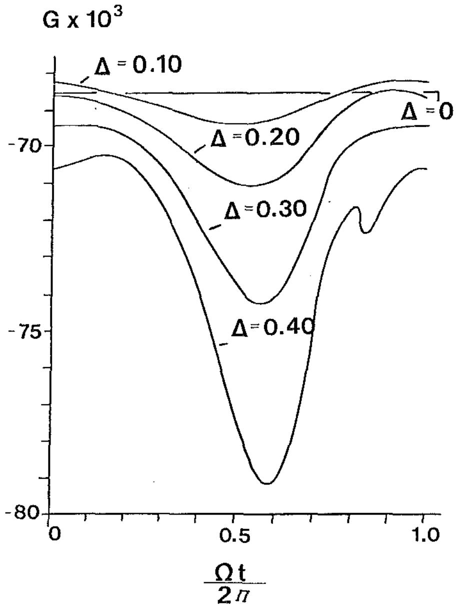


Fig. 4 Instantaneous growth rate versus time for least stable wall mode. R = 3000, β = 5.8

$$G \doteq \mathrm{Re}\{\lambda_0(t_0)\} \tag{21}$$

where $\lambda_0(t_0)$ is computed by equation (18). As expected, the results of the two methods of calculation of G do not differ significantly (Tozzi [13]).

Two features of the results of the calculation of the instantaneous growth rate G are significant. The first and most important feature concerns the wall mode in Fig. 4. The wall mode is significantly affected by the instantaneous velocity profile shape. However, in spite of the inflection ring criterion (1), which is a necessary condition for the instability of steady inviscid flows, the instantaneous profiles are even more stable than the inflectionless steady-flow profile. Furthermore, the stability increases with increasing values of Δ. This behavior of the flow is in accord with experimental results of Gilbrech and Combs [5] and Sarpkaya [7] up to certain values of Δ.

The second significant feature is displayed in Fig. 5 concerning the center mode. The instantaneous growth rate G of the center mode simply varies proportionately to the instantaneous Reynolds number because the central part of the instantaneous velocity profile retains its parabolic shape.

At this point it has been shown that, based on the wall modes which are deemed most relevant, the enhancement of the stability of the oscillatory flow is in accord with the experimental result that oscillations do delay transition. In order to discern some quantitative relationship between this stability theory and experimental transition studies, the following interpretation is proposed. The threshold growth (or decay) rate γ is defined by

$$\gamma = \mathrm{Re}\{\lambda - \lambda_{oc}\} \tag{22}$$

where λ_{oc} is the steady flow complex growth rate for the given value of α and the value of Reynolds number $R_c = 2200$. Natural transition in normal steady pipe flow experiments usually occurs at this value of Reynolds number.

Then, for fixed values of Reynolds number and oscillation amplitude, if $\gamma > 0$, the oscillations are said to destabilize the flow. If $\gamma < 0$, the oscillations are said to stabilize the flow. The value $\gamma = 0$ is defined as the threshold of stabilization. For fixed values of α the graph of the locus of points for which

$$\gamma(R, \Delta) = 0 \tag{23}$$

is determined. These critical values of R, denoted by $\bar{R}$, are plotted versus Δ in Fig. 6. The value of the critical Reynolds number $\bar{R}$ is a measure of the stabilization of disturbances with respect to a level which prevails in the steady flow on the verge of transition to turbulence. These values of $\bar{R}$ are compared to Sarpkaya's critical values of Reynolds number for transition to turbulence in Fig. 6. The curve labeled R_q denotes the quai-steady always unstable definition of critical Reynolds number proposed by Davis [22]. Davis' definition of R_q is

$$R_q = R_c\left(\frac{1}{1-\Delta}\right) \tag{24}$$

The value of R_q defines the value of Δ at which the smallest value of the instantaneous Reynolds number is equal to R_c = 2200. The value of R_q is an upper bound of the values of the critical Reynolds number of transition for low frequency oscillations.

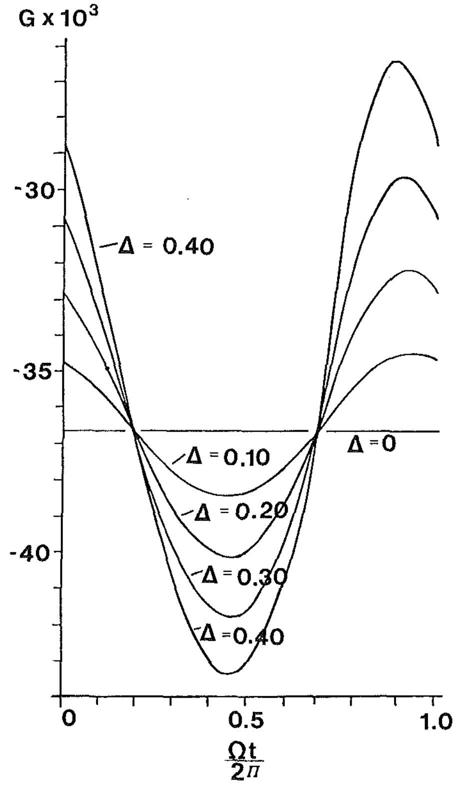


Fig. 5 Instantaneous growth rate versus time for least stable center mode. R = 3000, β = 5.8

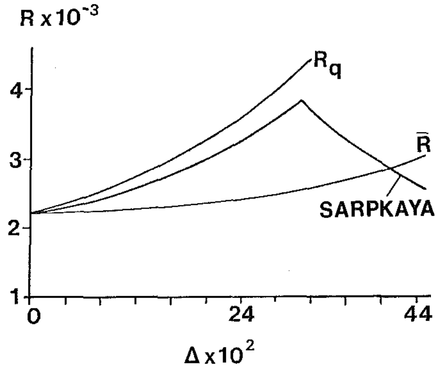


Fig. 6 Comparison of experimental results with locus of points for which $\gamma(R, \Delta)$ = 0. β = 5.8, α = 0.5. Rq is the always unstable quasi-steady bound.

Figure 6 compares the critical Reynolds numbers R_q, $\bar{R}$, and Sarpkaya's experimental results. It is evident that, although linear stability theory does correctly predict qualitatively an initial enhancement of stability of the oscillatory base flow with increasing oscillation amplitude, it does not seem to contain anything of quantitative significance concerning the mechanics of transition. In particular, the sudden break and diminishing stabilization as oscillation amplitude is increased which is exhibited by the experiments is not at all reflected by linear stability theory. Sarpkaya [7] speculates that this break in transition behavior is associated with occurrences of momentary reverse flow in the Stokes layer and the duration of the inflectional period of the velocity profiles. But this is definitely not the case, at least for axisymmetric disturbances, as shown by Figs. 4 and 5. However, Sarpkaya's surmise may in fact be true on the basis of nonaxisymmetric disturbances which may be much more unstable. It is important to note that disturbances introduced into experimental flows are three-dimensional.

5 Concluding Remarks

This study has shown that low frequency oscillatory Hagen-Poiseuille flow is slightly more stable to axisymmetric disturbances than the steady flow. More importantly, the highly inflectional instantaneous velocity profiles do not lead to instantaneous instability. Thus reasoning developed for steady flows must be applied with great caution to oscillatory flows. This lesson was learned earlier by von Kerczek and Davis [9] but on the basis of zero-mean oscillatory flows. Davis [22] argued that careful scale analysis must be performed before attempting to apply the stability results for steady flows to unsteady flows. However, in this case, even a scaling argument would lead to erroneous conjectures concerning the possible momentary instabilty of the flow. It is emphasized though that an instantaneous instability mechanism has not been completely precluded by the axisymmetric calculations presented here. There is evidence (Salwen and Grosch [1]) that nonaxisymmetric disturbances may play a much more important instability role in axisymmetric base flows. This problem ought to be investigated in the future.

Acknowledgment

This work was supported by the NSF under Grant No. CME-7900929.

References

1 Salwen, H., and Grosch, C. E., (1972), "The Stability of Poiseuille Flow in a Pipe of Circular Cross-section," *Journal of Fluid Mechanics*, Vol. 54, pp. 93-112.

2 Kerczek, C. von (1982, "The Instability of Oscillatory Plane Poiseuille Flow," *Journal of Fluid Mechanics*, Vol. 116, pp. 91—114.

3 Kerczek, C. von (1976), "Stability of Modulated Plane Couette Flow," *The Physics of Fluids*, Vol. 19, pp. 1288—1295.

4 Davey, A. (1973), "On the Stability of Plane Couette Flow to Infinitesimal Disturbances," *Journal of Fluid Mechanics*, Vol. 57, pp. 369-380.

5 Gilbrech, D. A., and Combs, G. D. (1963), "Critical Reynolds Numbers for Incompressible Pulsating Flow in Tubes," *Developments in Theoretical and Applied Mechanics*, Vol. 1, pp. 292-304.

6 Ramaprian, B. R., and Tu, S. W. (1980), "An Experimental Study of Oscillatory Pipe Flow at Transition Reynolds Numbers," *Journal of Fluid Mechanics*, Vol. 100, pp. 513-544.

7 Sarpkaya, T. (1966), "Experimental Determination of the Critical Reynolds Number for Pulsating Poiseuille Flow," ASME *Journal of Basic Engineering*, pp. 589-598.

8 Yellin, E. L. (1966), "Laminar-Turbulent Transition Process in Pulsatile Flow," *Circulation Research*, Vol. 19, pp. 791-804.

9 Kerczek, C. von, and Davis, S. H. (1974), "Linear Stability Theory of Oscillatory Stokes Layers," *Journal of Fluid Mechanics*, Vol. 62, pp. 753-773.

10 Yang, W. H., and Yih, C. S. (1977), "Stability of Time-periodic Flows in a Circular Pipe," *Journal of Fluid Mechanics*, Vol. 82, pp. 497-505.

11 Hino, M., Sawamoto, M., and Takasu, S. (1976), "Experiments on Transition to Turbulence in an Oscillatory Pipe Flow," *Journal of Fluid Mechanics*, Vol. 75, pp. 193-207.

12 Abramowitz, M., and Stegun, I. A., (Eds.), *Handbook of Mathematical Functions*, NBS Appl. Math. Series 55, U.S. Dept. of Commerce, Washington, D.C., 1964.

13 Tozzi, J. T. (1982), "The Stability of Pulsatile Flow in a Conduit of Circular Cross Section," Ph.D. dissertation, The Catholic University of America, University of Microfilm International, Ann Arbor, MI. 48106 (Order No. 82-21446).

14 Coddington, E. A., and Levinson, N., *Theooy of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

15 Gottlieb, D.,and Orszag, S. A. (1977), *Numerical Analysis of Spectral Methods: Theory and Applications*, Society for Industrial and Applied Mathematics, Philadelphia, 1977.

16 Shanks, D. (1955), "Nonlinear Transformation of Divergent and Slowly Convergent Sequences," *Journal of Mathematical Physics*, Vol. 34, pp. 1-42.

17 Smith, B. T. et al. (Authors) *Matrix Eigensystem Routines—EISPACK Guide*, (Ed.: Karlsruhe, G. and Hartmanis, J.), Springer Verlag, Berlin, 1974.

18 Davey, A., and Drazin, P. G. (1969), "The Stability of Poiseuille Flow in a Pipe," *Journal of Fluid Mechanics*, Vol. 36, pp. 209-218.

19 Schlichting, H. (1979), *Boundary-Layer Theory*, 7th Ed., McGraw-Hill, New York.

20 Lin, C. C., *The Theory of Hydrodynamic Stability*, Cambridge University Press, London, 1966.

21 Davey, A., and Nguyen, H. P. F. (1971), "Finite-amplitude Stability of Pipe Flow," *Journal of Fluid Mechanics*, Vol. 45, pp. 701-720.

22 Davis, S. H. (1976), "The Stability of Time-Periodic Flows," *Annual Review of Fluid Mechanics*, Vol. 8, pp. 57-74.

R. L. Adams
Manager, Hard Copy Research,
Tektronix, Inc.,
Beaverton, Ore. 97077;
and Associate Professor of
Mechanical Engineering,
Oregon State University,
Corvallis, Ore. 97331

J. Roy
Engineer, Hard Copy Research,
Tektronix, Inc.,
Beaverton, Ore.

A One-Dimensional Numerical Model of a Drop-On-Demand Ink Jet

MacCormack's predictor-corrector algorithm is used to solve one-dimensional model equations of drop development from a drop-on-demand ink jet. The calculation is done in Lagrangian coordinates, and the results are compared with calculations reported in which an axisymmetric marker-and-cell algorithm is used. The comparison indicates that, although drop velocities differ in the two cases, good qualitative results can be obtained with the less complex one-dimensional approach.

Introduction

The drop-on-demand ink jet printer is emerging as an inexpensive device for the production of high-resolution computer-graphics hard copy. The controlled production of single ink drops is achieved through the action of a sudden pressure pulse produced by a piezoelectric device. In order to enhance our fundamental understanding of the drop-development phenomenon, numerical solutions of the appropriate conservation equations together with stroboscopic flow-visualization studies have been pursued.

This paper proposes a one-dimensional model of the drop development based on a finite difference solution of the governing equations in Lagrangian coordinates. Although many studies of drop formation in continuous jets have been published (see the review of Bogy, 1979), the investigation of drop-on-demand jets has been limited to the semiempirical modeling of Kyser et al. (1981) and the axisymmetric Navier-Stokes analysis of Fromm (1982). The one-dimensional model proposed here offers an alternative approach of lesser complexity than Fromm's and of higher accuracy than Kyser's. Since the fluid domain is surrounded by a free surface, the Lagrangian approach also offers a direct simultaneous way to solve the governing equations and to define the fluid geometry. It is expected that an approach based upon the Eulerian frame would be more cumbersome because of the necessity to update the coordinate grid after each time step. Such calculations are done either with marked particles (marker-and-cell) that define the free surface in an Eulerian frame or with moving grids (see Hirt, 1971). Fromm (1982, 1984) uses a vorticity-stream function formulation of the Navier-Stokes equations in axisymmetric form together with the marker-and-cell approach in his calculations. The accuracy of the one-dimensional approach will be established by comparing the results with Fromm's calculations.

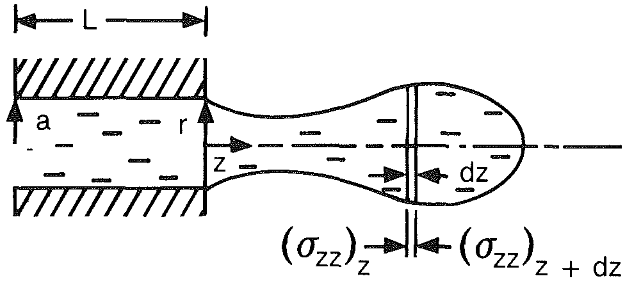


Fig. 1 Geometry of the liquid column emerging from an orifice

Analysis

Figure 1 shows a liquid column emerging without wetting from a nozzle to form a single drop. Following Lee's investigation of continuous jets (1974), one can develop a one-dimensional model of the motion and geometry of the liquid jet from conservation of mass and Newton's second law of motion. Bogy (1979) and Bogy et al. (1980), however, obtain different results by incorporating radial-inertia effects through a one-dimensional Cosserat theory. According to the simpler model by Lee (1974), the radius R and velocity V of an element of fluid in the column are governed by

$$\frac{1}{R^2}\frac{dR^2}{dt}+\frac{\partial V}{\partial z}=0 \tag{1}$$

$$\frac{dV}{dt}-\frac{1}{\rho}\frac{\partial}{\partial z}\sigma_{zz}=0 \tag{2}$$

where σ_{zz} is the total normal stress acting on the liquid element. Note that the time derivatives are expressed in Lagrangian form while the Eulerian form of the spatial derivatives is retained.

The normal stress includes viscous as well as hydrostatic contributions and is influenced by surface tension. The pressure of the liquid is related to the surrounding air pressure through

$$p=p_{\text{air}}+T\left(\frac{1}{R_N}+\frac{1}{R_T}\right)+2\mu\frac{\partial U_r}{\partial r} \tag{3}$$

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 two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June, 1984; final revision, February, 1985.

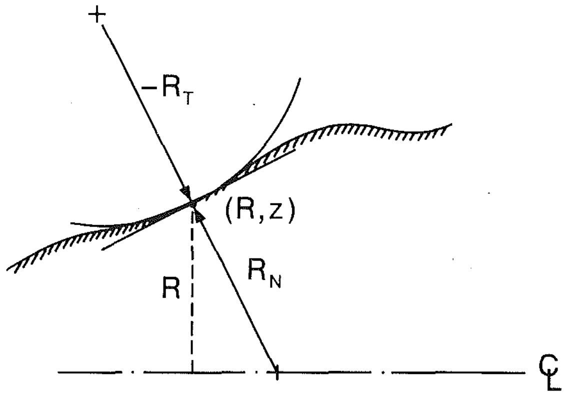


Fig. 2 Principal radii of curvature of the liquid column

where R_N and R_T are the principal radii of curvature of the column as shown in Fig. 2 and are given by

$$R_N = R\sqrt{1+\left(\frac{\partial R}{\partial z}\right)^2} \tag{4}$$

$$R_T = \frac{-\sqrt{\left[1+\left(\frac{\partial R}{\partial z}\right)^2\right]^3}}{\frac{\partial^2 R}{\partial z^2}} \tag{5}$$

According to equation (1)

$$U_r = \frac{dR}{dt} = -\frac{R}{2}\frac{\partial V}{\partial z} \tag{6}$$

thus

$$\frac{\partial U_r}{\partial r} = -\frac{1}{2}\frac{\partial V}{\partial z} \tag{7}$$

producing the total normal stress

$$\sigma_{zz} = -p + 2\mu\frac{\partial V}{\partial z} = -p_{\text{air}} - T\left(\frac{1}{R_N}+\frac{1}{R_T}\right) + 3\mu\frac{\partial V}{\partial z} \tag{8}$$

When fluid properties and air pressure are constant along the column, equation (2) can be written as

$$\frac{dV}{dt} - \frac{1}{\rho}\left[-T\frac{\partial}{\partial z}\left(\frac{1}{R_N}+\frac{1}{R_T}\right) + 3\mu\frac{\partial^2 V}{\partial z^2}\right] = 0 \tag{9}$$

Following Fromm (1982), one can cast (1) and (9) in nondimensional form using the characteristic velocity

$$V_0 = \sqrt{\frac{\sigma}{\rho a}} \tag{10}$$

where a is the radius of the orifice, to produce

$$\frac{1}{R^{+2}}\frac{dR^{+2}}{dt^+} + \frac{\partial V^+}{\partial z^+} = 0 \tag{11}$$

$$\frac{dV^+}{dt^+} + \frac{\partial}{\partial z^+}\left(\frac{1}{R_N^+}+\frac{1}{R_T^+}\right) - 3\frac{\text{We}}{\text{Re}}\frac{\partial^2 V^+}{\partial z^{+2}} = 0 \tag{12}$$

In equation (12), the Weber to Reynolds number ratio, We/Re is $\mu/\rho V_0 a$.

The Eulerian form of equations (11) and (12) is a system of hyperbolic differential equations with the dynamic wavespeed dependent upon V_0 and the radii of curvature of the drop. Thus it seems appropriate to use a numerical algorithm of the Lax-Wendroff type in order to obtain a numerical solution, and MacCormack's method (see e.g., Roache, 1982) was selected because of its superior performance and wide use in gas dynamic calculations. Also, the method is second-order accurate in both space and time.

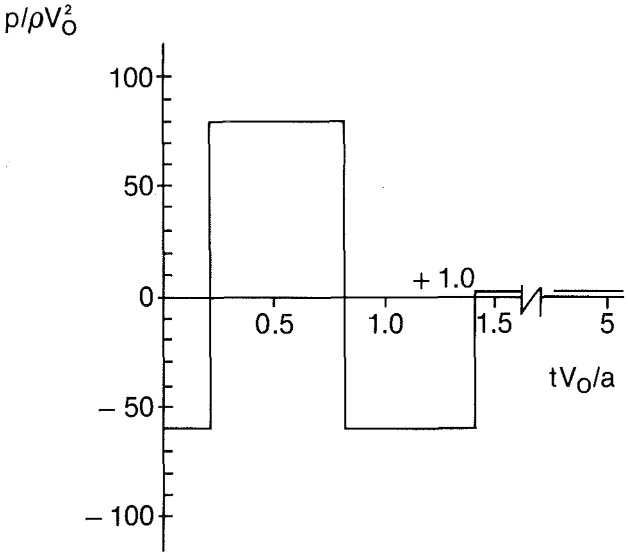


Fig. 3 Drive pressure for Fromm's calculations and with Re/We = 5

For the numerical analysis, (11) and (12) are expressed as the vector equation

$$\frac{d\mathbf{U}^+}{dt^+} + \frac{\partial \mathbf{F}^+}{\partial z^+} = 0 \tag{13}$$

where

$$\mathbf{U}^+ = \begin{Bmatrix} \ln R^{+2} \\ V^+ \end{Bmatrix} \tag{14}$$

and

$$\mathbf{F}^+ = \begin{Bmatrix} V^+ \\ \left(\frac{1}{R_N^+}+\frac{1}{R_T^+}\right) - 3\frac{\text{We}}{\text{Re}}\frac{\partial V^+}{\partial z^+} \end{Bmatrix} \tag{15}$$

Then the solution of (13) is accomplished using MacCormack's algorithm (Roache, 1982) in the form

$$\mathbf{U}_i^{*n+1} = \mathbf{U}_i^n - \Delta t^+ \frac{\mathbf{F}_i^n - \mathbf{F}_{i-1}^n}{z_i^+ - z_{i-1}^+} \tag{16}$$

$$\mathbf{U}_i^{n+1} = \frac{1}{2}\left(\mathbf{U}_i^n + \mathbf{U}_i^{*n+1} - \Delta t^+ \frac{\mathbf{F}_{i+1}^{*n+1} - \mathbf{F}_i^{n+1}}{z_{i+1}^+ - z_i^+}\right) \tag{17}$$

where the superscript denotes the time step number, the subscript is associated with the ith element of the liquid column and $\mathbf{F}^*$ is to be evaluated in terms of $\mathbf{U}^*$. The position coordinates of the elements z_i^+, are held fixed until the calculation is completed for a given time step, then updated once new velocities are obtained using the trapezoidal approximation

$$z_i^{+n+1} = z_i^{+n} + \frac{(V_i^{+n} + V_i^{+n+1})\Delta t^+}{2} \tag{18}$$

The time step Δt^+ for the integration of equations (15-17) is based upon the capillary wavespeed of the smallest resolvable wavelength, as proposed by Foote (1973). For the present case, this wavelength is twice the minimum distance between nodes, and the resulting velocity is

$$c^+ = \frac{3}{2}\sqrt{\frac{\pi}{|z_i^+ - z_{i-1}^+|_{\min}}} \tag{19}$$

so that the time step is determined from

$$\frac{\Delta t^+}{|z_i^+ - z_{i-1}^+|_{\min}} \le \frac{1}{c^+} \tag{20}$$

From equation (20), $\Delta t^+ < 0.0266$ for $|z_i^+ - z_{i-1}^+|_{\min} = 0.1$. Numerical experiments indicate that $\Delta t^+ < 0.01$ is necessary for stability.

As the calculation proceeds, elements of the column may spread or contract along the direction of motion and lead to nonuniform spacing. Spreading results in a loss of resolution of the radius and velocity, while contraction necessitates a diminishing time step to maintain stability. To avoid these problems, elements are added or deleted so that element spacing remains nearly uniform. When nodes are to be added, numerical inaccuracies in the radius calculation are avoided by using a cubic spline fit following Daly (1969). This procedure results in a tridiagonal system for the second derivative of the radius in the form

$$(z_i - z_{i-1})\frac{r''_{i-1}}{6} + (z_{i+1} - z_{i-1})\frac{r''_i}{3} + (z_{i+1} - z_i)\frac{r''_{i+1}}{6} = \frac{r_{i+1} - r_i}{z_{i+1} - z_i} - \frac{r_i - r_{i-1}}{z_i - z_{i-1}} \qquad (21)$$

where $r'' = \partial^2 R/\partial z^2$ and the following boundary conditions are used

$r'' = 0$ at the orifice

and r'' = const at the head or tail of the drop.

Once r''_i is obtained from equation (21), the radius of added nodes is calculated from

$$r(z) = r''_{i-1}\frac{(z_i - z)^3}{6(z_i - z_{i-1})} + r''_i\frac{(z - z_{i-1})}{6(z_i - z_{i-1})} + \left(\frac{r_i}{z_i - z_{i-1}} - r''_i\frac{z_i - z_{i-1}}{6}\right)(z - z_{i-1}) + \left(\frac{r_{i-1}}{z_i - z_{i-1}} - r''_{i-1}\frac{z_i - z_{i-1}}{6}\right)(z_i - z) \qquad (22)$$

The regularity with which nodes are added to or subtracted from the fluid column varies during the course of drop formation and is dependent on the drive pressure shape. Typically, nodes are added once every ten time steps.

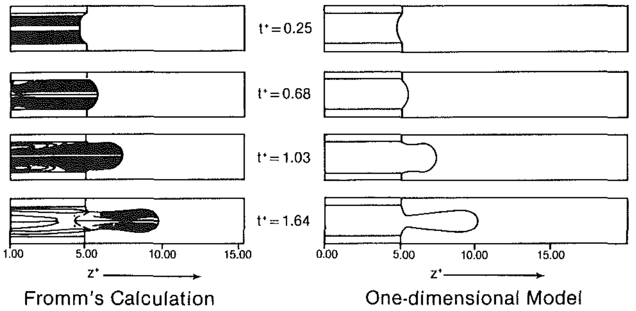


Fig. 4 Drop formation sequence for the drive pressure of Fig. 3

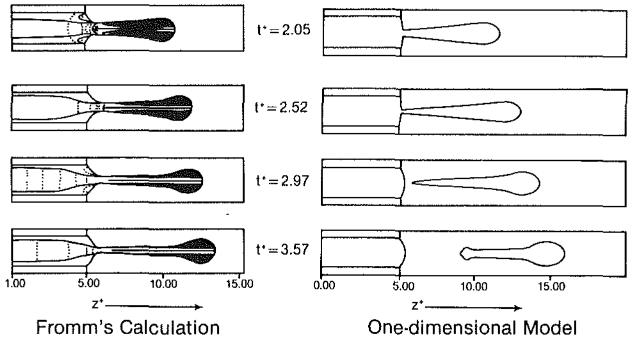


Fig. 5 Continuation of Fig. 4

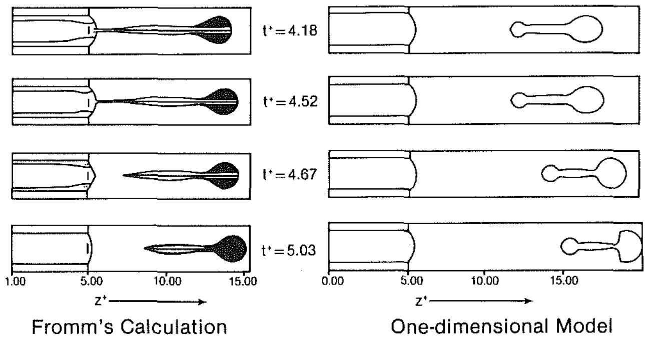


Fig. 6 Continuation of Fig. 4

Comparison With Axisymmetric Calculations

Fromm (1982) uses an axisymmetric marker-and-cell algorithm to simulate drop formation from a nozzle with specified driving pressure. The calculation neglects the effect of orifice wetting, and the lack of drive-pressure data caused him to assume the distribution, with time, to be as shown in Fig. 3. The results of Fromm's calculation will be used to assess the accuracy of the one-dimensional approach.

For the one-dimensional analysis, according to the approach just described, the necessary outflow boundary condition at the orifice exit is established from a linear analysis of the transient viscous flow within the nozzle. Taking the flow to be one dimensional, the local velocity is governed by

$$\frac{\partial u^+}{\partial t^+} = -\frac{d}{dz^+}(p^+(t^+)) + \frac{1}{\mathrm{Re}}\left(\frac{\partial^2 u^+}{\partial r^{+2}} + \frac{1}{r^+}\frac{\partial u^+}{\partial r^+}\right) \qquad (23)$$

with $u^+(r^+, 0) = 0$ (initial condition)

and $u^+(1, t^+) = 0$ (no slip condition)

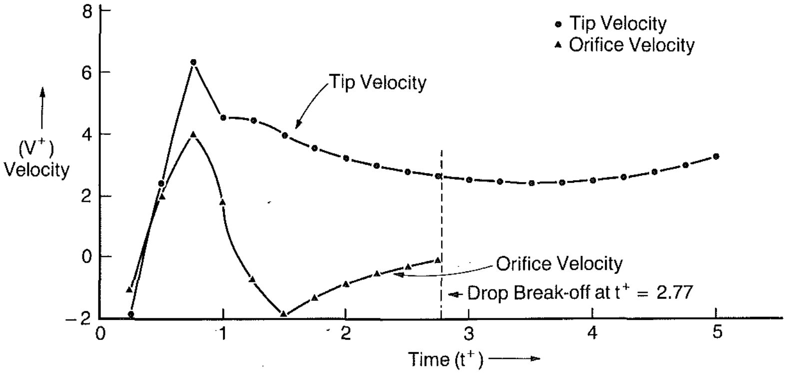


Fig. 7 Tip velocity and orifice exit velocity for the drive pressure of Fig. 3 and Re/We = 5

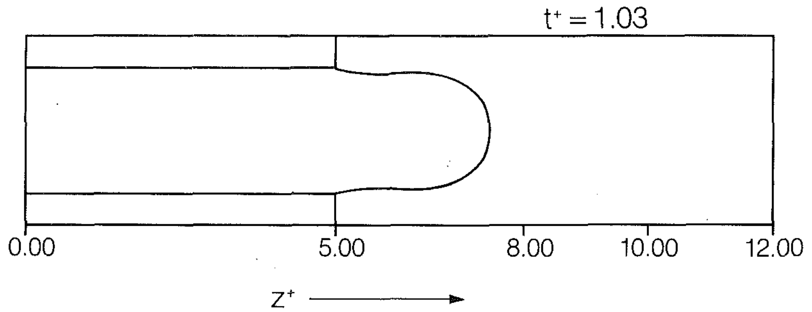


Fig. 8 Drop profile for the drive pressure of Fig. 3 and Re/We = ∞ at $t^+ = 1.03$

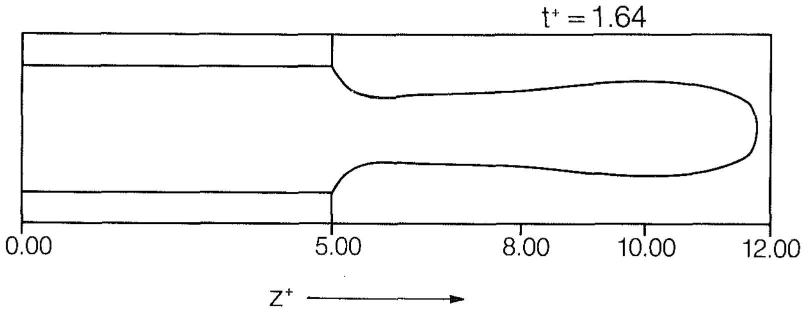


Fig. 9 Drop profile for the drive pressure of Fig. 3 and Re/We = ∞ at $t^+ = 1.64$

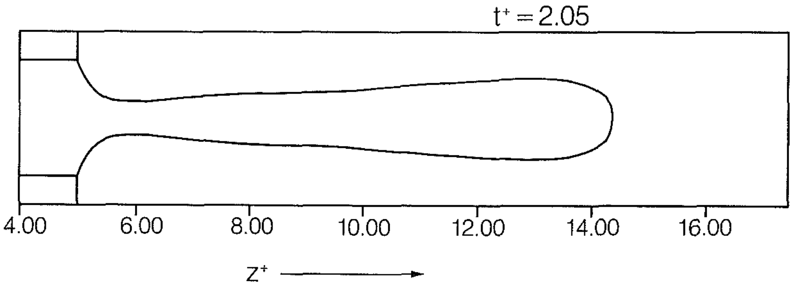


Fig. 10 Drop profile for the drive pressure of Fig. 3 and Re/We = ∞ at $t^+ = 2.05$

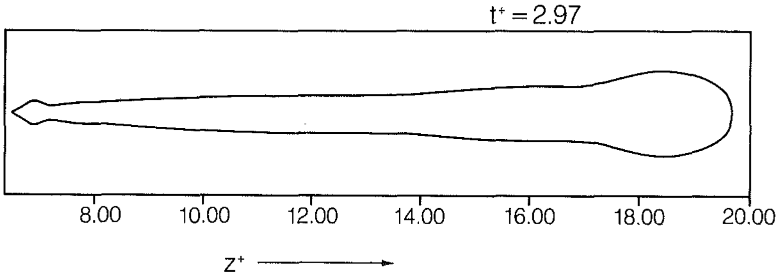


Fig. 11 Drop profile for the drive pressure of Fig. 3 and Re/We = ∞ at $t^+ = 2.97$

The solution of equation (23) for a step change in pressure gradient is given in White (1974). This result can be generalized through the use of the superposition integral to obtain the centerline velocity and the spatially averaged velocity

$$u^+(o,t^+) = \frac{2}{L^+}\int_0^{t^+} \Delta p^+(s)\left\{\sum_1^{\infty} \frac{\exp[-\lambda_n^2(t^+-s^+)/\mathrm{Re}]}{\lambda_n J_1(\lambda_n)}\right\}ds^+ \tag{24}$$

$$u^+_{\mathrm{ave}}(t^+) = \frac{4}{L^+}\int_0^{t^+} \Delta p^+(s)\left\{\sum_1^{\infty} \frac{\exp[-\lambda_n^2(t^+-s^+)/\mathrm{Re}]}{\lambda_n^2}\right\}ds^+ \tag{25}$$

where

$$\Delta p^+ = p^+_{\mathrm{drive}} - \left(\frac{1}{R_N^+} + \frac{1}{R_T^+}\right)_{\mathrm{orifice}}$$

and λ_n represents the zeros of J_0. Furthermore, the meniscus is assumed to be a frustrum of a sphere until its positive displacement becomes equal to the nozzle radius. Thus

$$R_N^+ = R_T^+ = \frac{1}{2}\left(z^+ + \frac{1}{z^+}\right), \quad \text{for } -1 < z^+ < 1 \tag{26}$$

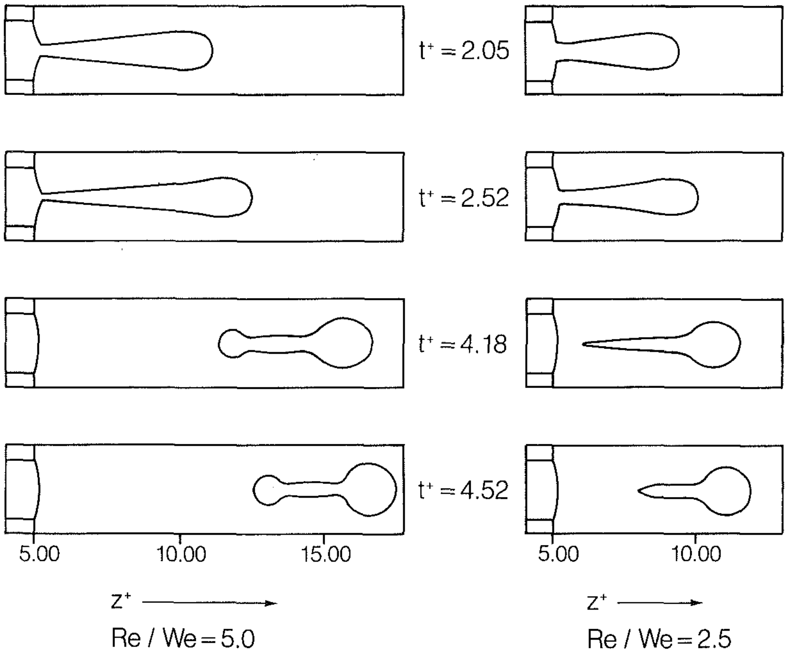


Fig. 12 Drop formation sequence for the drive pressure of Fig. 3

Equations (24) and (26) are used to describe the positive and negative meniscus shape until the positive radius of the spherical cap meniscus becomes equal to the nozzle radius. At this point, the hemisphere is divided into ten elements and the integration of equations (16) and (18) commences with equation (25) providing the boundary conditions at the orifice exit. Additional elements are added at the nozzle exit when the node adjacent to the nozzle has displaced a distance equal to the initial node spacing (0.10 in this case). The leading and second nodes are assumed to move at the same velocity, and this proceudre is used for the trailing element of a drop once break-off occurs. Break-off is taken to be the first time step at which the local radius drops below $0.05a$, as assumed by Fromm.

The results of the one-dimensional calculation are shown in sequence along with Fromm's calculations for $Re/We = 5$ in Figs. 4-6. Shown is the geometry of the liquid column at several positions from the orifice. Comparison of the two results indicates that good agreement is obtained up to the point of break-off, although the one-dimensional model predicts break-off earlier than the full Navier-Stokes calculation. Beyond break-off, the one-dimensional results indicate higher velocities of the drop so that, by the end of the sequence, the head positions differ by about 40 percent. Events leading to satellite formation are also more rapid for the one-dimensional model. It seems that although the one-dimensional model produces higher velocities than the full Navier-Stokes calculation, it does reveal the essential qualitative features of the main and satellite drops that are formed. The velocity at the tip of the liquid column and the average velocity at the orifice exit are plotted for this case in Fig. 7.

Drop formation sequences for the inviscid case ($Re/We = \infty$) and two viscous cases ($Re/We = 2.5$ and 5) are shown in Figs. 8-12. In the inviscid calculations (Figs. 8-11) and the two viscous cases (Fig. 12), only the coefficient of viscosity of the fluid was changed. Therefore, the comparison of these three cases illustrates the action of viscosity in retarding the liquid column. Also note that the numerical algorithm used in these calculations was developed for hyperbolic partial differential equations so that the procedure works well over the range of Re/We.

Summary

A one-dimensional model of the motion of a liquid jet has been used to simulate drop formation from a drop-on-demand ink jet. The surface-tension effect results in a hyperbolic system of partial-differential equations, which has been solved numerically using MacCormack's predictor-corrector algorithm. Results of calculations using the model compare qualitatively with calculations using the marker-and-cell Navier-Stokes code by Fromm (1982).

References

Bogy, D. B., 1979, "Drop Formation in a Circular Liquid Jet," *Annual Review of Fluid Mechanics*, Vol. 11, pp. 207-227.

Bogy, D. B., Shine, S. J., and Talke, F. E., 1980, "Finite Difference Solution of the Cosserat Fluid Equations," *Journal of Computational Physics*, Vol. 38, pp. 294-326.

Daly, B. J., 1969, "A Technique for Including Surface Tension Effects in Hydrodynamic Calculations," *Journal of Computational Physics*, Vol. 4, pp. 97-117.

Foote, G. B., 1973, "A Numerical Method for Studying Liquid Drop Behavior: Simple Oscillation," *Journal of Computational Physics*, Vol. 11, pp. 507-530.

Fromm, J., 1982, "A Numerical Study of Drop-on-Demand Jets," in *Proceedings of the Second International Colloquium on Drops and Bubbles*, pp. 54-82.

Fromm, J. E., 1984, "Numerical Calculation of the Fluid Dynamics of Drop-On-Demand Jets," *IBM Journal of Research and Development*, Vol. 28, May, pp. 322-333.

Hirt, C. W., 1971, "An Arbitrary Lagrangian-Eulerian Computing Technique," in *Proceedings of the Second International Conference on Numerical Methods in Fluid Dynamics, Lecture Notes in Physics*, Vol. 8, Springer-Verlag, New York.

Kyser, E. L., Collins, L. F., and Herbert, N., 1981, "Design of an Impulse Ink Jet," *Journal of Applied Photographic Engineering*, Vol. 7, pp. 73-79.

Lee, H. C., 1974, "Drop Formation in a Liquid Jet," *IBM Journal of Research and Development*, July, pp. 364-369.

Roache, P. J., 1982, *Computational Fluid Dynamics*, Hermosa Publishers.

White, F. M., 1974, *Viscous Fluid Flow*, McGraw-Hill.

The Effect of the Earth's Rotation on Channel Flow

C. G. Speziale[1]
Stevens Institute of Technology,
Hoboken, N.J. 07030

The influence that the rotation of the earth has on laminar channel flow is investigated theoretically. The full nonlinear Navier-Stokes equations relative to a reference frame rotating with the earth are solved numerically for laminar flow in a rectangular channel whose axis is aligned east-west: the orientation which yields the most drastic effect. It is demonstrated that for channels of moderate width (less than 1 ft for the flow of most liquids), the rotation of the earth can give rise to a roll instability which has a severe distortional effect on the classical parabolic velocity profile. Consequently, the usual assumption made of neglecting the effect of the earth's rotation in the calculation of channel flow can lead to serious errors unless the channel is substantially smaller than this size. It is briefly shown that similar effects would be expected for turbulent channel flow when the channel width is approximately an order of magnitude larger.

1 Introduction

Approximately thirty years ago, Benton [1956] conducted an interesting study which suggested that the rotation of the earth gives rise to a secondary motion in laminar pipe flow which could have a significant distortional effect on the usual parabolic velocity profile in pipes of moderate size (e.g., a pipe with a radius of 1 in.). This secondary flow, which causes a lateral transport of momentum, yields an asymmetric axial velocity profile in contrast to the classical parabolic velocity profile obtained for this viscous flow in an inertial framing. While it had been known by geophysicists that the rotation of the earth cannot be neglected in the study of fluid flow in estuaries, rivers, and oceans (cf, Pedlosky 1979), it was, nevertheless, quite surprising that it could have a non-negligible effect on a standard laboratory pipe flow of significantly smaller dimensions. Although it would appear that these results provided a strong motivation for examining the effect of the earth's rotation on other analogous laboratory flow configurations, few studies, if any, along these lines have been subsequently reported.

The purpose of the present paper is to examine the effect that the rotation of the earth has on the pressure-driven flow of a viscous fluid in a rectangular channel with a large aspect ratio – the experimental configuration used to simulate Poiseuille flow. Here, the axis of the channel will be aligned east-west with the side walls of the channel aligned parallel to the axis of rotation of the earth (the orientation that yields the most drastic effect from the earth's rotation). The full nonlinear Navier-Stokes equations relative to a reference frame rotating with the earth will be solved numerically by the same finite difference code that was developed by Speziale and Thangam [1983] for rotating channel flow (the present paper represents an application of that more general study). Provided that the channel is sufficiently small, the rotation of the earth merely gives rise to a weak double-vortex secondary flow (see Fig. 1), which has very little effect on the axial velocity profiles in the interior of the channel which are parabolic in a strong approximate sense. However, it will be shown that for

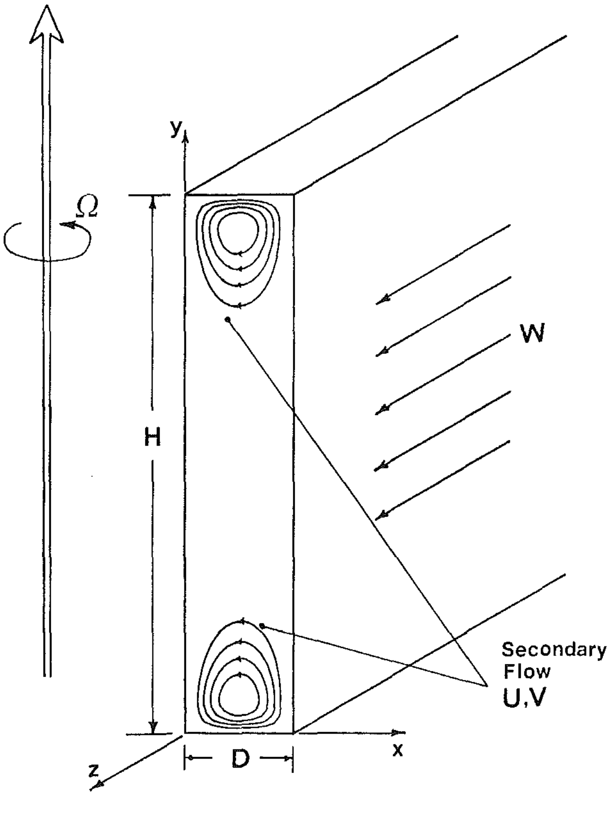


Fig. 1 Secondary flow in a laminar channel flow subjected to a spanwise rotation

---

[1] Present address: School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 18, 1985; final revision, July 26, 1985.

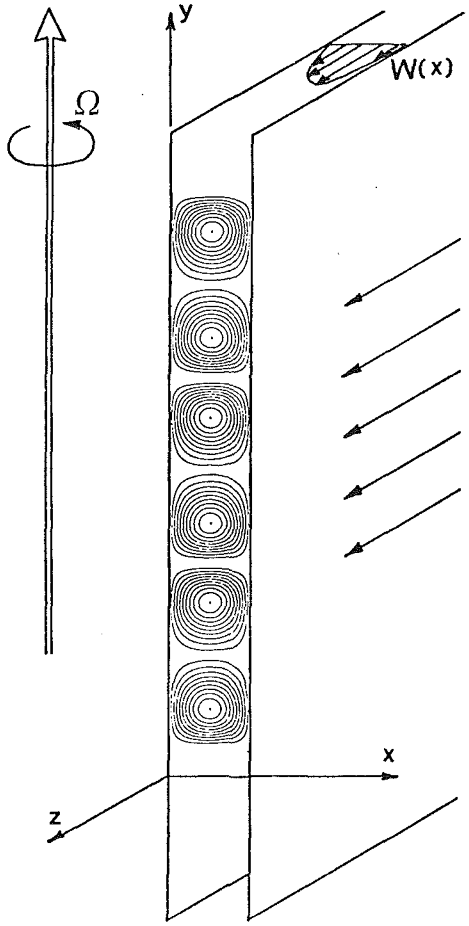


Fig. 2 Roll instability in a laminar Poiseuille flow subjected to a spanwise rotation

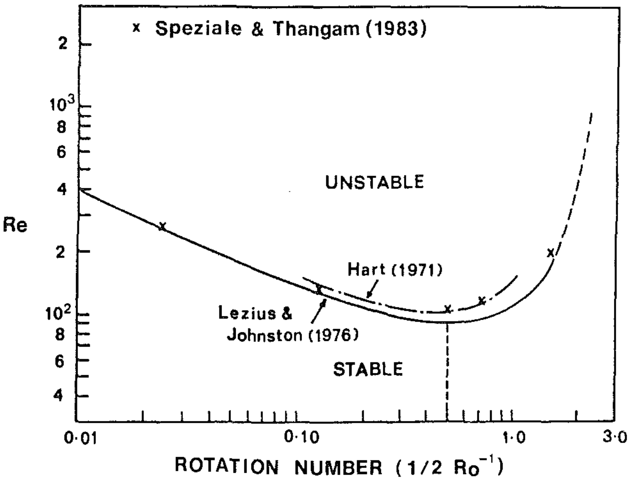


Fig. 3 Stability boundaries for laminar channel flow subjected to a spanwise rotation

channels of moderate size (i.e., for channel widths greater than a few inches), the rotation of the earth can give rise to a roll instability which has a severe distortional effect on the classical parabolic velocity profile by making it noticeably *asymmetric*. Consequently, for this case, the usual assumption made of neglecting the effect of the earth's rotation in the calculation of a laboratory channel flow can lead to serious error. This rather surprising result, which is analogous to that found by Benton [1956], will be validated by making comparisons with previously conducted stability analyses and experimental investigations on rotating channel flow (see Hart 1971 and Lezius and Johnston 1976). Extensions of these results to the turbulent regime will also be discussed briefly along with the prospects for future research.

2 Channel Flow in the Rotating Framework of the Earth

The effect of the earth's rotation on the fully-developed laminar flow of an incompressible viscous fluid in a rectangular channel with a large aspect ratio H/D (the experimental configuration used to simulate plane Poiseuille flow) will be considered as illustrated in Fig. 1. The flow is driven by a constant axial pressure gradient

$$\frac{\partial P}{\partial z} = -G \tag{1}$$

and all flow properties are assumed to be independent of the axial coordinate z. For simplicity, the sides of the channel are aligned parallel to the axis of rotation of the earth and the axial flow is oriented east-west. This orientation gives rise to the most drastic effect since it constitutes a pure spanwise rotation; axial rotations are of little consequence since, relative to an observer moving with the channel, they leave the flowfield unaffected. The velocity field $\mathbf{V}$ takes the form

$$\mathbf{V} = u(x,y)\mathbf{i} + v(x,y)\mathbf{j} + w(x,y)\mathbf{k} \tag{2}$$

(where w is the axial velocity and u and v are the secondary flow velocities) since for non-zero rotation rates Ω it no longer is possible to maintain a unidirectional flowfield (cf, Benton 1956 and Hart 1971). Of course, in an inertial framework (i.e., for $\Omega = 0$), the velocity field is of the unidirectional form

$$\mathbf{V} = w(x,y)\mathbf{k}. \tag{3}$$

The axial velocity profile w in (3) is, in a strong approximate sense, parabolic along the horizontal centerline $y = H/2$ of the channel, provided that H/D is large (it becomes exactly parabolic for plane Poiseuille flow which is approached as $H/D \to \infty$). The flow properties in the region of interest near the horizontal centerline of the channel depend weakly on H/D provided that this parameter is greater than 6 or 7. The equations of motion for rotating channel flow can be written in a modified vorticity-stream function formulation (see Speziale and Thangam 1983) which takes the dimensionless form

$$u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} = C + \frac{1}{\mathrm{Re}}\nabla^2 w + \frac{1}{\mathrm{Ro}}u \tag{4}$$

$$u\frac{\partial \zeta}{\partial x} + v\frac{\partial \zeta}{\partial y} = \frac{1}{\mathrm{Re}}\nabla^2 \zeta + \frac{1}{\mathrm{Ro}}\frac{\partial w}{\partial y} \tag{5}$$

$$\nabla^2 \psi = \zeta \tag{6}$$

$$u = -\frac{\partial \psi}{\partial y}, \quad v = \frac{\partial \psi}{\partial x} \tag{7}$$

where ψ is the secondary flow stream function, $\zeta \equiv \partial v/\partial x - \partial u/\partial y$ is the axial vorticity, and the flow variables are nondimensionalized with respect to the integrated average axial velocity W_o and the channel width D. In equations (4) and (5),

$$\mathrm{Re} = \frac{W_o D}{\nu}, \quad \mathrm{Ro} = \frac{W_o}{2\Omega D}, \quad C = \frac{GD}{\rho W_0^2} \tag{8}$$

are, respectively, the Reynolds number, Rossby number, and dimensionless pressure gradient where ρ is the density of the fluid and ν is the kinematic viscosity. Equations (4)-(7) must be solved subject to the boundary conditions

$$u = 0, \quad v = 0, \quad w = 0, \quad \psi = 0 \tag{9}$$

on the walls of the channel.

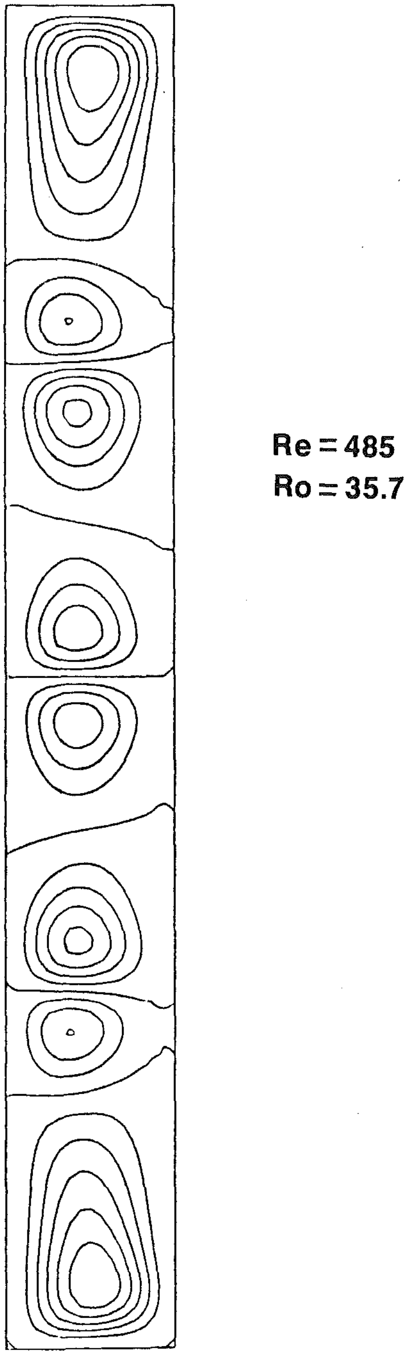


Fig. 4 Fully-developed secondary flow streamlines in a laminar channel flow subjected to a spanwise rotation: Re = 485, Ro = 35.7

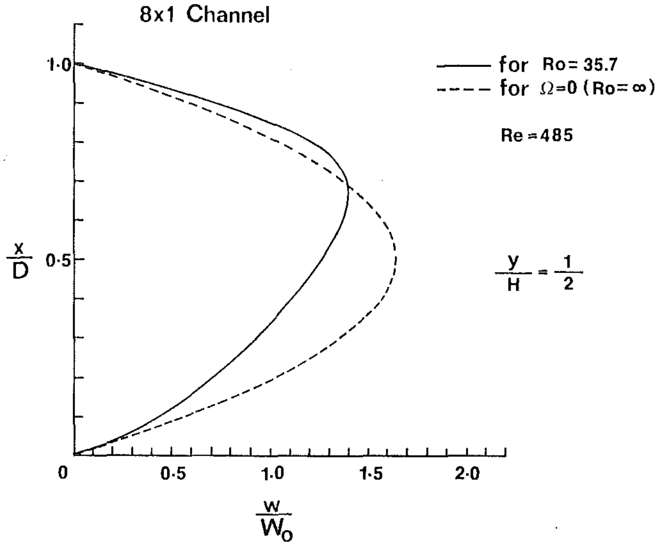


Fig. 5 Fully-developed axial velocity profile along the horizontal centerline of a channel subjected to a spanwise rotation: Re = 485, Ro = 35.7

The equations of motion (4)-(7) will be solved using the explicit finite difference code that was developed by Speziale [1982] and Speziale and Thangam [1983]. In this approach, the convective terms are formulated using Arakawa's scheme; the viscous diffusion terms are formulated using the DuFort-Frankel scheme; and the Poisson equation for the stream function is solved using cyclic reduction. This explicit scheme, which is second-order accurate, has excellent numerical stability characteristics (see Speziale 1982 for more details on this numerical method). Furthermore, unlike in the linear stability analyses presented by Hart [1971] and Lezius and Johnston [1976], it is possible to precisely determine the distortional effect that the secondary flow has on the axial velocity since the full nonlinear Navier-Stokes equations are solved. Numerical results will be presented in the next section for several Reynolds numbers and Rossby numbers in the range

$$100 < \mathrm{Re} < 2000, \quad 10 < \mathrm{Ro} < 1000 \tag{10}$$

for which roll instabilities can occur. Calculations will be conducted in a channel with an aspect ratio $H/D = 8$, which is close in size to that used by Hart [1971] and Lezius and Johnston [1976] in their experimental studies.[2] The channel will be discretized into a 16×128 finite difference mesh – the same mesh used by Speziale and Thangam [1983].

3 Numerical Results

Before presenting the specific numerical results obtained in this study, it would be advantageous to briefly review the nature of the results obtained in previous theoretical and experimental studies on rotating channel flow. The results of these previous studies (see Hart 1971, Lezius and Johnston 1976, and Speziale and Thangam 1983) indicated that there were three flow regimes in channel flow subjected to a spanwise rotation: a weak double vortex secondary flow at slow rotation rates; a roll instability at intermediate rotation rates (see Fig. 2); and a restabilized Taylor-Proudman regime at rapid rotation rates. The critical disturbance mode for roll instabilities occurs at a Reynolds number $\mathrm{Re} \approx 100$ and a Rossby number $\mathrm{Ro} \approx 1$ (no such instabilities will occur at significantly smaller Reynolds numbers). In Fig. 3, the stability boundaries for the onset of roll instabilities in rotating channel flow obtained by Hart [1971], Lezius and Johnston [1976] and Speziale and Thangam [1983] are shown in terms of the Reynolds number and rotation number (i.e., the dimensionless quantity $1/2\,R_o^{-1}$). Since these results, which were obtained by completely different procedures, are in a close range of one another, there is strong reason to have confidence in them. It is clear from Fig. 3 that a roll instability can occur at relatively large Rossby numbers. For instance, at a Reynolds number $\mathrm{Re} = 400$, a roll instability would occur for a Rossby number $\mathrm{Ro} \approx 50$. Furthermore, it will be shown that for $\mathrm{Re} = 1860$, a roll instability can occur when $\mathrm{Ro} \approx 667$ which is a rather large Rossby number (this result, whose validity would be expected

---

[2] The specific value of $H/D = 8$ was chosen since it is a power of 2. In this fashion, a uniform mesh size in x and y could be used and the requirement of the cyclic reduction scheme that the number of subintervals be a power of 2 would be satisfied identically (see Speziale [1982]).

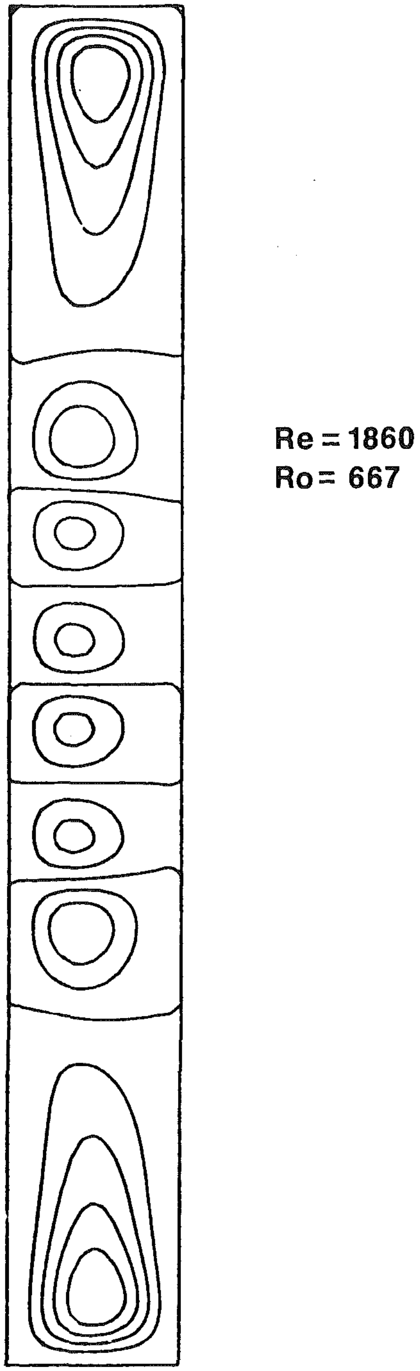


Fig. 6 Fully-developed secondary flow streamlines in a laminar channel flow subjected to a spanwise rotation: Re = 1860, Ro = 667

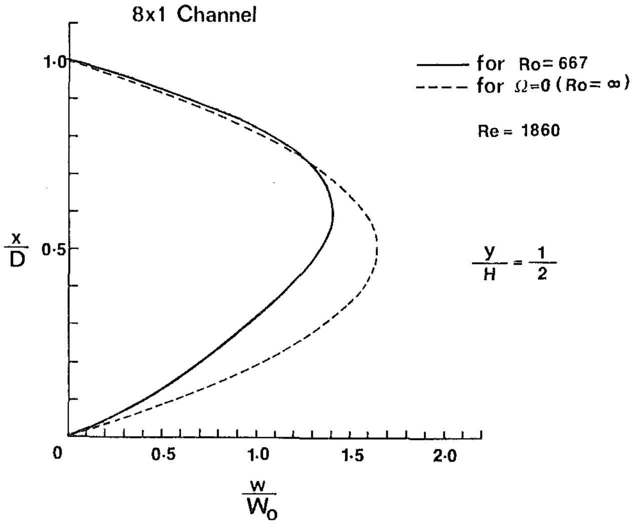


Fig. 7 Fully-developed axial velocity profile along the horizontal centerline of a channel subjected to a spanwise rotation: Re = 1860, Ro = 667

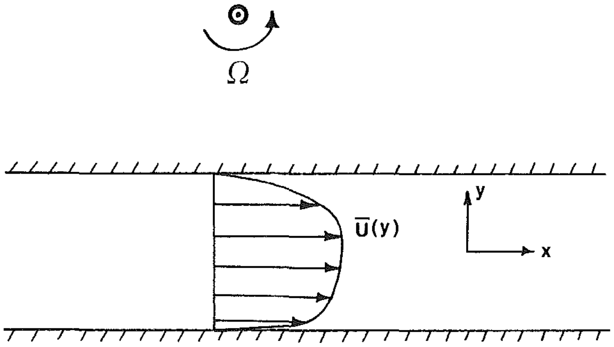


Fig. 8 Fully-developed mean axial velocity profile in a turbulent channel flow subjected to a spanwise rotation (after Lezius and Johnston 1976)

by a simple extrapolation of the stability boundary given in Fig. 3, will be demonstrated numerically in the latter part of this section). It is thus quite clear that a roll instability can occur for Rossby numbers $\mathrm{Ro} > 100$. For this range of Reynolds numbers, such a Rossby number can be induced by the rotation of the earth for the flow of most common liquids (e.g., water) in a channel whose width D is less than 1 ft.

Now, specific numerical calculations will be presented to clearly demonstrate that for $\mathrm{Re} > 100$, there is sufficient inertia so that a roll instability which is induced by the rotation of the earth can have a significant distortional effect on the axial velocity profiles in the interior of the channel. The computed streamlines of the secondary flow for rotating channel flow are shown in Fig. 4 for $\mathrm{Re} = 485$ and $\mathrm{Ro} = 35.7$. It is clear that there is a roll instability at this particular Reynolds and Rossby number in agreement with previous studies (see Fig. 3). The corresponding axial velocity profile along the horizontal centerline of the channel $y = H/2$ is shown in Fig. 5. It is obvious from this graph that the secondary flow has a severe distortional effect on the axial velocity profile, making it asymmetric with the maximum velocity shifted toward the high-pressure side of the channel (i.e., the side of the channel which is farthest from the axis of rotation). From equation (8), it follows that

$$D = \left(\frac{\mathrm{Re}\ \nu}{2\Omega\ \mathrm{Ro}} \right)^{1/2} \tag{11}$$

and hence, given that we are considering the earth's rotation for which

$$\Omega = 7.292 \times 10^{-5} \mathrm{rad/sec}, \tag{12}$$

it is possible to explicitly calculate the width D of the channel corresponding to a particular choice of Re and Ro once the kinematic viscosity ν of the fluid is specified. The following results are obtained (at room temperature) for this case of $\mathrm{Re} = 485$ and $\mathrm{Ro} = 35.7$:

$D = 46.3$ in. (for air)

$D = 12.1$ in. (for water)

$D = 3.84$ in. (for mercury).

It is thus clear that for most common liquids, the classical parabolic velocity profile obtained for laminar channel flow (if the effect of the earth's rotation is neglected) can be in serious error if the channel has a width greater than a few inches.

Now, the computed solution for a substantially higher Reynolds number will be presented for which the results are even more dramatic. In Fig. 6, the computer generated secondary flow streamlines are shown for rotating channel flow at a Reynolds number $\mathrm{Re}=1860$ and Rossby number $\mathrm{Ro}=667$. There is a roll instability present as would be expected on physical grounds from a simple extrapolation of the linear stability analyses presented in Fig. 3. This roll instability has a severe distortional effort on the axial velocity profile along the horizontal centerline of the channel which shows a dramatic departure from the classical parabolic profile as can be seen in Fig. 7. To be more specific, this axial velocity profile is highly asymmetric with the maximum velocity shifted toward the high pressure side of the channel. By utilizing equations (11)-(12), it is possible to calculate the width of the channel (for a given fluid) that corresponds to these flow parameters. For $\mathrm{Re}=1860$ and $\mathrm{Ro}=667$, we have the following results at room temperature:

$$D=20.99 \text{ in. (for air)}$$

$$D=\ 5.50 \text{ in. (for water)}$$

$$D=\ 1.74 \text{ in. (for mercury).}$$

These results are rather dramatic and surprising in that they indicate that the earth's rotation can induce a roll instability which has a severe distortional effect on the classical parabolic velocity profile in a laboratory channel flow with a characteristic width less than 2 in.[3]

At this point, some comments should be made as to why the surprising and striking results presented above have not been observed in previously conducted laboratory experiments on channel flow. To begin with, a considerable number of such laboratory experiments have been conducted with *air* (Johnston 1984) in channels with a characteristic width $D\leq 2$ in., so that the fluid velocity would not have to be extremely small to maintain laminar flow conditions. The results presented in this paper clearly demonstrate that this size channel is an order of magnitude smaller than that required for the earth's rotation to cause a roll instability and hence the effects shown in Fig. 5 and Fig. 7 would not have arisen. A roll instability would not occur in a laminar channel flow with water if $D<5$ in. (this value for D can be obtained by an extrapolation of the stability results given in Fig. 3) which constitutes a channel width that is larger than those considered in previously published studies. Of course, in the absence of a roll instability, the secondary flow that results from the earth's rotation will be confined to a region which is far enough removed from the channel centerline (see Fig. 1) so that it will have a negligible effect on the axial velocity profile there.

Finally, the extension of these results to the turbulent regime will be briefly discussed. The experiments on turbulent channel flow subjected to the spanwise rotation depicted in Fig. 1 indicate that a roll instability occurs at a Rossby number $\mathrm{Ro}=22.7$ for Reynolds numbers Re between 6000 and 35,000 (see Johnston, Halleen, and Lezius 1972 and Lezius and Johnston 1976). These experiments indicated that this roll instability can have a profound distortional effect on the usual symmetric and flat velocity profile observed in turbulent channel flow (see Fig. 8). By utilizing equations (11)-(12), it is a simple matter to show that the earth's rotation can cause distortions in the axial velocity profile similar to those shown in Fig. 8 for the following values of D:

$$D>17.03 \text{ ft. (for air)}$$

$$D>\ 4.46 \text{ ft. (for water)}$$

$$D>1.41 \text{ ft. (for mercury)}$$

at room temperature. These channel widths are approximately an order of magnitude larger than those for the corresponding laminar case. Hence, the earth's rotation would have a negligible effect on the overwhelming majority of turbulent channel flows which could be constructed in the laboratory. Nevertheless, it could play a role in certain special cases (e.g., a large and low speed wind tunnel or an unusually large cooling duct).

4 Conclusion

Results have been presented in this paper which clearly indicate that the rotation of the earth can induce a roll instability in the laminar flow of fluids in channels of moderate width (i.e., for D substantially less than 1 ft for the flow of certain liquids). This roll instability was shown to have a severe distortional effect on the axial velocity profiles at the horizontal centerline $y=H/2$ of the channel which are obtained when the effect of the earth's rotation is neglected. Hence, the usual assumption of neglecting the effect of the earth's rotation in the calculation of a laboratory channel flow in the laminar regime could lead to serious errors. This surprising result, which was documented by comparisons with previously conducted stability analyses on rotating channel flow, is similar to that which was obtained by Benton [1956] several decades earlier.

Future theoretical research is needed on the effect of the earth's rotation on turbulent channel flow. Such a study, which would be quite difficult because of the lack of reliable turbulence models for rotating flows, could be of considerable value. Since the results of this paper indicate that the rotation of the earth could play an important role in determining the structure of the turbulent flow of water in channels with a width $D>5$ ft, there could be some important applications in problems of interest to civil engineers. Future research which accounts for different orientations of the channel would also be of value. In this paper, the channel was aligned east-west (with the sides of the channel parallel to the axis of rotation of the earth) for simplicity since this is the orientation which yields the most drastic effect. A study along these lines, which would incorporate the effects of latitude and variable axial alignment of the channel, is highly detailed and beyond the scope of the present study.

Considerable evidence has been presented which suggests that the rotation of the earth can play a non-negligible role in the calculation of certain standard laboratory flows. This rather surprising result, which seems to have gone unnoticed since the initial work of Benton [1956], warrants substantial future research from both a theoretical and experimental standpoint.

Acknowledgment

The author is indebted to Professor S. Thangam for his assistance with some of the computations presented in this paper.

References

Benton, G. S., 1956, "The Effect of the Earth's Rotation on Laminar Flow in Pipes," ASME JOURNAL APPLIED MECHANICS, Vol. 23, pp. 123-127.

Hart, J. E., 1971, "Instability and Secondary Motion in a Rotating Channel Flow," *J. Fluid Mech.*, Vol. 45, pp. 341-351.

Johnston, J. P., Halleen, R. M. and Lezius, D. K., 1972, "Effects of a Spanwise Rotation on the Structure of Two-Dimensional Fully Developed Turbulent Channel Flow," *J. Fluid Mech.*, Vol. 56, pp. 533-557.

Johnston, J. P., 1984, Private communication.

Lezius, D. K. and Johnston, J. P., 1976, "Roll-Cell Instabilities in Rotating Laminar and Turbulent Channel Flow," *J. Fluid Mech.*, Vol. 77, pp. 153-175.

Pedlosky, J., 1979, *Geophysical Fluid Dynamics*, Springer-Verlag, Berlin.

Speziale, C. G., 1982, "Numerical Study of Viscous Flow in Rotating Rectangular Ducts," *J. Fluid Mech.*, Vol. 122, pp. 251-271.

Speziale, C. G. and Thangam, S., 1983, "Numerical Study of Secondary Flows and Roll-Cell Instabilities in Rotating Channel Flow," *J. Fluid Mech.*, Vol. 130, pp. 377-395.

---

[3] This channel width $D<2$ in. corresponds to the flow of mercury and could have important ramifications in magnetohydrodynamic experiments (e.g., the Hartmann problem).

D. Nixon

Y. Liu[1]

Nielsen Engineering & Research, Inc.,
Mountain View, CA 94043

The Mechanisms of Determining Shock Locations in One and Two Dimensional Transonic Flows

The mechanism that locates a shock wave in a transonic flow in one and two dimensions is examined. It is found that in one dimension the shock is located by specifying the downstream pressure whereas in two dimensions the shock is located by the application of an entropy condition at the sonic line.

Introduction

The discipline of computational fluid dynamics (CFD) has grown enormously over the last decade. Flows that are quite complex can now be computed routinely. One such type of flow is transonic flow with shock waves and it is one aspect of such a flow that is considered here.

In spite of the progress of CFD some confusion regarding the nature of transonic flows is apparent, in particular the mechanism that fixes waves. The present note is concerned with the mechanisms that determine the shock wave location in computational solutions of the Euler equations in both one and two dimensions. It is found that shock locations are determined by entirely different means in each case so caution is advised in the use of one dimensional analysis in two dimensional transonic flows.

Analysis

a **One Dimensional Flow.** For one dimensional nozzle flows the conservation of mass equation can be written as

$$\rho US = \rho_\infty U_\infty S_\infty = A \qquad (1)$$

where A is a constant and S is the area distribution. The density is given in terms of the pressure by

$$\rho = p^{\frac{1}{\gamma}e^{-\Delta S/C_p}} \qquad (2)$$

where the reference values of p and ρ are assumed unity, ΔS is the entropy jump through a shock wave and C_p is the specific heat at constant pressure.

From the usual density/velocity relation

$$U = U_\infty \left\{ \frac{2}{(\gamma-1)M_\infty^2}[1 - \rho^{\gamma-1}\exp(\Delta S/C_v)] + 1 \right\}^{1/2} \qquad (3)$$

Combining equations (1), (2), and (3) then gives

$$\rho US = SU_\infty \rho_\infty \left\{ \frac{2}{(\gamma-1)M_\infty^2} \left[1 - p^{\frac{\gamma-1}{\gamma}} \right. \right.$$

$$\left. \left. \bullet \exp[-\Delta S(2\gamma-1)/C_p \right] + 1 \right\}^{1/2} p^{\frac{1}{\gamma}} \exp(-\Delta S/C_p) = A \qquad (4)$$

The only means of entropy production is the shock wave and hence ΔS is constant downstream of the shock. Thus if p is specified at the downstream boundary the conservation of mass equation gives the entropy and hence the shock strength. If the shock strength is known then the location can be found. If p and ΔS are known at the boundary then U, ρ can be found from equations (2) and (3).

If equation (1) is differentiated with respect to x then

$$\frac{\partial \rho U}{\partial x} = \frac{-A}{S^2}\frac{\partial S}{\partial x} \qquad (5)$$

Ahead of the shock wave the flow is isentropic. For isentropic flow the density is given in terms of the Mach number M by

$$\rho = \rho_o \left\{ 1 + \frac{\gamma-1}{2}M^2 \right\}^{\frac{-1}{\gamma-1}} \qquad (6)$$

and

$$U = aM = a_o \left\{ 1 + \frac{\gamma-1}{2}M^2 \right\}^{\frac{-1}{2}} M \qquad (7)$$

where a subscript o denotes stagnation conditions.

Combination of equations (1), (6), and (7) thus gives

$$a_o\rho_o \left\{ 1 + \frac{\gamma-1}{2}M^2 \right\}^{\frac{-(\gamma+1)}{2(\gamma-1)}} M = A/S \qquad (8)$$

Differentiation gives

$$a_o\rho_o \left\{ 1 + \frac{\gamma-1}{2}M^2 \right\}^{\frac{1-3\gamma}{2(\gamma-1)}} \{1 - M^2\}\frac{\partial M}{\partial x} = \frac{-A}{S^2}\frac{\partial S}{\partial x} \qquad (9)$$

and hence equation (9) will give an infinite value of $\partial M/\partial x$ at sonic conditions unless $\partial S/\partial x$ is zero, that is, the sonic point is at the nozzle throat. Thus, the bounding of $\partial M/\partial x$ in the flow outside the shock wave simply requires that the sonic point occurs at the nozzle throat; it does not control the shock location. In this paper the bounding of $\partial M/\partial x$ is called the entropy condition for convenience since for an accelerating flow an in-

---

[1] Presently at Computational Fluid Dynamics Branch, NASA Ames Research Center, Moffett Field, CA.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 29, 1985; final revision, May 28, 1985.

finite value of $\partial M/\partial x$ denotes an expansion shock which permits a non-physical decrease in entropy.

b **Two Dimensional Flow.** In potential theory a very effective way to examine the characteristics of the equation is to write it in integral form; a detailed discussion of this procedure is given in Nixon (1976). In the next section this procedure is performed for the Euler equations. Thin airfoil boundary conditions are assumed. The analysis given here is for nonlifting airfoils.

Let a stream function Ψ be defined as

$$\Psi_y = \rho U, \Psi_x = -\rho V \tag{10}$$

Now let a perturbation stream function ψ be defined as

$$\begin{aligned} \Psi &= \rho_\infty U_\infty y + \psi \\ \psi_y &= \rho U - \rho_\infty U_\infty \\ \psi_x &= -\rho V \end{aligned} \tag{11}$$

The governing equation is

$$U_y - V_x = \omega \tag{12}$$

where ω is the vorticity; from Crocco's theorem it is a function of entropy. This equation can be rewritten as

$$\left(\frac{\rho_\infty U_\infty + \psi_y}{\rho}\right)_y + \left(\frac{\psi_x}{\rho}\right)_x = \omega \tag{13}$$

or

$$\nabla^2 \psi = \rho\omega + U\rho_y - V\rho_x = f \tag{14}$$

The tangency boundary condition is

$$\psi_x(x, \pm 0) = -Y_s'(x)\rho(x, \pm 0)U_\infty \tag{15}$$

where $y_s'(x, \pm 0)$ is the airfoil slope. On the wake

$$\psi_x(x, \pm 0) = 0$$

This can be written in integral form using Green's identity to give

$$\begin{aligned} \psi = &\int_{c_\infty} (K\psi_R - K_R\psi)dC_\infty \\ &+ \int_o^\infty (K\Delta\psi_\eta - K_\eta\Delta\psi)d\xi + \int_S\int KfdS \end{aligned} \tag{16}$$

where the kernel function K is the solution of the equation

$$\nabla^2 K = \delta(R)$$

and is given by

$$K = \frac{1}{2\pi} ln[(x-\xi)^2 + (y-\eta)^2]^{1/2}$$

and Δ denotes a jump across the x axis; (ξ, η) are coordinates in the x and y directions, respectively; the domain S is shown in Fig. 1. It is required that the first integral be bounded as $R = \infty$ which gives

$$\psi_R \rightarrow 0 \text{ faster than } \frac{1}{R}$$

Note that the integral around the shock vanishes (Nixon, 1976).

For a symmetric airfoil the integral equation becomes

$$\psi = -\int_o^\infty K_\eta \Delta\psi d\xi + \int_S\int KfdS \tag{17}$$

Differentation of equation (17) with respect to y gives

$$\psi_y = \rho U - \rho_\infty U_\infty = -\int_o^1 K_{\eta y}\Delta\psi d\xi + \int_S\int K_y fdS$$

or, integrating the first integral by parts to give

$$\rho U - \rho_\infty U_\infty = \int_o^1 K_\xi \Delta\psi_\xi d\xi + \int_S\int K_y fdS \tag{18}$$

The upper limit in the first integral is unity since $\Delta\psi_\xi = 0$ downstream of the airfoil. The first integral is known from the tangency boundary condition. The field integral can be integrated by parts to give

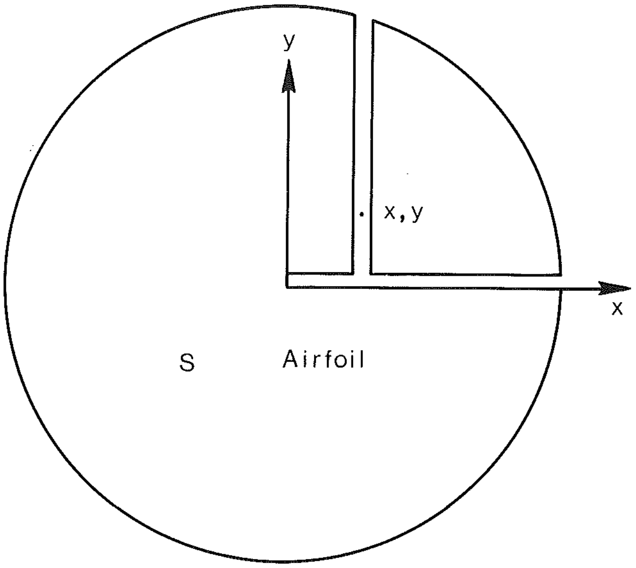


Fig. 1 Domain of integration *S*

$$\begin{aligned} \int_S\int K_y fdS = \frac{1}{2\pi}\Bigg\{ &\int_{-\infty}^{x-\xi} \frac{-F(\xi, +0)}{y^2 + (x-\xi)^2} d\xi \\ &- \int_{x+\xi}^{\infty} \frac{F(\xi, +0)y}{y^2 + (x-\xi)^2} d\xi \\ &+ \int_{-\infty}^{\infty} \frac{F(\xi, -0)y}{y^2 - (x-\xi)^2} d\xi \Bigg\} - \int_S \int K_{y\eta} FdS \end{aligned} \tag{19}$$

where

$$F(\xi, \eta) = \int^\eta f(\xi, \eta')d\eta' \tag{20}$$

If $y \neq 0$ the first part of equation (19) becomes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{[F(\xi, -0) - F(\xi, +0)]y}{y^2 + (x-\xi)^2} d\xi \tag{21}$$

which is zero for a symmetric $F(\xi, \eta)$

If $y \rightarrow 0$ the integral only has a value for ξ close to x and hence

$$\begin{aligned} \int_S \int K_y fdS = &-\lim_{y\rightarrow 0} \lim_{\Delta\rightarrow 0} \lim_{\epsilon\rightarrow 0} \frac{1}{2\pi} \\ &\times \Bigg\{ -F(x, +0)\left[\tan^{-1}\left(-\frac{\epsilon}{y}\right) - \tan^{-1}\left(\frac{-\Delta}{y}\right) \right. \\ &\left. + \tan^{-1}\left(\frac{\Delta}{y}\right) - \tan^{-1}\left(\frac{\epsilon}{y}\right)\right] \\ &+ F(x, -0)\left[\tan^{-1}\left(\frac{\Delta}{y}\right)\right]\Bigg\} \\ &- \int_S \int K_{\eta y} FdS \end{aligned} \tag{22}$$

Taking the limits gives

$$\int_S\int K_y fdS = -\int_S\int K_{\eta y} FdS \tag{23}$$

Hence for a symmetric airfoil

$$\rho U - \rho_\infty U_\infty = \int_0^1 K_\xi \Delta\psi_\xi d\xi - \int_S\int K_{\eta y} FdS \tag{24}$$

This is the integral form of the two dimensional Euler equations. The right-hand side is continuous through a normal shock (Nixon, 1976). Although a much more detailed analysis is given in the quoted reference it is easy to justify this statement by noting that ρU is continuous through a normal shock and hence the right-hand side of equation (24) must also be continuous.

Consider now the case where $x \to \infty$ in equation (24). Let the far field boundary condition be

$$\frac{\partial F}{\partial x} \to 0 \text{ as } x \to \infty \tag{25}$$

Now

$$\int_S\int K_{\eta y} F dS = -\int_S\int K_{\xi x} F dS \tag{26}$$

and

$$\begin{aligned} -\int_S \int K_{\xi x} F dS &= -\int_o^\infty \left[K_x \Big|_{-\infty}^{x-\epsilon} + K_x \Big|_{x+\epsilon}^{\infty} F(x,\eta) \right] d\eta \\ &= -\frac{1}{2\pi} \int_o^\infty \left[\frac{2\epsilon}{\epsilon^2 + (y-\eta)^2} \right] F(x,\eta) d\eta \end{aligned} \tag{27}$$

As $\epsilon \to 0$ this integral only has a value when $(y-\eta)$ is small and hence

$$\begin{aligned} -\int_S \int K_{\xi x} F dS &= 2F(x,y) \lim_{\delta \to 0} \lim_{\epsilon \to 0} \frac{1}{2\pi} \tan^{-1} \frac{\eta - y}{\epsilon} \Big|_{y-\delta}^{y+\delta} \\ &= -F(x,y) \end{aligned} \tag{28}$$

Since the line integral in equation (24) approaches zero as $x \to \infty$, equation (24) becomes, as $x \to \infty$,

$$\rho U - \rho_\infty U_\infty = F(x,y) = \int^y (\rho\omega - \rho_\eta U - \rho_x V) d\eta \tag{29}$$

when equation (14) is used.

Now if the boundary condition as $x \to \infty$ is that all x derivatives are zero, then equation (29) becomes

$$\rho U - \rho_\infty U_\infty = \int^y (\rho\omega + \rho_\eta U) d\eta \tag{30}$$

Where ω is specified by the entropy production at the shock and Crocco's theorem.

Differentiation of equation (30) gives

$$\rho U_y = \rho\omega \tag{31}$$

and ω is a function of entropy. Equation (31) is valid for any shock strength.

If the pressure is specified at the downstream boundary and the reference values of p, ρ are assumed to be unity then

$$\rho = p^{\frac{1}{\gamma}} \exp(-\Delta S / C_p) \tag{32}$$

Using the above relation the energy equation can be written as

$$\frac{\gamma}{\gamma-1} p^{\frac{\gamma-1}{\gamma}} \exp(\Delta S/C_p) + \frac{1}{2}U^2 = \frac{\gamma}{\gamma-1} + \frac{1}{2}U_\infty^2 \tag{33}$$

Note that since the flow is nonlifting V is assumed to be zero as $x \to \infty$. Differentiation with respect to y at constant pressure gives

$$U_y = -\frac{1}{R} \frac{\rho^{\gamma-1}}{U} \exp(\Delta S/C_v) \frac{\partial \Delta S}{\partial y} \tag{34}$$

where R is the gas constant.

From Crocco's theorem the vorticity is given by

$$\omega = -\frac{1}{R} \frac{\rho^{\gamma-1}}{U} \exp(\Delta S/C_v) \frac{\partial \Delta S}{\partial y} \tag{35}$$

and hence if the pressure is specified downstream then equation (31) is consistent. If a similar analysis is performed with the density fixed downstream then

$$U_y = -\frac{\gamma}{R} \frac{\rho^{\gamma-1}}{U} \exp(\Delta S/C_v) \frac{\partial \Delta S}{\partial y} \tag{36}$$

which is not consistent with the result of equation (31). Consequently if a zero gradient boundary condition is imposed the pressure must be specified.

It can be seen from the above analysis that the boundary condition for a specified downstream pressure is satisfied by any shock strength or entropy and, hence, the shock strength is not fixed by the downstream conditions. An alternative procedure must therefore be found.

The integral formulation of the two-dimensional Euler equations, given by equation (24) can be written as

$$\rho U - \rho_\infty U_\infty = G \tag{37}$$

where G is the right-hand side of the integral equation and is continuous except across an oblique shock; G can be thought of as the change in area of the streamlines. It has been noted earlier, in connection with equation (24) that G is continuous through a normal shock since ρU is continuous there. If, however, the shock is oblique, ρU (and hence G) is discontinuous. Ahead of the shock the flow is isentropic and hence, if freestream values of ρ and p are unity,

$$\rho U = \gamma \rho^{\frac{\gamma+1}{2}} M \tag{38}$$

where M is the local Mach number. This can be written using the isentropic density relation as

$$\rho U = \sqrt{\gamma} \left\{ \left(1 + \frac{\gamma-1}{2} M^2 \right)^{\frac{-(\gamma+1)}{2(\gamma-1)}} M \right\} \rho_o^{\frac{\gamma+1}{2}} = G + \rho_\infty U_\infty \tag{39}$$

Differentiation with respect to x gives

$$\sqrt{\gamma} \rho_o^{\frac{\gamma+1}{2}} \left\{ \frac{1 - M^2}{\left[1 + \frac{\gamma-1}{2} M^2 \right]^{\frac{3\gamma-1}{2(\gamma-1)}}} \right\} \frac{\partial M}{\partial x} = \frac{\partial G}{\partial x} \tag{40}$$

Hence, if $M \to 1$, $\partial M/\partial x$ is infinite unless

$$\underset{M \to 1}{\text{limit}} \frac{\partial G}{\partial x} \to 0 \tag{41}$$

At an accelerating sonic point $\partial M/\partial x \to \infty$ implies that an expansion shock would exist unless the condition of equation (41) is satisfied. Such an expansion shock would give rise to a physically unacceptable entropy decrease. In addition, at the sonic line

$$G = \sqrt{\gamma} \left(\frac{\gamma+1}{2} \right)^{\frac{-(\gamma+1)}{2(\gamma-1)}} \rho_o^{\frac{\gamma+1}{2}} - \rho_\infty U_\infty \tag{42}$$

The physical implication of equations (41) and (42) is that expansion shocks must be eliminated from the solution, that is, no entropy decrease is allowed at the sonic line. There is only one choice to alter G and that is the shock location. Hence equations (41) and (42) give the shock location; the second unknown is the sonic line location.

Conclusions

An analysis to determine the mechanism to locate shock waves in transonic flow has been performed. It is found that the mechanism for locating the shock wave in two dimensions is different from that in one dimension. In two dimensions the shock location is fixed by the behavior at the sonic line whereas in one dimension it is fixed by the downstream boundary conditions.

References

Nixon, D., 1976, Calculation of Transonic Flows Using Integral Equation Methods, Ph.D. Thesis, University of London.

Mohamed Gad-el-Hak
Senior Research Scientist,
Flow Research Company,
Kent, Wash. 98032
Mem. ASME

The Response of Elastic and Viscoelastic Surfaces to a Turbulent Boundary Layer

The unstable response of elastic and viscoelastic surfaces to a turbulent boundary layer was experimentally investigated in an 18-m towing tank. The compliant surface deformation was measured using a remote optical technique. The "Laser Displacement Gauge" employs a Reticon camera equipped with a linear array of 256 photodiodes spaced 25 microns apart. The device was used to measure the characteristics of two classes of hydroelastic instability waves that form on elastic or viscoelastic surfaces as a reuslt of the interaction with a turbulent boundary layer. The instability waves developing on an elastic surface are symmetric and have a relatively high phase speed and a small wavelength, as compared to the slow and highly nonlinear "static-divergence" waves observed on the viscoelastic surface. The experimentally determined wave characteristics are compared to existing theories on compliant surface instabilities.

1 Introduction

The motion of a fluid over a surface that complies to the flow offers the potential for a rich variety of fluid/surface interactions. Compliant surfaces are currently finding many engineering applications, such as sound absorption in aeroengines, vibration reduction in Naval vessels, and noise shielding in sonar arrays. Moreover, intensive research is currently being conducted to find compliant surfaces that will reduce the skin-friction drag on moving vehicles.

The design of a compliant coating to achieve a particular objective is a complex task requiring the determination of the surface response to a specific flow disturbance. This response is excited by the hydrodynamic forces and results in a surface motion that in turn acts on the flow field near the interface. Waves that form on the compliant surface can be stable, unstable, or neutral.

There exists a need for the development of reliable techniques to measure the compliant surface response under a variety of flow conditions. Bushnell, Hefner and Ash (1977), in their excellent review article, state that "extensive wall motion measurements must be made before any theoretical approach to the problem can be reasonably validated." The device needed to measure the surface deformation should be accurate, have a fast response, and not interfere with the observed phenomenon. Very few such devices exist today. Grosskreutz (1971) used a schlieren apparatus to measure the motion of a homogeneous but nonisotropic compliant surface made of rubber and subjected to a turbulent boundary layer in a water tunnel. He computed the frequency and the wavenumber dependence of the flexibility of the compliant wall using the third-octave-spectra of the surface motion. Ash et al. (1975) used a similar schlieren method to provide flash photos of a compliant surface deformation in a wind tunnel. Dinkelacker et al. (1977) placed a 97-mm pressure transducer containing several hundred membranes under a turbulent boundary layer. The device served as the mirror in a Michelson interferometer. High-speed photographs of the fringe patterns in the interferometer were used to compute the dimensions and the speed of convected turbulent pressure fluctuations. More recently, Rathsam et al. (1983) measured the "pre-instability," microscopic surface motion on a PVC plastisol in a turbulent boundary layer. Their laser/optics system sensed the instantaneous slope and the frequency of motion on the compliant surface where a focused laser beam was reflected. This device is incapable of directly measuring the amplitude of the surface motion. However, Rathsam et al. (1983) inferred the amplitude from the measured slope spectra by assuming a dispersion relation for the compliant surface response.

Unstable, flow-induced deformations have been observed experimentally on viscoelastic surfaces (Hansen et al., 1980a; Gad-el-Hak et al., 1984). No corresponding experimental data are available for the hydroelastic instability that is theoretically predicted to exist on an elastic surface (Benjamin, 1963).

In the present investigation, a remote optical technique was used to measure the flow-induced motion of a compliant surface. The technique is particularly suited for studying the two classes of hydroelastic instability waves that are theoretically predicted to form on an elastic or a viscoelastic surface as a result of the interaction with a turbulent boundary layer. The wave's amplitude, wavelength and phase speed are directly measured with this linear device. Our experiment

---

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, New Orleans, December 9-14, 1984, Symposium on Flow-Induced Vibrations.

Discussion on this paper should be addressed to the Publishing Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July, 1984; final revision, March, 1985.

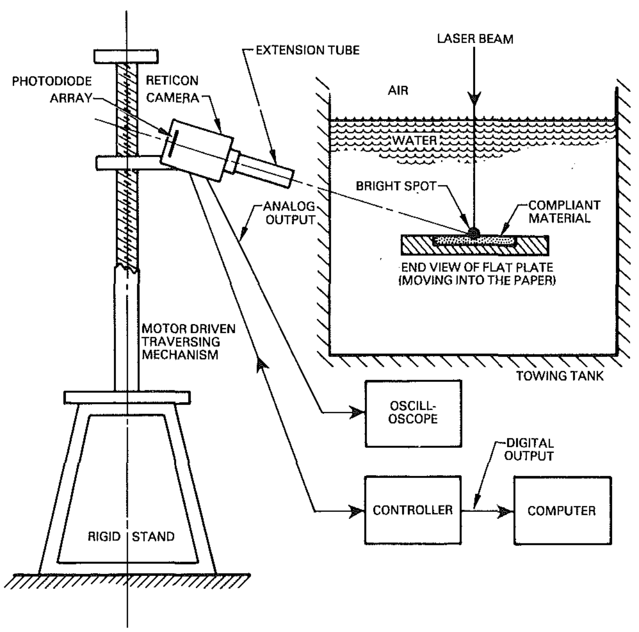


Fig. 1 Setup for the Reticon camera

using viscoelastic coatings have been reported by Gad-el-Hak et al. (1984). Here, we present new results using elastic coatings. The two publications provide a data base against which existing theories on the stability of both kinds of surfaces can be validated.

2 Experimental Equipment and Procedure

2.1 The Laser Displacement Gauge. The laser displacement gauge (LDG) is a remote optical device used in the present investigation to measure the vertical displacement of the compliant surface. The technique was originally developed by Liu et al. (1982) and Liu and Lin (1982) for measuring wind waves. Its first use in measuring compliant surface deformation was reported by Gad-el-Hak et al. (1982; 1984). A schematic of the setup is shown in Fig. 1. The system employs a Reticon camera (Model LC 600V) driven by a controller (Reticon Corporation, Model RS605). An optical interface is created at the surface of the compliant material, which contains minute amounts of Rhodamine-6G fluorescent dye,[1] by projecting a vertical beam of a 5-watt argon-ion laser (Spectra Physics, Model 164-05) having a diameter of 1 mm. The displacement of this optical interface is measured by electronically scanning the photodiode array housed in the Reticon camera.

The axis of the photodiode array is aligned at a given angle with respect to the vertical laser beam above the fluid/compliant coating interface. The optical interface is imaged onto the photodiode array via a set of lenses and extension tubes. The linear photodiode array is composed of 256 elements spaced 25 μm apart. The aperture width of the array is also 25 μm. The spatial resolution, which is the same in both the vertical and longitudinal directions, depends on the field of view. For example, the spatial resolution is 0.01 cm for a field of view of 2.5 cm. In this case, the horizontal spatial resolution is only about one-tenth the diameter of the laser beam. The scanning rate of the array ranges from 0.4 to 40 ms. The LDG is a digital device with practically no electronic drift. The digital output from the controller is a time series of integers from 1 to 256 updated at a frequency of the scanning rate. Each integer corresponds to the *n*th photodiode on which the optical interface is imaged during each scan. The digital output is recorded and analyzed on-line with a NOVA minicomputer system.

Calibration of the LDG is made by displacing the Reticon camera, which is fixed on an accurate traverse mechanism, to several vertical positions with predetermined increments. A second-degree polynomial is best-fitted through the calibration points to account for nonlinearity resulting from the aberration of the optical lenses. The ratio of the coefficients of the nonlinear and linear terms was typically 10^{-4}. For practical purposes, the displacements may be considered to be linearly proportional to the LDG output.

The Reticon camera is mounted so that it looks down onto the compliant surface at a nearly horizontal angle (about 15 deg above horizontal). This arrangement minimizes blockage of the optical interface by the wave crests between the laser beam and the tank wall on the side where the camera is mounted. This blockage occurs most often near the troughs of the waves, where the wave profiles are relatively smooth. Whenever a blockage occurs, the photodiode array loses its object (i.e., the optical interface) and the maximum diode number of 256 is registered by the controller. Therefore, either a sharp jump or a sharp spike, depending on the duration of the blockage, appears on the measured wave profiles. To remove the sharp jumps or spikes, the computer was programmed to replace them with a straight line that connects the points before and after each jump or spike.

In the present experiment, the laser displacement gauge was set to have a frequency response of 1 kHz and to resolve vertical displacements as low as 0.002 cm.[2] The surface deformations were also recorded using a 16-mm movie camera moving with the plate. For the elastic surface, the camera was mounted to the side to capture a side view of the instability waves. For the viscoelastic surface, a top view was more suitable for observing the instabilities developing on such a surface.

2.2 Flow Facility. The Flow Research 18-m towing tank was used in the present experiments. The 1.2-m-wide, 0.9-m-deep water channel has been described by Gad-el-Hak et al. (1981). To generate a turbulent boundary layer, a flat plate is rigidly mounted under a carriage that rides on two tracks mounted on top of the towing tank. During towing, the carriage is supported by an oil film to ensure a vibrationless tow, having an equivalent freestream turbulence of about 0.1 percent. The carriage is towed by two cables driven through a reduction gear by a 1.5 hp Boston Ratiotrol motor. The towing speed is regulated within an accuracy of 0.1 percent. For the present study, the system was able to achieve towing speeds between 20 and 140 cm/s.

The flat plate used in the present experiment had an aluminum frame that provided a flat bed for the Plexiglas working surface. The gaps in the aluminum frame were filled with lightweight styrofoam, and the frame was painted with marine enamel to prevent corrosion. The whole structure was buoyant in water and was flat to within 0.2 mm. Care was taken to avoid leading-edge separation and premature transition by having an elliptic leading edge and an adjustable lifting flap at the trailing edge. The flap was adjusted so that the stagnation line near the leading edge was located on the working surface, which was smooth and was 210 cm long by 106 cm wide. A 45 cm by 95 cm well was built into the working surface for placing compliant materials of up to 1 cm in thickness.

Trips were used to generate a fully-developed turbulent boundary layer. The trips were brass cylinders with 0.32-cm diameters and 0.25-cm heights placed 20 cm behind the leading edge and having their axes perpendicular to the flat plate. During towing, the plate and the movie camera moved

---

[1] About 0.05 percent by weight.

[2] The field of view is then about 0.5 cm.

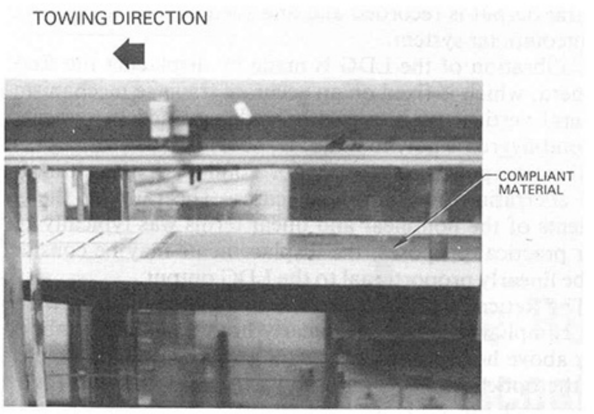


Fig. 2 Side view of the flat plate with the compliant coating insert

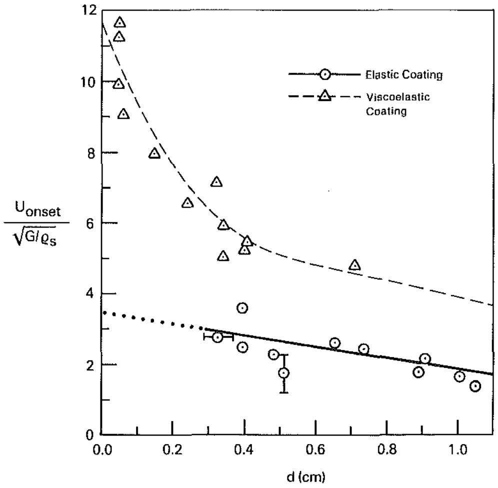


Fig. 3 Onset speed dependence on thickness

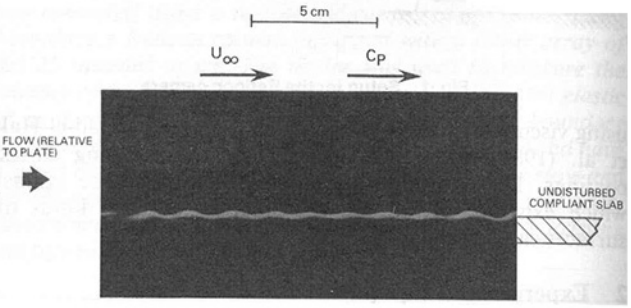


Fig. 4 Instability waves on the elastic surface $U_\infty = 70$ cm/s, $U_\infty/\sqrt{G/\rho_s} = 1.8, d = 1.05$ cm

at a speed U_∞, while the Reticon camera and the vertical laser beam were fixed in space.

2.3 Compliant Material. A nearly ideal elastic compliant surface and an incompressible viscoelastic one were used in the present investigation. The elastic coating was made of commercially available Knox gelatin. The gelatin powder was dispersed in boiling water, followed by the addition of an equal amount of water at room-temperature. The concentration of the gelatin was varied in the range of 1 to 6 parts of weight of gelatin per 100 parts of water. The mixture was poured into the well in the flat plate and allowed to gel for 16 hours before using for a maximum of eight hours, then a new coating was formed for the next series of runs. Care was taken to ensure that the compliant surface was smooth and flush with the rest of the Plexiglas working surface. Figure 2 is a side view of the flat plate containing the compliant surface submersed in the water tank.

Whenever a new coating was poured, a 0.6 cm × 10 cm × 10 cm sample was produced from the same mixture to measure the modulus of rigidity, G. The shear modulus of rigidity was measured with an automated strain gauge/LVDT device that subjected the sample to a prescribed shear force and then measured the displacement. The force-versus-displacement curve was always linear in spite of the fact that displacements as high as 50 percent of the thickness were used. The modulus value was quite sensitive to small differences in the mixing process and ranged in value from 400 dyne/cm^2 at the lowest concentration used to 25,000 dyne/cm^2 at the highest concentration.

The viscoelastic coating used in the present investigation was a plastisol gel made by heating to 160°C a mixture of polyvinyl chloride resin (PVC), dioctyl-phthalate (plasticizer), and dibutyl tin maleate (stabilizer). The mixture was poured in a heated aluminum pan and allowed to cool gradually to complete the gelation process. The pan was then placed inside the well in the working surface of the flat plate, and its height was adjusted from the bottom to ensure a flush, smooth surface. Unlike the gelatin, the PVC plastisol solidified rather quickly, particularly when the percentage of PVC in the mix was increased.

Several recent studies have used similar PVC plastisols to study their interactions with laminar and turbulent flows [Hansen and Hunston (1974a; 1974b; 1976; 1983), Hansen et al. (1980a; 1980b); Hoyt (1981)]. In the present experiments, the modulus of rigidity of the PVC plastisol was varied in the range of 50 to 125,000 dyne/cm^2 by changing the percentage of PVC from 3 to 25 percent in the mixture. The stabilizer was always 10 percent of the PVC by weight. To check the viscoelasticity of the plastisol, a dynamic test was conducted on a sample using the strain gauge/LVDT device. A shear stress was applied until the system came into equilibrium, and then the stress was suddenly released. The value of the subsequent strain was observd using the LVDT and a Nicolet digital oscilloscope (Model 4094). The observed time history indicated that the plastisol could be modeled as a viscoelastic solid of the generalized Kelvin type (Jaeger and Cook, 1976). The time constant of the plastisol decreased as the percentage of PVC in the mix increased. For the three-percent plastisol, the relaxation time was about 1 second, indicating a relatively strong damping. Since the mechanical properties of the material change considerably during gelling, the compliant surface and the sample were allowed to sit in air for 16 hours before testing and using in the tank. Typically, a coating was used for eight hours before a new one was formed for the next series of tests.

3 Results

The slowest traveling free wave speed on the surface of either an elastic or a viscoelastic solid is given approximately by the transverse wave speed $c_t = \sqrt{G/\rho_s}$ where G is the shear modulus of rigidity[3] and ρ_s is the density of the solid, which was very close to that of water for all the coatings used in the present investigation. Whenever the freestream velocity, U_∞, becomes sufficiently large compared to c_t, unstable waves appear on the solid surface. The onset speed of the two classes of hydroelastic instability was determined from the Reticon camera records and from visual observation of the elastic and viscoelastic surfaces. The results for different moduli of rigidity and different thicknesses are shown in Fig. 3. For a particular coating thickness, the onset of instabilities on the

---

[3] For an elastic solid, the shear modulus is real, whereas for a viscoelastic solid it is complex and frequency-dependent; the real part is the shear storage modulus G and the imaginary part is the shear loss modulus G'.

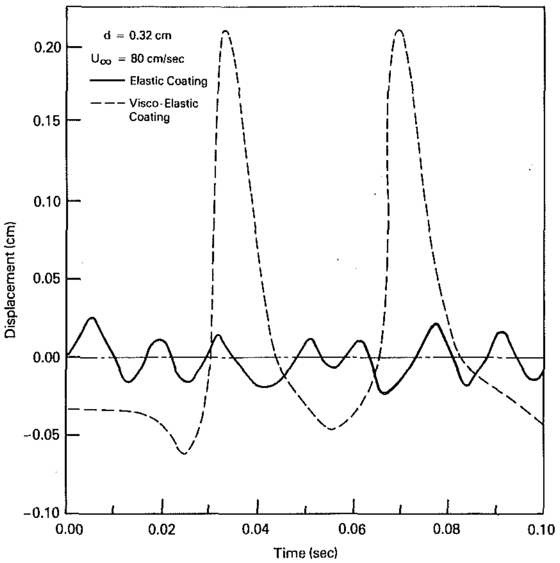


Fig. 5 Typical surface displacement

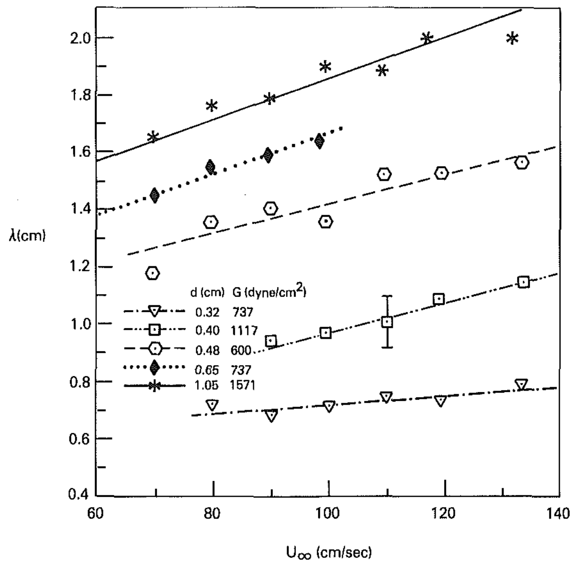


Fig. 6(*a*) Dimensional plots

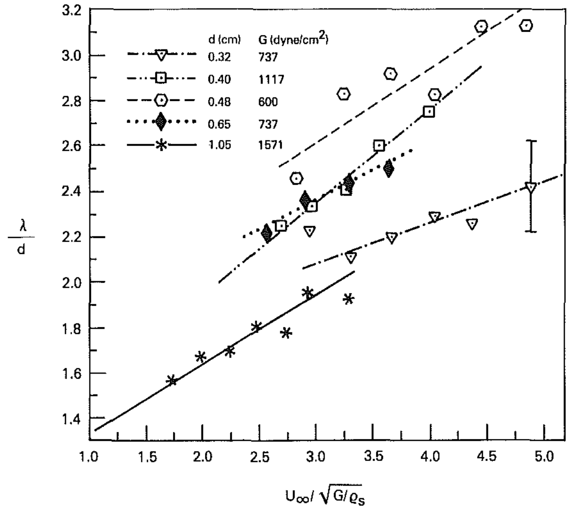


Fig. 6(*b*) Normalized plots

Fig. 6 Wavelength dependence on thickness and flow speed

viscoelastic coating is generally observed at higher ratios of the onset speed to the transverse wave speed than for the elastic coating. For both coatings, U_{onset}/c_t decreases as the thickness, d, increases. In other words, thick surfaces are more susceptible to hydroelastic instability than thin ones.

To visualize the instability waves on the elastic surface, a vertical sheet of laser light parallel to the towing direction was projected onto the flat plate. The laser excited the fluorescent dye mixed with the gelatin, and a bright, horizontal line formed at the undisturbed compliant/fluid interface. When a wave is propagated on the elastic surface, the bright line deformed correspondingly. A 16-mm movie camera, outside the tank but moving with the flat plate, recorded the motion of the bright line, as shown in the side view depicted in Fig. 4. A sketch of the undisturbed compliant slab is included in the figure for reference. In this run, the towing speed was $U_\infty = 70$ cm/s, the modulus of rigidity was $G = 1571$ dyne/cm^2, and the coating thickness was $d = 1.05$ cm. The highly asymmetric waves forming on the viscoelastic surface were more readily visualized from the top, using conventional flood lights for illumination (Fig. 5 of Gad-el-Hak et al., 1984).

Typical examples of the instability waves on the elastic and the viscoelastic coatings, as recorded by the Reticon camera, are shown in Fig. 5. Both coatings have a thickness of $d = 0.32$ cm, and the freestream speed was $U_\infty = 80$ cm/s. The modulus of rigidity for the elastic coating was $G = 740$ dyne/cm^2 and for the viscoelastic coating, $G = 50$ dyne/cm^2. A well-defined average wavelength and amplitude are apparent. The elastic waves have smaller wavelength and amplitude as compared to the waves excited on the viscoelastic surface. The peaks of the waves on the viscoelastic coating are sharp and the valleys are shallow and broad, while the elastic waves are more or less symmetric. The waveform on the viscoelastic surface appears to be nonsinusoidal, with higher harmonics phase-locked with the fundamental wave.

In the viscoelastic coating case, small-amplitude waves always grew very rapidly to large amplitude waves; consequently, a wave train of small amplitude could never be recorded. With the elastic coating at low flow velocities, small-amplitude waves existed. The growth mechanism for the two kinds of instabilities appears to be different.

The average wavelength was measured from the cine films. By averaging over several frames, the statistical scatter of this random phenomenon was reduced to a standard deviation of less than 20 percent. For both the elastic surface and the viscoelastic surface, the wavelength has a strong dependence upon the depth of the coating and upon the flow speed. The results for the viscoelastic surface were reported by Gad-el-Hak et al. (1982; 1984). Here, we primarily present the results for the elastic surface. Figure 6(*a*) is a dimensional plot of the wavelength for five gelatin coatings. The elastic waves have about half the wavelength of that of the viscoelastic waves for comparable moduli and thicknesses, indicative of the two different types of instability. The wavelength increases as the flow speed and the coating thickness increase for both coatings. However, a maximum wavelength is observed for each of the five viscoelastic coatings (see Fig. 13 of Gad-el-Hak et al., 1984). The flow speed at which this maximum is observed coincided with the appearance of a three-dimensional wave structure superimposed on the normally two-dimensional viscoelastic waves. As the velocity continued to increase, small irregularities along the wave crests seemed to spawn new crescent-shaped waves downstream. As these additional waves appeared over the viscoelastic surface, the average wavelength decreased. No similar phenomenon was observed for the elastic coating at the speeds achieved in the present experiments.

The data of Fig. 6(*a*) are normalized with the thickness, d, and the transverse wave speed, c_t, and are replotted in Fig.

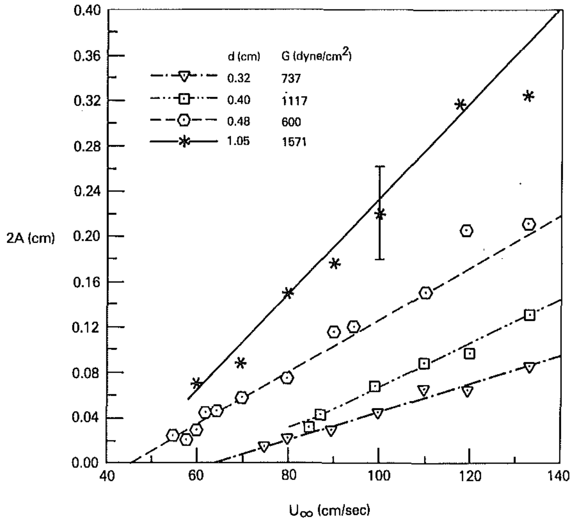


Fig. 7(*a*) Dimensional plots

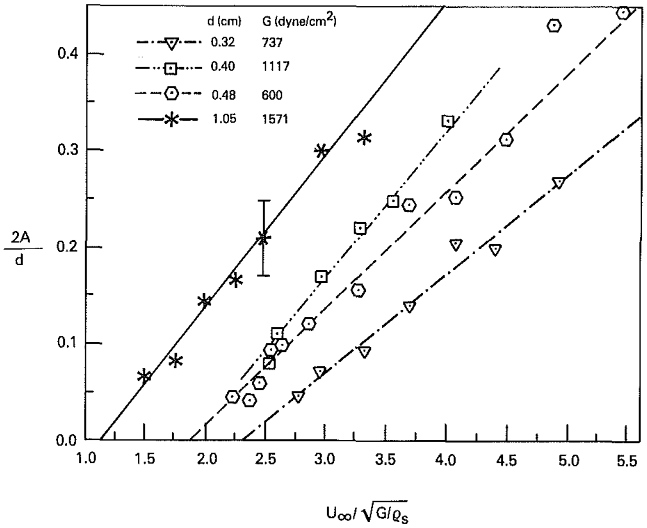


Fig. 7(*b*) Normalized plots

Fig. 7 Wave-amplitude dependence on thickness and flow speed

6(*b*). Scaling the wavelength with the thickness reduces the range of the elastic coating data but does not appear to collapse the data, suggesting that another length scale may be relevant to the problem. For the viscoelastic coatings, the data collapse reasonably well for the two-dimensional waves, while the three-dimensional wave data do not collapse (Fig. 12 of Gad-el-Hak et al., 1984).

The average peak-to-trough amplitude, 2A, was computed from the Reticon camera's output. The results are shown in Figs. 7(*a*) and 7(*b*) for the elastic coating's dimensional and normalized data, respectively. The peak-to-trough amplitude for the waves on both the elastic and viscoelastic surfaces increases monotonically with both the thickness and the flow speed. The elastic wave data do not collapse when the amplitude is normalized with the coating thickness, while the viscoelastic waves do scale with the thickness indicating, perhaps, that the maximum amplitude is limited by the thickness (Fig. 15 of Gad-el-Hak et al., 1984). However, confidence in the latter result should be tempered by the limited amount of data in the figure.

As mentioned before, very few small-amplitude waves were ever observed on the viscoelastic coating. No measurable surface deformation was observed as long as the velocity was below the onset speed; immediately above the threshold velocity, waves with amplitudes of typically $2A/d=0.5$ appeared. With the elastic coating at velocities near onset, small-amplitude waves existed. The growth rate for the two waves is evidently different.

One of the greatest differences between waves on the elastic and the viscoelastic coatings is the phase speed. The phase speed, c_p, was determined from the record of the Reticon and the movie cameras using the relation:

$$\lambda = P(U_\infty - c_p),$$

where λ is the wavelength measured from the ciné films, P is the period measured from the Reticon camera's record, and U_∞ is the flow speed. As shown in Fig. 8, the phase speed for the elastic waves is between 25 and 50 percent of U_∞. As noted by the error bars, there is some uncertainty in the data. Nevertheless, it appears that the phase speed for elastic waves is a constant percentage of U_∞ independent of U_∞/c_t. On the other hand, the viscoelastic waves have an extremely low phase speed compared to other characteristic velocities in the fluid. The maximum value of c_p for these waves was five

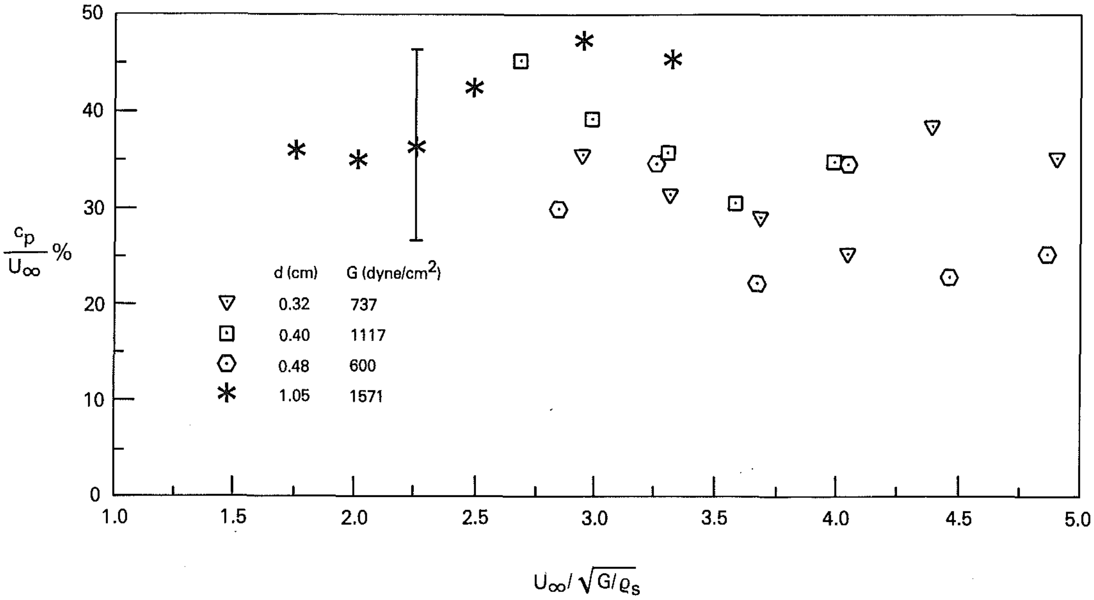


Fig. 8 Phase speed for the unstable waves

percent of U_∞ (Fig. 14 of Gad-el-Hak et al., 1984). For the waves on the viscoelastic surface, the phase speed increases as the flow speed or the thickness increases. The dependence on flow speed for these coatings appears to be given approximately by a power law:

$$c_p \sim U_\infty^{2.6}$$

4 Discussion

A solid in a vacuum can sustain free surface waves that may be modeled as a linear combination of waves having displacements perpendicular to or parallel to the propagation direction. These are called transverse and longitudinal displacement waves, respectively. For a linear-elastic solid, the propagation velocity of the transverse waves is $c_t = \sqrt{G/\rho_s}$, and that of the longitudinal waves is $c_l = \sqrt{(\lambda + 2G)/\rho_s}$, where G and λ are elastic constants and ρ_s is the density of the solid. For a nearly-incompressible solid, $\lambda >> G$ and $c_l \rightarrow \infty$. The free surface wave dispersion relationship for a finite thickness solid has been reported by Gad-el-Hak et al. (1984).

To determine the effects of the fluid motion on the compliant surface, the analysis should be extended to include the surface stresses induced by the fluid moving over the compliant coating. Some general aspects of this case have been addressed by, among others, Benjamin (1960; 1963), Landahl (1962), and Kaplan (1964). Benjamin and Landahl have conducted stability analyses and have established that three types of instability waves may exist. The first type, labeled Class A, is a wave that is destabilized by the addition of dissipation or damping in the system. Duncan et al. (1982) have suggested that pressure phase lags transfer energy from the flow to the interfacial wave system, thus stabilizing these waves. Static-divergence waves, commonly observed on viscoelastic surfaces excited with a turbulent boundary layer of sufficient strength, appear to be members of this class. The second type, Class B waves, is stabilized by damping and destabilized by pressure effects, as for example in the case of wind waves. The third type, Class C, corresponds to a Kelvin-Helmholtz type of instability, where the waves grow or decay primarily through reversible processes. Kaplan (1964) has computed solutions for specific cases.

Experimentally, conditions have been identified in which flow-induced deformations occur on a viscoelastic surface. Boggs and Hahn (1962) were probably the first to point to the existence of a large-amplitude, spanwise wave structure on a compliant surface/fluid interface due to the fluid motion. These "static-divergence" waves (Class A instability) appeared after the freestream speed exceeded an onset velocity threshold. The term "static divergence" is derived from the analogous static instabilities that precede flutter on a flat plate exposed to a high-speed flow (Weaver and Unny, 1970; 1973). In a series of experiments, Hansen & Hunston (1974a; 1974b; 1976; 1983) and Hansen et al. (1980a; 1980b) established several quantitative characteristics of the static-divergence waves, such as the conditions for their initiation, their propagation speed and their influence on hydrodynamic drag. Gad-el-Hak et al. (1982; 1984) presented definitive data on the instabilities of a viscoelastic coating.

No corresponding experimental data are available for Class B instabilities. It is anticipated from the theoretical work of Benjamin (1963) and others that these instabilities may appear on compliant surfaces having little or no damping. Hence, the use of elastic coatings in the present investigation was intended to provide a data base for existing theories on the stability of elastic coatings.

The complexity of the problem of fluid/solid interaction necessitates several restrictive assumptions in formulating a theory. Assuming that the surface stress of primary importance due to the fluid is pressure, the effect of the fluid motion on the compliant surface can be simply modeled by considering the basic flow over the coating as inviscid and unsheared, and, hence, using potential flow theory to determine the surface pressure in terms of the surface displacement. Duncan et al. (1985) have recently explored the dispersion relation for a one-layer, viscoelastic solid. Duncan and Hsu (1984) extended the one-layer analysis to determine the response of a two-layer coating to pressure disturbances from a turbulent boundary layer. The model of Duncan et al. (1985) couples a homogeneous, isotropic Voigt material of uniform thickness and infinite horizontal extent with a modified potential flow. The actual mean pressure distribution in a turbulent or laminar boundary layer flow is represented by modulating a potential flow pressure equation to allow for a reduced magnitude and phase change. Although this may be a reasonable assumption in view of the experimental data of Kendall (1970) and the theoretical work of Benjamin (1959), the results obtained numerically by Duncan et al. (1985) are only in qualitative agreement with the present experimental data and those obtained by Gad-el-Hak et al. (1984).

By selecting "appropriate" values for the model's parameters, Duncan et al. (1985) give detailed comparison between their calculations and existing experimental data. In the case of a viscoelastic coating with high damping, they find that the first instability occurring with increasing flow speed is a damping instability (Class A), which has phase speeds of a few percent of U_∞. When the damping is reduced sufficiently to approximate an elastic coating, the first instability found with increasing flow speed is a phase-lag instability (Class B), which has a much larger phase speed. Their model gives reasonable predictions in both magnitude and trend for the onset flow velocity and phase speed of the waves under turbulent conditions; however, predictions of wavelength were not as satisfactory. Duncan et al. attributed the differences between theory and experiment to the extremely complex viscoleastic properites of the compliant materials used in the experiments and to the difference between pressure fluctuations in an actual boundary layer flow and those represented by their modified potential flow. They concluded that more accurate predictions may be obtained by increasing the complexity of the flow and solid models without changing the basic physics of the coupling between the two.

5 Conclusions

The flow-induced motion of a compliant surface was measured using a remote optical technique. Two different classes of hydroelastic instabilities were observed on the elastic surface and on the viscoelastic surface. The onset speed for these instabilities depends upon the coating's geometrical and mechanical properties. The elastic surface instability has a relatively high phase speed and a small wavelength, and its wave profile is symmetric as compared to the slow and highly nonlinear "static-divergence" waves observed on the viscoelastic surface. Thus, it appears that the addition of viscous damping to a compliant material can produce a dramatic difference in the characteristics of the waves on the solid/fluid interface when subjected to the perturbation of a turbulent boundary layer. The experimentally determined wave characteristics compare qualitatively to existing theories on fluid/compliant surface interactions.

The experimental results presented in Section 3 agree qualitatively with the theory by Duncan et al. (1985). The onset speed and the slow phase speed observed for the waves on the viscoelastic surface are the same as that predicted for Class A instability (static-divergence waves). On the other hand, the theory predicts a much higher phase speed for Class B instability occuring on surfaces with little or no damping.

The phase speed of the elastic waves is an order of magnitude larger than that for the viscoelastic waves as seen in Fig. 8 of the present paper and Fig. 14 of Gad-el-Hak et al. (1984).

Acknowledgment

This work is supported by the U.S. Office of Naval Research, under Contract N00014-81-C-0453, and monitored by Dr. M. M. Reischman. The author wishes to acknowledge the many helpful discussions with Drs. R. F. Blackwelder, J. H. Duncan, H.-T. Liu, and J. J. Riley. A version of this paper was presented at the 1984 winter annual meeting of The American Society of Mechanical Engineers (Gad-el-Hak, 1984).

References

Ash, R. L., Bushnell, D. M., Weinstein, L. M., and Balasubramanian, R. (1975), "Compliant Wall Surface Motion and its Effects on the Structure of a Turbulent Boundary Layer," *Proc. Fourth Biennial Symp. on Turbulence*, eds. J. L. Zakin and G. K. Patterson, Univ. Missouri-Rolla, pp. 220-243.

Benjamin, T. B. (1959), "Shearing Flow Over a Wavy Boundary," *J. Fluid Mech.*, Vol. 6, pp. 161-205.

Benjamin, T. B. (1960), "Effects of a Flexible Boundary on Hydrodynamic Stability," *J. Fluid Mech.*, Vol. 9, pp. 513-532.

Benjamin, T. B. (1963), "The Threefold Classification of Unstable Disturbances in Flexible Surfaces Bounding Inviscid Flows," *J. Fluid Mech.*, Vol. 16, pp. 436-450.

Boggs, F. W., and Hahn, E. R. (1962), "Performance of Compliant Skins in Contact with High Velocity in Water," *Proc. Seventh Joint Army-Navy-Air Force Conference on Elastomer Research and Development*, Vol. 2, San Francisco, pp. 443-464.

Bushnell, D. M., Hefner, J. N., and Ash, R. L. (1977), "Effect of Compliant Wall Motion on Turbulent Boundary Layers," *Phys Fluids*, Vol. 20, pp. S31-48.

Dinkelacker, A., Hessel, M., Meier, G. E. A., and Schewe, G. (1977), "Investigation of Pressure Fluctuations Beneath a Turbulent Boundary Layer by Means of an Optical Method," *Phys. Fluids*, Vol. 20, pp. S216-224.

Duncan, J. H., and Hsu, C. C. (1984), "The Response of a Two-Layer Viscoelastic Coating to Pressure Disturbances from a Turbulent Boundary Layer," AIAA Paper No. 84-0535.

Duncan, J. H., Waxman, A. M., and Tulin, M. P. (1982), "Dispersion Relationships for Waves at the Interface Between a Single Layer Visco-Elastic Compliant Coating and a Turbulent Flow," Hydronautics Tech. Report No. 8111-1.

Duncan, J. H., Waxman, A. M., and Tulin, M. P. (1985), "The Dynamics of Waves at the Interface Between a Viscoelastic Coating and a Fluid Flow," *J. Fluid Mech.*, in press.

Gad-el-Hak, M. (1984), "An Optical Technique for Measuring the Flow-Induced Motion of a Compliant Surface," in *Symposium on Flow-Induced Vibrations*, Vol. 5, eds. M. P. Paidoussis and A. J. Kalinowski, ASME, pp. 9-22.

Gad-el-Hak, M., Blackwelder, R. F., and Riley, J. J. (1981), "On the Growth of Turbulent Regions in Laminar Boundary Layers," *J. Fluid Mech.*, Vol. 110, pp. 73-95.

Gad-el-Hak, M., Blackwelder, R. F., and Riley, J. J. (1982), "Interaction of Compliant Surfaces with Transitional and Turbulent Boundary Layers," in *Structure of Complex Turbulent Shear Flow*, eds. R. Dumas and L. Fulachier, Springer, pp. 20-30.

Gad-el-Hak, M., Blackwelder, R. F., and Riley, J. J. (1984), "On the Interaction of Compliant Coatings with Boundary Layer Flows," *J. Fluid Mech.*, Vol. 140, pp. 257-280.

Grosskreutz, R. (1971), "Wechselwirkungen Zwischen Turbulenten Grenzschichten und Weichen Wanden," Max-Planck-Institut fur Stromungsforschung, Gottingen, No. 53.

Hansen, R. J., and Hunston, D. L. (1974a), "An Experimental Study of the Hydrodynamic Drag on Compliant Surfaces: Fluid Property Effects," *Proc. Eighth Int. Congress on Acoustics*, Vol. 2, pp. 579-579.

Hansen, R. J., and Hunston, D. L. (1974b), "An Experimental Study of Turbulent Flows over Compliant Surfaces," *J. Sound & Vibration*, Vol. 34, pp. 297-308.

Hansen, R. J., and Hunston, D. L. (1976), "Further Observations on Flow-Generated Surface Waves in Compliant Surfaces," *J. Sound & Vibration*, Vol. 46, pp. 593-596.

Hansen, R. J., and Hunston, D. L. (1983), "Fluid Property Effects on Flow-Generated Waves on a Compliant Surface," *J. Fluid Mech.*, Vol. 133, pp. 161-177.

Hansen, R. J., Hunston, D. L., Ni, C. C., and Reischman, M. M. (1980a), "An Experimental Study of Flow-Generated Waves on a Flexible Surface," *J. Sound & Vibration*, Vol. 68, pp. 317-334.

Hansen, R. J., Hunston, D. L., Ni, C. C., Reischman, M. M., and Hoyt, J. W. (1980b), "Hydrodynamic Drag and Surface Deformations Generated by Liquid Flows over Flexible Surfaces," in *Viscous Flow Drag Reduction*, ed. G. R. Hough, AIAA, pp. 439-452.

Hoyt, J. W. (1981), "A Flow-Visualization Study of Turbulent Spots on Solid and Compliant Surfaces," *Proc. Seventh Biennial Symp. on Turbulence*, Eds. J. L. Zakin and G. K. Patterson, Univ. Missouri-Rolla, Paper No. 32.

Jaeger, J. C., and Cook, N. G. W. (1976), *Fundamentals of Rock Mechanics*, second edition, Chapman & Hall.

Kaplan, R. E. (1964), "The Stability of Laminar Incompressible Boundary Layers in the Presence of Compliant Boundaries," Sc.D. thesis, MIT.

Kendall, J. M. (1970), "The Turbulent Boundary Layer Over a Wall with Progressive Waves," *J. Fluid Mech.*, Vol. 41, pp. 259-281.

Landahl, M. T. (1962), "On the Stability of a Laminar Incompressible Boundary Layer over a Flexible Surface," *J. Fluid Mech.*, Vol. 13, pp. 609-632.

Liu, H.-T., Katsaros, K. B., and Weissman, M. A. (1982), "Dynamic Response of Thin-Wire Wave Gauges," *J. Geophys. Res.*, Vol. 87, pp. 5686-5698.

Liu, H.-T., and Lin, J.-T. (1982), "On the Spectra of High-Frequency Wind Waves," *J. Fluid Mech.*, Vol. 123, pp. 165-185.

Rathsam, A. D., Mastny, G. F., Phillips, T. E., and Reischman, M. M. (1983), "Laser/Optics System for Measuring Microscopic Motion on Compliant Materials in Turbulent Boundary Layers," *Proc. Eighth Biennial Symp. on Turbulence*, eds. J. L. Zakin and G. K. Patterson, Univ. Missouri-Rolla, Paper No. 30.

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

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

Journal of Applied Mechanics

Design Data and Methods

Stresses and Displacements on the Boundaries of Circular Rings Diametrically Loaded

A. J. Durelli[1,2] **and Y. H. Lin**[2]

The paper deals with stresses and displacements in circular rings of rectangular cross-section, loaded in the plane and perpendicular to the boundary. Values are given for all points at the inside and outside boundaries, are presented parametrically for rings for which the ratio of diameters ID/OD varies from 0 to very close to 1, and have been obtained from several sources, mainly Nelson's equations. References to some previous contributions are included. The information presented in the paper was not available in a complete manner and will be useful in numerous structural applications. The analysis corresponding to loads applied tangentially to the boundary could be approached in a similar manner.

Introduction

The circular ring of rectangular cross-section with a concentric hole is probably the most commonly treated geometry in the stress analysis literature, and the most commonly considered loading condition of that geometry is the pair of loads diametrically applied to the outside boundary. Applications are numerous because this type of ring is the transverse cross-section of tubes widely used in many kinds of construction. Many contibutions to the solution of this problem can be found in the literature. Several of these are referred to in treatises like the ones by Timoshenko [1], Roark [2], and Peterson [3], and they include some basic theoretical work like Filon's [4] as well as experimental results [5]. Among other important contributions the work of Billevicz [6], Horger [7], Bell [8], and Ripperger and David [9] may be mentioned. Probably the most thorough treatment of the subject has been conducted by Nelson [10] in his thesis. Some of the other contributions, by one of the authors, to this subject can be found in [11].

For many applications, in particular the use of rings as dynamometers, the necessary information from the many papers mentioned is limited to the maximum and minimum stress on the inside boundary of the ring when the ring is loaded by a pair of opposite forces. There are other applications, however, like the use of rings as dynamometers for the simultaneous measurements of several pairs of opposite loads, which require the knowledge of stresses at all points of the inside boundary, and the possession of this knowledge parametrically for many ratios *ID/OD* of the inside to the outside boundary.

The authors are proposing in another publication the use of coefficients of influence to evaluate the multiple loads applied to the rings. The object of this note is to present parametrically in a manner that permits a direct use of the data the values of stresses and displacements on the inside and outside boundaries for ratios *ID/OD* from 0 to 0.92.

The results presented are drawn from both theoretical and experimental contributions. Advantage has been gained by the cross plotting of curves to increase precision. Advantage has also been gained by the use of the two limiting cases, the small circular hole in an infinite plate (Kirsch solution) and the curved beam formulae for the very thin ring. Both give highly accurate results. It is believed that this is the first time parametrical information has been presented covering practically the whole field. The scope of the paper is limited however to the elastic behavior of materials, and small deformations of the rings. Considerable information for the case of finite deformations can be found in another publication [12].

Stresses on the Inside and Outside Boundaries

Only the stresses at four points of the boundaries for some typical $\alpha = ID/OD$ were calculated by Nelson [10]. The complete analysis was conducted here using his equations:

$$(\sigma_\theta)_i = -M_o \frac{P}{\pi R_o t} + \frac{P}{\pi R_o t}\,(-M_2 \cos 2\theta^* + M_4 \cos 4\theta^* - M_0 \cos 6\theta^* + \ldots)$$

and

$$(\sigma_\theta)_o = \frac{M'P}{\pi R_o t} - M_o{}' \frac{P}{\pi R_o t} + \frac{P}{\pi R_o t}\,(M_2{}' \cos 2\theta^* - M_4{}' \cos\theta^* + M_6{}' \cos 6\theta^* \ldots)$$

where $M' = 1$ when $\theta^* = \pi/2$ or $\dfrac{3\pi}{2}$

$$M' = 0 \text{ when } \theta^* \neq \frac{\pi}{2} \text{ or } \frac{3\pi}{2}$$

---

[1] Fellow ASME.

[2] University of Maryland, College Park, Md. 20742.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July, 1984; final revision, April, 1985.

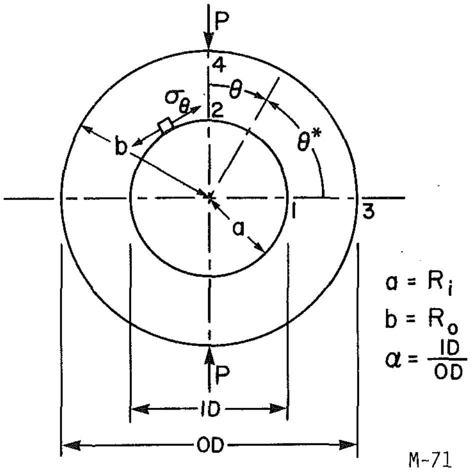


Fig. 1 Geometry and loading of the rings

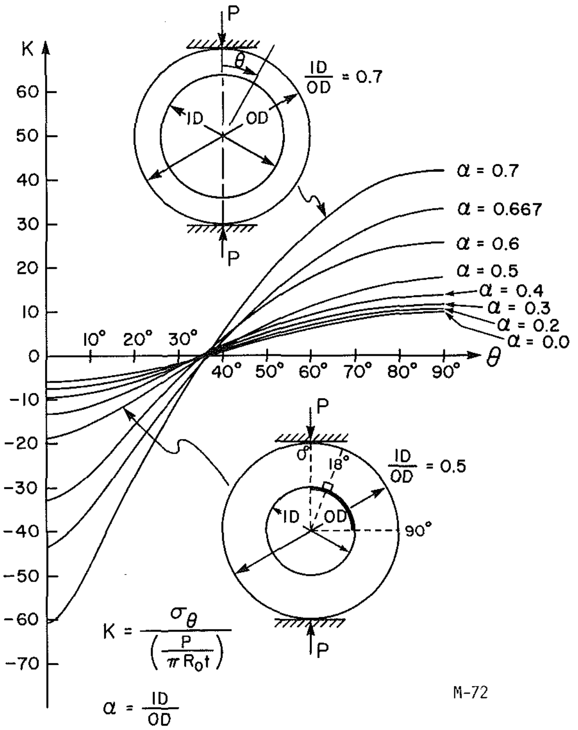


Fig. 2 Stress factors at the inner boundary of circular rings diametrically loaded (0 < α < 0.7; α as parameter)

$$M_o = \frac{2}{(1-\alpha^2)}\,;\quad M_o{}' = \frac{2\alpha^2}{(1-\alpha^2)}$$

$$M_n = \frac{4n\alpha^{n-2}(1-\alpha^2)(1-\alpha^{2n})}{Q_n} \qquad \text{for } n=2,4,6,\ \ldots$$

$$M_n{}' = \frac{4(1-\alpha^2)^2}{\alpha^2}\,\frac{n^2\alpha^{2n}}{Q_n} \qquad \text{for } n=2,4,6,\ \ldots$$

and

$$Q_n = (1-\alpha^{2n})^2 - n^2\alpha^{2n-2}(1-\alpha^2)^2$$

The meaning of other symbols is explained in Fig. 1.

There is good agreement between the more extensive computations conducted here (Figs. 2–11), and those shown by Nelson. It should be pointed out, however, that as a consequence of the series nature of the solution, a jump was found for the values of the stresses in the neighborhood of the point of load application for the low values of α. For the computations reported here, sixty terms were used in the series. A larger number of terms may decrease the amount of the jump (see for instance Fig. 9). It was also observed that Nelson's values for the inside boundary at the vertical axis are slightly higher than some of the ones reported here.

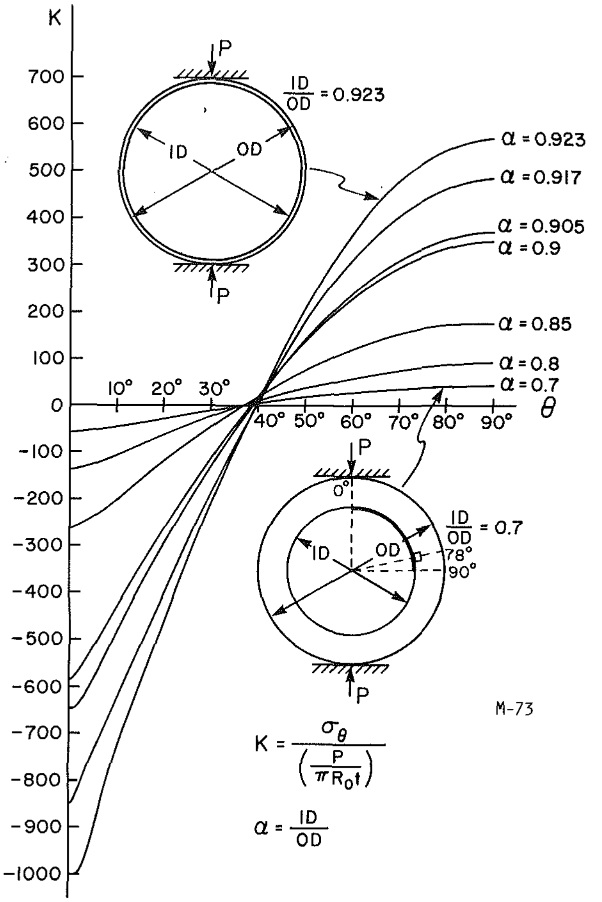


Fig. 3 Stress factors at the inner boundary of circular rings diametrically loaded (0.7 < α < 0.91; α as parameter)

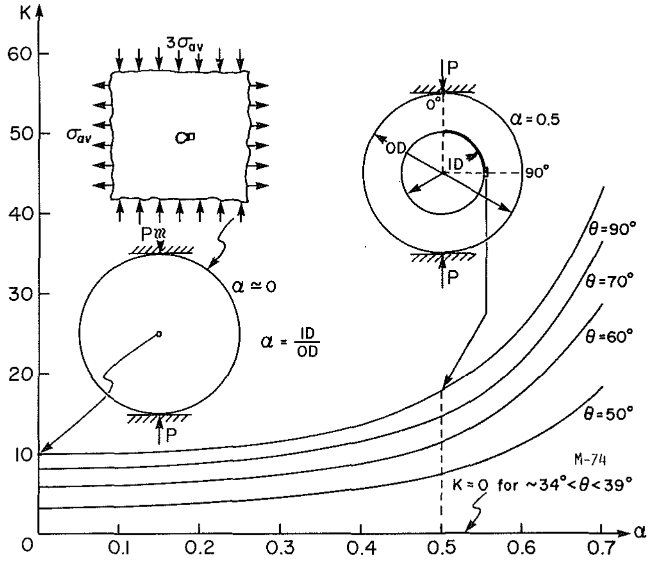


Fig. 4 Stress factors along the inner boundary of circular rings diametrically loaded (0 < α < 0.7; ~ 40 deg < θ < 90 deg; θ as parameter)

Displacements on the Inside and Outside Boundaries

Nelson also gives the displacements for the four

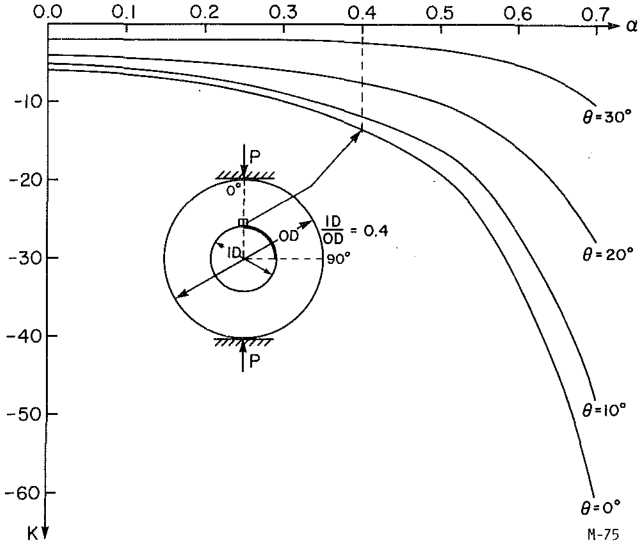


Fig. 5 Stress factors along the inner boundary of circular rings diametrically loaded (0 < α < 0.7; 0 deg < θ < ~ 40 deg; θ as parameter)

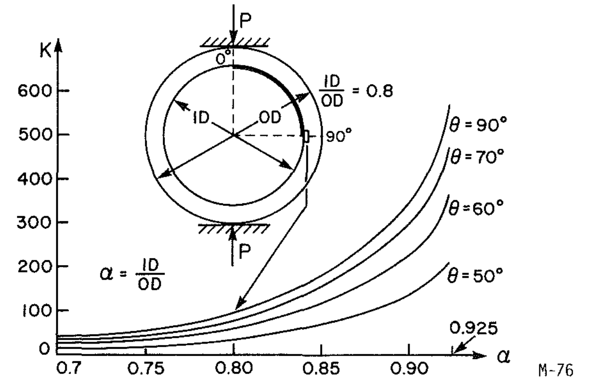


Fig. 6 Stress factors along the inner boundary of circular rings diametrically loaded (~ 40 < θ < 90 deg; 0.7; < α < 0.923; θ as parameter)

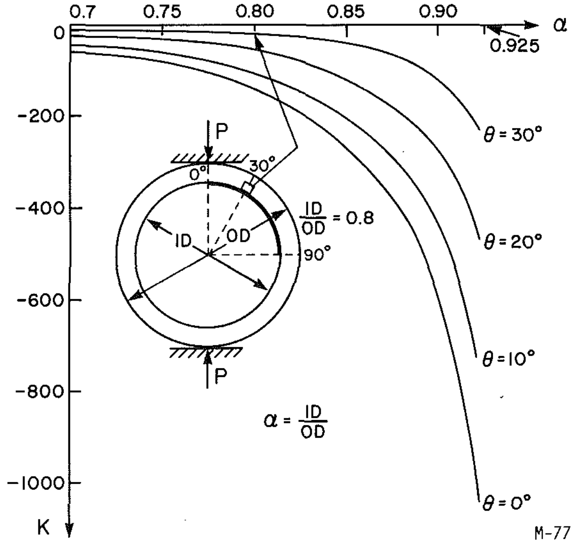


Fig. 7 Stress factors along the inner boundary of circular rings diametrically loaded (0 deg < θ < ~ 40 deg; 0.7; < α < 0.923; θ as parameter)

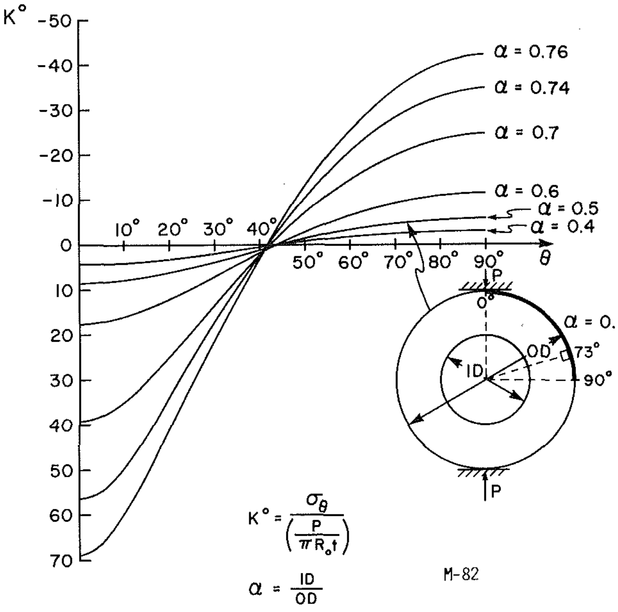


Fig. 8 Stress factors at the outer boundary of circular rings diametrically loaded (0 < α < 0.76 as parameter)

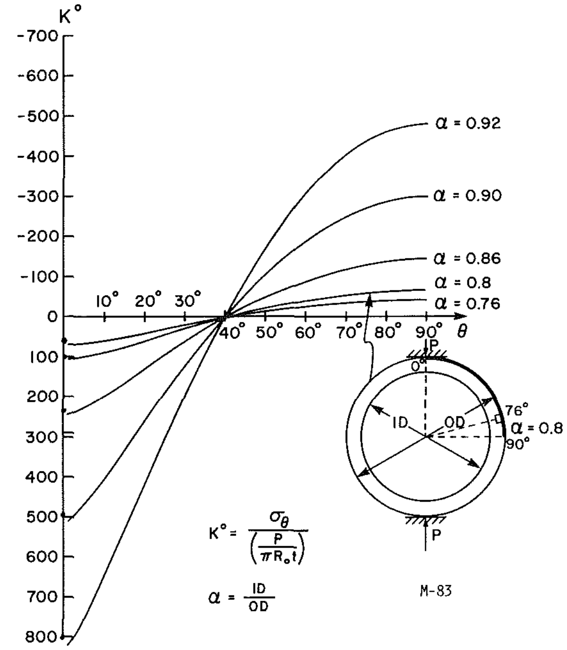


Fig. 9 Stress factors at the outer boundary of circular rings diametrically loaded (0.76 < α < 0.92 as parameter)

aforementioned points (Fig. 1). The values for all points on the boundries, given in Figs. 12–21, were calculated using the following equations for $(u_r)_i$ and $(u_r)_o$[10].

$$(u_e)_o = (u)_{R_i=0,R_o} - N_o'\frac{P'}{\pi E} + \frac{P'}{\pi E}(N_2'\cos 2\theta^* - N_4'\cos 4\theta^* + N_o'\cos 6\theta^* \ldots)$$

where

$$N_o = 2\alpha/1-\alpha^2$$

$$N_n = \frac{4n}{n^2-1}\alpha^{n-1}[n(1-\alpha^2)(1+\alpha^{2n}) + (1+\alpha^2)(1-\alpha^{2n})]/Q_n \quad \text{for } n=2,4,6\ldots$$

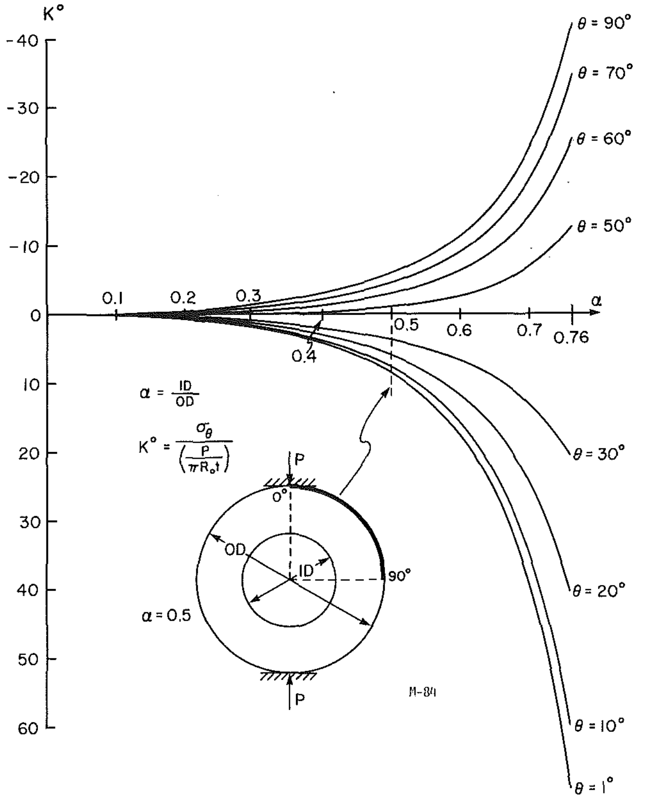


Fig. 10 Stress factors at the outer boundary of circular rings diametrically loaded (0.0 < α < 0.76; θ as parameter)

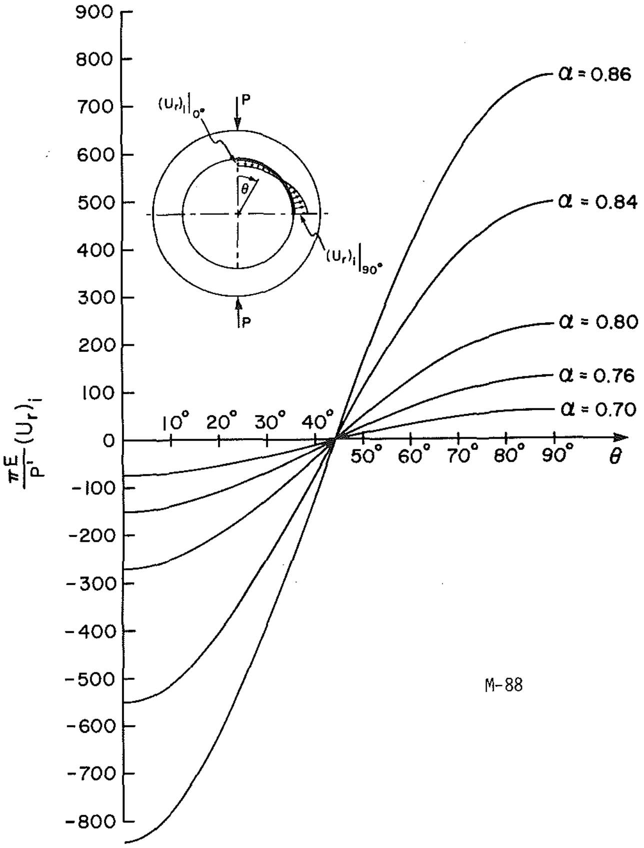


Fig. 12 Radial displacements at the inner boundary of circular rings diametrically loaded (0.0 < α < 0.7 as parameter)

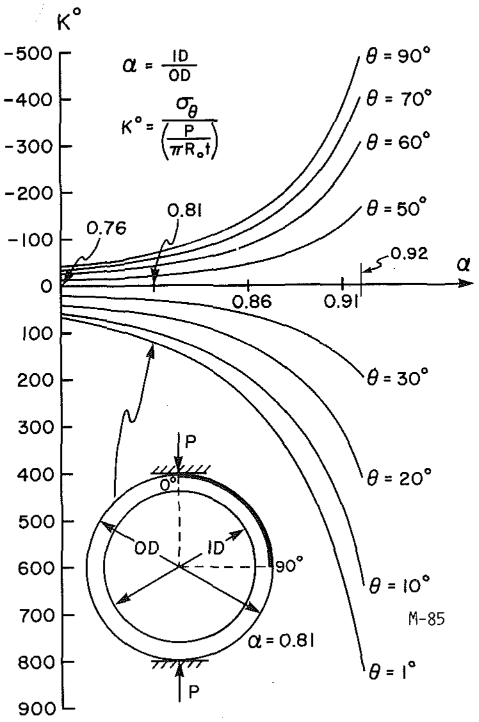


Fig. 11 Stress factors at the outer boundary of circular rings diametrically loaded (0.76 < α < 0.92; θ as parameter)

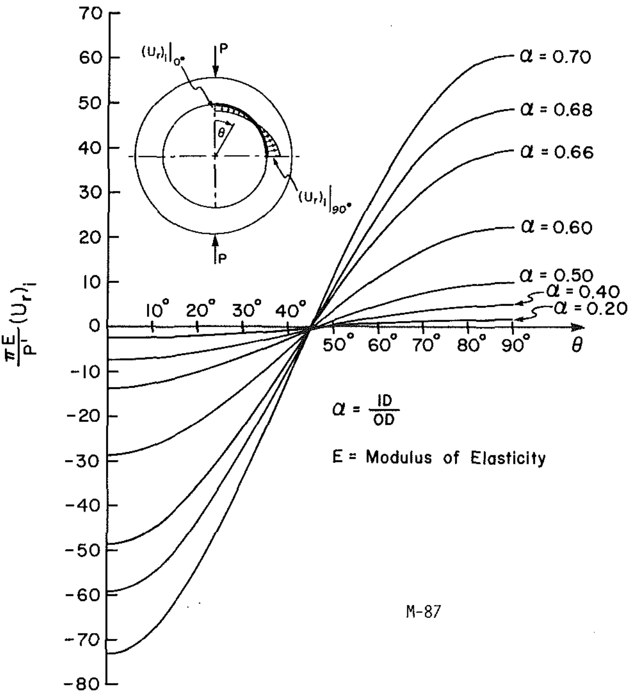


Fig. 13 Radial displacements at the inner boundary of circular rings diametrically loaded (0.7 < α < 0.86 as parameter)

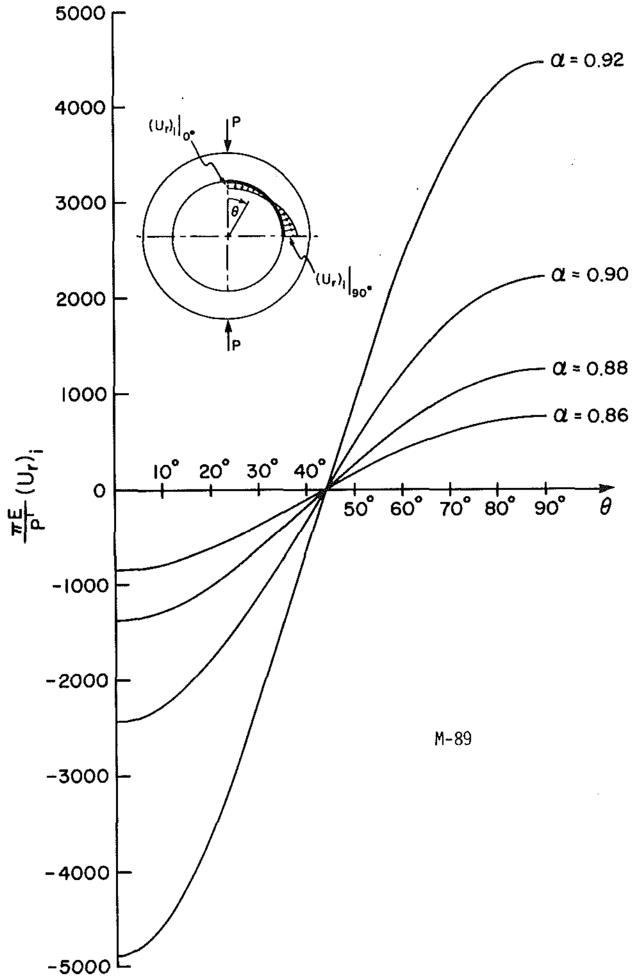


Fig. 14 Radial displacements at the inner boundary of circular rings diametrically loaded (0.86 < α < 0.92 as parameter)

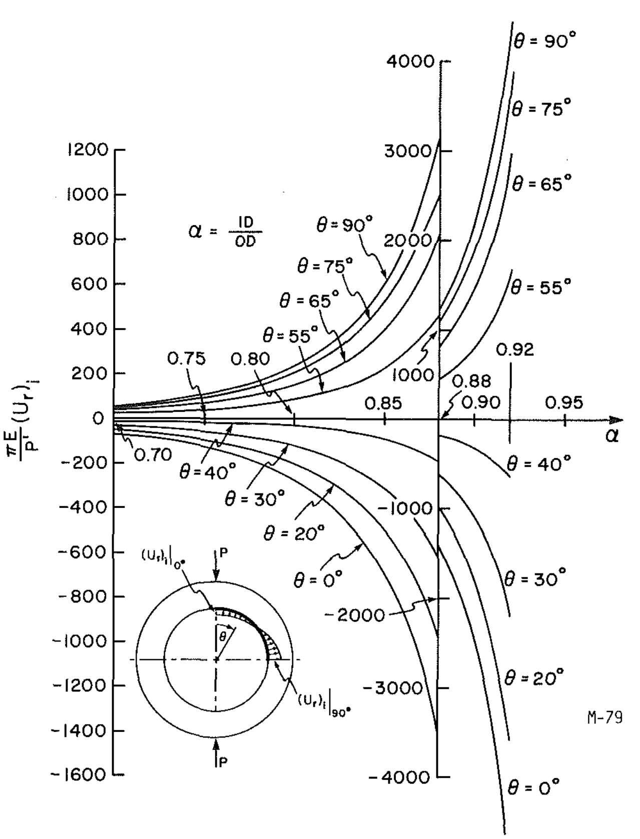


Fig. 16 Radial displacements at the inner boundary of circular rings diametrically loaded (0.7 < α < 0.92; θ as parameter)

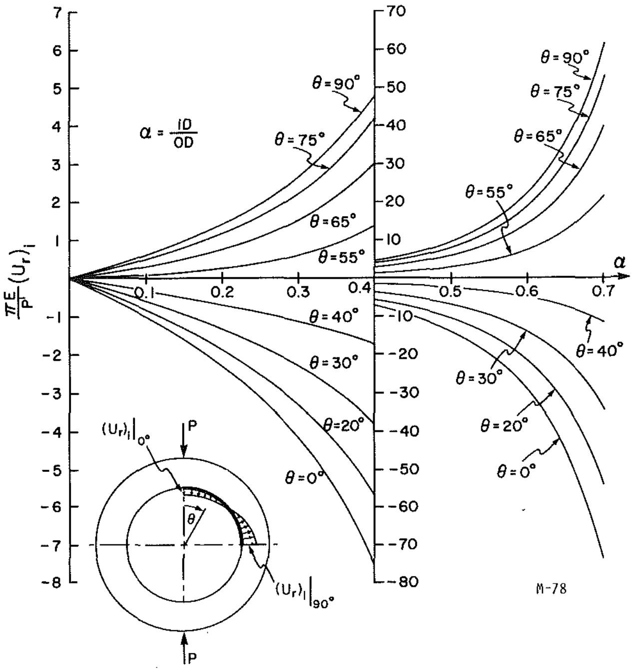


Fig. 15 Radial displacements at the inner boundary of circular rings diametrically loaded (0 < α < 0.7; θ as parameter)

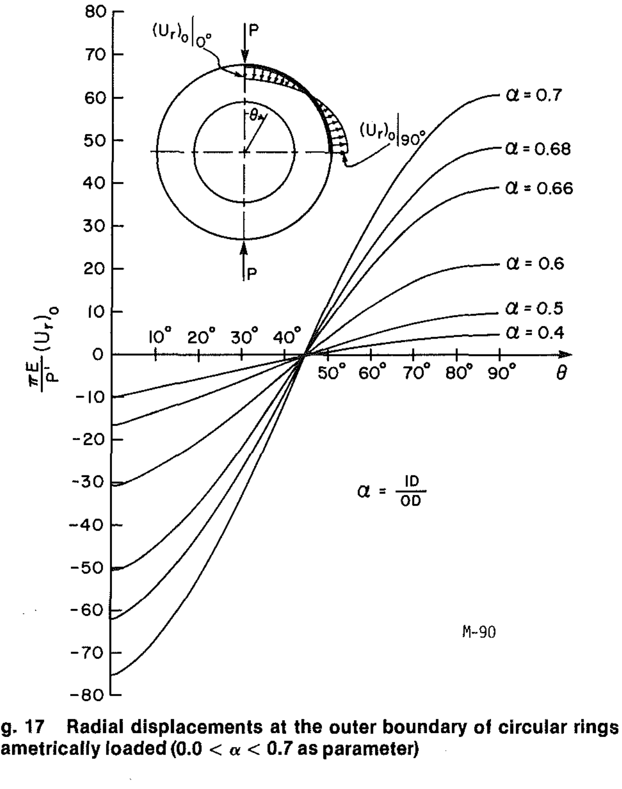


Fig. 17 Radial displacements at the outer boundary of circular rings diametrically loaded (0.0 < α < 0.7 as parameter)

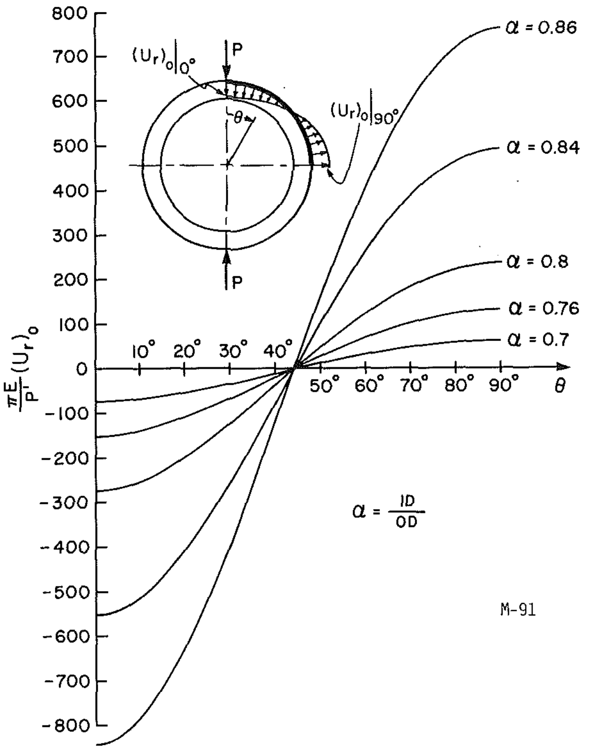


Fig. 18 Radial displacements at the outer boundary of circular rings diametrically loaded (0.7 < α < 0.86 as parameter)

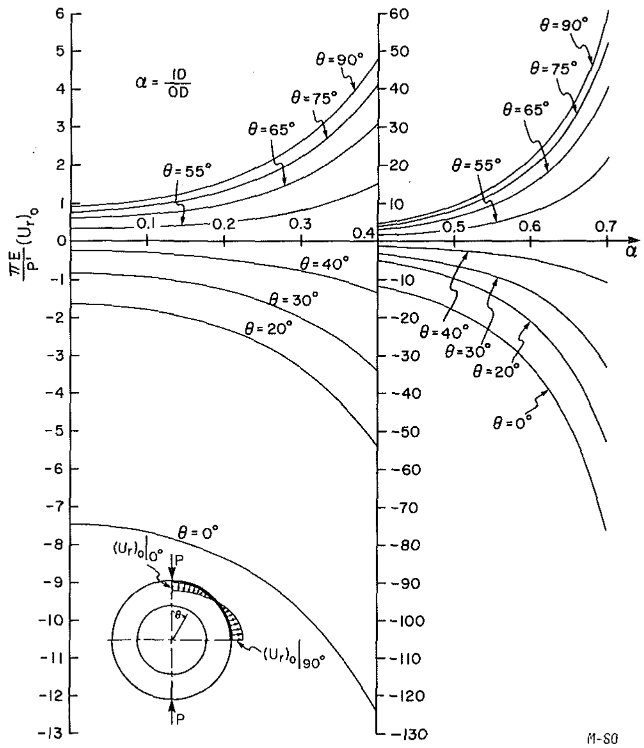


Fig. 20 Radial displacements at the outer boundary of circular rings diametrically loaded (0.0 < α < 0.7 θ as parameter)

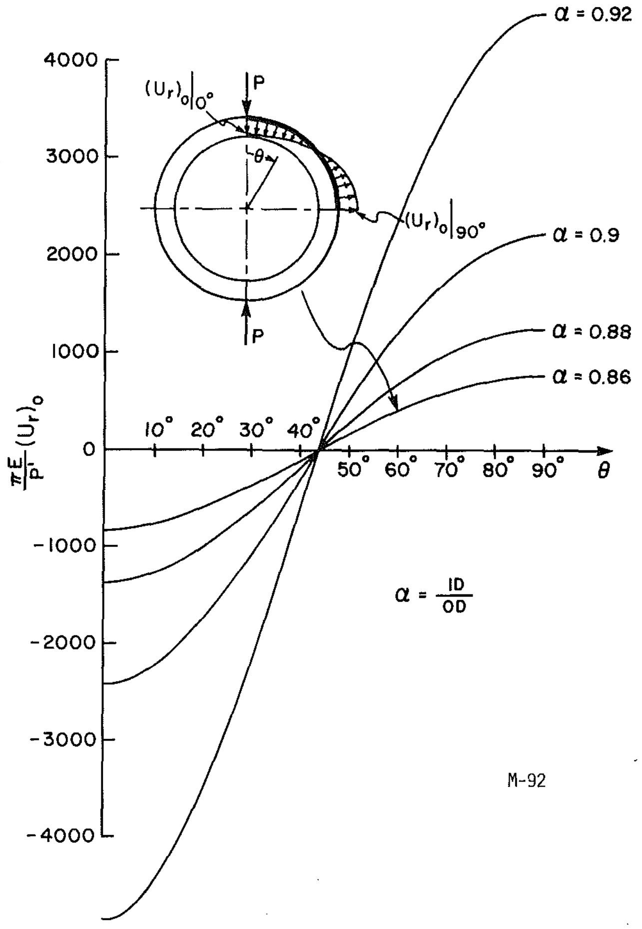


Fig. 19 Radial displacements at the outer boundary of circular rings diametrically loaded (0.86 < α < 0.92 as parameter)

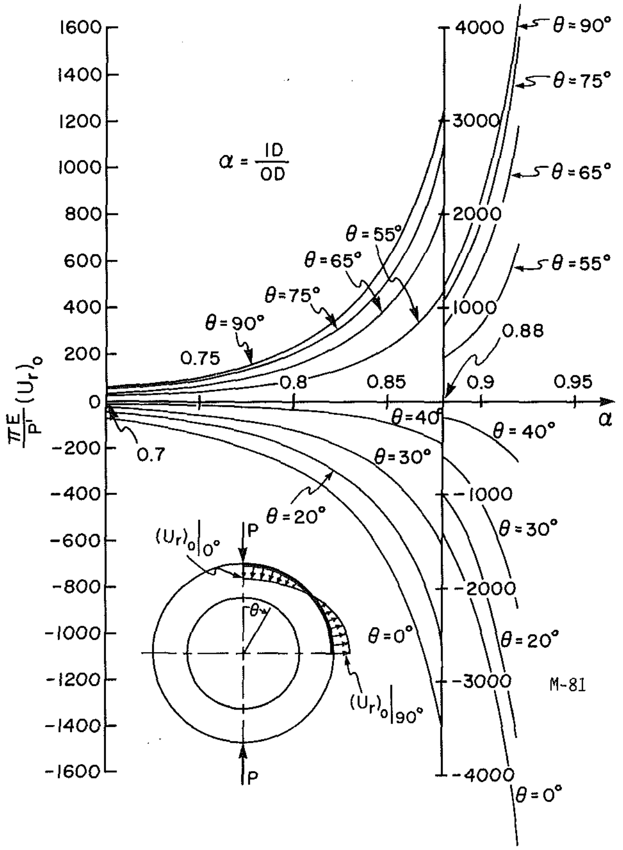


Fig. 21 Radial displacements at the outer boundary of circular rings diametrically loaded (0.7 < α < 0.92; θ as parameter)

$$N_o' = 2\alpha^2/1-\alpha^2$$

$$N_n' = \frac{4n}{n^2-1}\left\{\frac{[1-\alpha^{4n}]+n\alpha^{2n-2}(1-\alpha^4)]}{Q_n}-1\right\}$$

$$\text{for } n=2,4,6\ .\ .$$

$$Q_n = (1-\alpha^{2n})^2 - n^2\alpha^{2n-2}(1-\alpha^2)^2$$

$$(u)_{R_i=0,R_o} = \frac{2P'}{\pi E} + \frac{P'}{\pi E}\sin\theta^*\log\frac{(1-\sin\theta^*)}{(1+\sin\theta^*)} - \frac{(1-\nu)P'}{2E}\cos\theta^*$$

$$\text{for} \qquad -\pi/2 \leq \theta^* \leq \pi/2$$

$$\theta^* = \pi/2-\theta$$

Acknowledgment

Professor Durelli is grateful to the chairman of the Mechanical Engineering Department, Prof. W. Fourney, for the opportunity of conducting this study, part of a series of studies on the analysis of stress and displacement. Credit should be given to Steve Thorssell for the drawings and to Thelma Miller for reproducing the manuscript. Y. H. Lin acknowledges the financial support of the Minta Martin Fund at the Unversity of Maryland.

References

1 Timoshenko, S., "On the Distribution of Stress in a Circular Ring Compressed by Two Forces Along a Diameter," *Philosophical Magazine*, Vol. 44, 1922, p. 1014.

2 Roark, R., *Formulas for Stress and Strain* (3rd ed., McGraw Hill, 1954, p. 156.

3 Peterson, R. E., *Stress Concentration Factors*, Wiley, 1974, p. 263.

4 Coker, E. G., and Filon, L. N. G., *A Treatise on Photo-Elasticity*, Cambridge University, Press, 1931, p. 373.

5 Frocht, M., *Photoelasticity*, Vol. I, John Wiley & Sons, New York, 1941.

6 Billevicz, V., "Analysis of Stress in Circular Rings," Ph.D. thesis, University of Michigan, 1931.

7 Horger, O. J., and Buckwalter, T. V., "Photoelasticity as Applied to Design Problems," *Jrn. Age*, Vol. 145, Pt. II. No. 21, 1940, p. 42.

8 Bell, G., "Theorie and Versuche ueber einige Faelle von Spannungsverteilung in ringfoermigen Koerperen," *Zeitshcrift fuer angew. Math. and Mech.*, Band 10, Heft 1, Feb. 1930, pp. 52–73.

9 Ripperger, E. A., and David N., "Critical Stresses in a Circular Ring," *Proc. ASCE*, Feb. 1946.

10 Nelson, C. W., "Stresses and Displacements in a Hollow Circular Cylinder," Ph.D. thesis, University of Michigan, 1939.

11 Durelli, A. J., and Parks, V. J., *Moire Analysis of Strain*, Prentice-Hall, New York, 1970.

12 Durelli, A. J., "Stress Concentrations," in *Experimental Evaluation of Stress Concentration and Intensity Factors*, G. C. Sih, ed., Martenus Nijhoff, 1981, p. 50–55.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N. Y. 10017, or to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

On the Singularities in Reissner's Theory for the Bending of Elastic Plates

W. S. Burton[1] **and G. B. Sinclair**[1]

Wedge-shaped elastic plates under bending, with the edges forming the wedge vertex being either stress-free, clamped or simply supported, are characterized as to possible singular behavior within the context of Reissner's plate theory.

Introduction

Probably the first singularity analysis of an angular elastic plate under bending is William's treatment using the *classical theory* [1]. In the classical theory it is possible to satisfy stress-free conditions at an edge solely in an approximate way, since only two boundary conditions can be enforced and there are three stress resultants. As the boundary conditions play an important role in governing singular behavior at the vertex of any corner in a plate it is to be expected that *Reissner's theory* [2], which admits three, physically-natural, boundary conditions on an edge, may offer an improved, albeit singular, representation in these instances. This is the expectation that possibly motivated other analysts (e.g., Wang [3]) to perform analyses of complete, individual, crack problems using Reissner's theory rather than the classical, and indeed more physically sensible results are derived in these analyses. Specifically, for the crack-tip on the tensile face of the plate the same hydrostatic singular field ahead of the crack as occurs in the extensional case of a cracked plate under tension is found; this is in contrast to the classical bending theory in which the principal stresses ahead of the crack differ in sign and magnitude. As a result it would seem reasonable to attempt the analogue of Williams' study for the classical theory [1] and explore the singularities in Reissner's plate theory for a wider range of geometries than that investigated elsewhere and for a full range of boundary conditions; this is the intent of the present note.

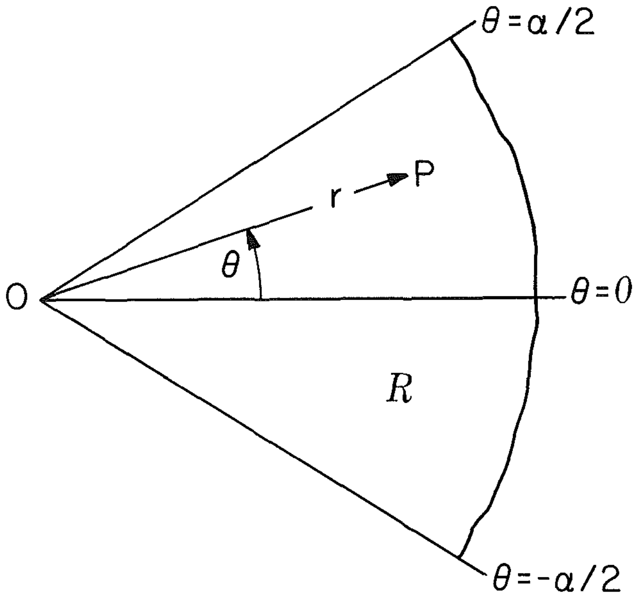


Fig. 1 Geometry and coordinates for the plate

We begin by formulating a class of problems for a wedge-shaped elastic plate under bending within the context of Reissner's theory. Next we establish suitable solution forms for the dominant asymptotic response near the wedge vertex and set down conditions for the existence of these fields. The conditions basically involve the analysis of an eigenequation for each pair of edge conditions considered. The note concludes by displaying these eigenequations and discussing the eigenvalues satisfying them which give rise to singularities.

Formulation

The plate has thickness h and occupies the open wedge region, R,

$$R=\{(r,\theta)\,|\,0<r<\infty,-\alpha/2<\theta<\alpha/2\}(0<\alpha\leq 2\pi), \qquad (1)$$

where (r,θ) are the polar coordinates of a point P in the wedge with respect to the origin, 0, at the wedge vertex, and α is the vertex angle (Fig. 1).

The plate is comprised of a homogeneous, isotropic, and

---

[1] Department of Mechanical Engineering, Carnegie-Mellon University, Pittsburgh, PA 15213.

Manuscript received by ASME Applied Mechanics Division, December 19, 1983; final revision July 15, 1985.

Table 1 Eigenequations for Reissner's plate theory generating singular moment resultants

Edge Conditions	Eigenequation	Constants
Stress-free/stress-free	$\sin\lambda\alpha = C_1\lambda$	$C_1 = (-)^k \sin\alpha$
Clamped/clamped	$\sin\lambda\alpha = C_2\lambda$	$C_2 = (-1/\kappa)^{k+1}\sin\alpha$
Clamped/stress-free	$\sin^2\lambda\alpha = C_3 - C_4\lambda^2$	$C_3 = 4/\kappa(1+\nu)^2$ $C_4 = (1/\kappa)\sin^2\alpha$
Stress-free/simply supported	$\sin 2\lambda\alpha = C_5\lambda$	$C_5 = \sin 2\alpha$
Clamped/simply supported	$\sin 2\lambda\alpha = C_6\lambda$	$C_6 = (-1/\kappa)\sin 2\alpha$
Simply supported/simply supported	$\cos\lambda\alpha = C_7$	$C_7 = (-)^k\cos\alpha$

Note: $k = 1,2$ for symmetric or anti-symmetric loading, respectively, and $\kappa = (3-\nu)/(1+\nu)$.

linear elastic material having Young's modulus, E, and Poisson's ratio, ν. For this plate the stress resultants and rotations of Reissner's plate theory, in the absence of surface loading, can be expressed in terms of the out of plane deflection, $w(r,\theta)$, and a single stress potential, $\chi(r,\theta)$. That is,

$$V_r = \frac{1}{r}\frac{\partial\chi}{\partial\theta}, \quad V_\theta = -\frac{\partial\chi}{\partial r},$$

$$M_\theta = 2\gamma\left(\frac{1}{r^2}\frac{\partial\chi}{\partial\theta} - \frac{1}{r}\frac{\partial^2\chi}{\partial r\partial\theta}\right)$$

$$-D\left(\nu\frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{1}{r^2}\frac{\partial^2 w}{\partial\theta^2}\right),$$

$$M_r = 2\gamma\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\chi}{\partial\theta}\right) - D\left(\frac{\partial^2 w}{\partial r^2} + \frac{\nu}{r}\frac{\partial w}{\partial r} + \frac{\nu}{r^2}\frac{\partial^2 w}{\partial\theta^2}\right),$$

$$M_{r\theta} = \gamma\left(\frac{1}{r^2}\frac{\partial^2\chi}{\partial\theta^2} - r\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\chi}{\partial r}\right)\right) \tag{2}$$

$$-(1-\nu)D\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial w}{\partial\theta}\right),$$

$$\beta_\theta = -\frac{2\gamma}{D(1-\nu)}\frac{\partial\chi}{\partial r} - \frac{1}{r}\frac{\partial w}{\partial\theta},$$

$$\beta_r = \frac{2\gamma}{D(1-\nu)}\frac{1}{r}\frac{\partial\chi}{\partial\theta} - \frac{\partial w}{\partial r},$$

on R, where V_r, V_θ, the shear resultants, M_θ, M_r, $M_{r\theta}$, the moment resultants and β_θ, β_r, the rotations, are functions of r, θ, defined in the usual manner, with $\gamma = h^2/10$, $D = Eh^3/12(1-\nu^2)$, the last being the flexural rigidity. Then the *field equations* of Reissner's theory are reduced to the Cauchy-Riemann equations for the functions $\chi - \gamma\nabla^2\chi$ and $D\nabla^2 w$, i.e.,

$$\frac{\partial}{\partial r}(\chi - \gamma\nabla^2\chi) = \frac{1}{r}\frac{\partial}{\partial\theta}(D\nabla^2 w),$$

$$\frac{1}{r}\frac{\partial}{\partial\theta}(\chi - \gamma\nabla^2\chi) = \frac{-\partial}{\partial r}(D\nabla^2 w), \tag{3}$$

on R. Here ∇^2 is the Laplacian operator in cylindrical polar coordinates.

On each wedge face three homogeneous boundary conditions are to be satisfied. These three boundary conditions are combined in sets of *edge conditions* to model various edges as follows:

$$\begin{array}{ll} \text{Stress-free} & M_\theta = M_{r\theta} = 0,\ V_\theta = 0, \\ \text{Clamped} & \beta_\theta = \beta_r = 0,\ w = 0, \\ \text{Simply supported} & M_\theta = 0,\ \beta_r = 0,\ w = 0, \end{array} \tag{4}$$

on $\theta = \pm\alpha/2$ $(0 \le r < \infty)$. These three cases combine to give six distinct problems for the wedge. When the same conditions apply on each face, it is possible to distinguish between symmetric and anti-symmetric contributions. Thus, in effect, nine problems are considered. Ensuring bounded displacements concludes our formulation and limits the singular behavior admitted-while this formulation is then still not complete, it suffices given our objective of characterizing *possible* singular fields at the wedge vertex. We next consider the construction of suitable sets of asymptotic solutions.

Singularity Analysis

From the form of the relations in (2), observe that if $w = 0(r^{\lambda+1})$, $\chi = 0(r^{\lambda+1})$, as $r \rightarrow 0$ on R, λ a constant, then the shear resultants and rotations are $0(r^\lambda)$, while the moment resultants are $0(r^{\lambda-1})$. Accordingly we seek to construct separable solutions for w, χ satisfying the governing equations (3) which furnish six independent constants multiplying these dominant contributions as $r \rightarrow 0$, thereby providing a means of satisfying the six boundary conditions contained in any pair of the edge conditions (4). To this end, and noting that $\chi - \lambda\nabla^2\chi$ and $D\nabla^2 w$ are harmonic functions and the interrelations in (3), we therefore take as our asymptotic solution forms for χ and w, the biharmonic functions[2]

$$\chi = r^{\lambda+1}F(\lambda,\theta) + 0(r^{\lambda+3}),$$
$$w = r^{\lambda+1}G(\lambda,\theta) + 0(r^{\lambda+3}), \tag{5}$$

as $r \rightarrow 0$ on R, where

$$F(\lambda,\theta) = (b_1\cos(\lambda+1)\theta + b_2\sin(\lambda+1)\theta$$
$$+ b_3\cos(\lambda-1)\theta + b_4\sin(\lambda-1)\theta),$$
$$G(\lambda,\theta) = (b_5\sin(\lambda+1)\theta + b_6\cos(\lambda+1)\theta$$
$$-\gamma b_3\sin(\lambda-1)\theta + \gamma b_4\cos(\lambda-1)\theta)/D.$$

The stress and moment resultants and the rotations in (2) may then be written as

$$V_r = r^\lambda F' + 0(r^{\lambda+2}), \quad V_\theta = -(\lambda+1)r^\lambda F + 0(r^{\lambda+2}),$$
$$M_\theta = r^{\lambda-1}[-2\gamma\lambda F' - D((\lambda+1)(1+\nu\lambda)G + G'')] + 0(r^{\lambda+1}),$$
$$M_r = r^{\lambda-1}[2\gamma\lambda F' - D((\lambda+1)(\lambda+\nu)G + \nu G'')] + 0(r^{\lambda+1}),$$
$$M_{r\theta} = r^{\lambda-1}[\gamma(F'' - (\lambda+1)(\lambda-1)F) - D(1-\nu)\lambda G'] + 0(r^{\lambda+1}),$$
$$\beta_\theta = r^\lambda\left[\frac{-2\gamma}{D(1-\nu)}(\lambda+1)F - G'\right] + 0(r^{\lambda+2}), \tag{6}$$
$$\beta_r = r^\lambda\left[\frac{2\gamma}{D(1-\nu)}F' - (\lambda+1)G\right] + 0(r^{\lambda+2}),$$

as $r \rightarrow 0$ on R, where the primes denote differentiation with respect to θ.

With this set of separable functions for w and χ the singularity analysis proceeds in a manner similar to that developed by Williams [4] for power singularities and by Dempsey and Sinclair [5] for logarithmic singularities. Imposing the displacement regularity requirements on (6) and con-

---

[2] Note that one cannot use Reissner's solution [2] and have a sufficient number of independent constants available for the asymptotic analysis.

fining attention to singular solutions, the resulting conditions may be summarized by

$$\mathbf{M} = 0(r^{\lambda-1}) \text{ for real } \lambda \text{ satisfying } D=0,\ 0<\lambda<1,$$

$$\mathbf{M} = 0\left(r^{\lambda-1} \begin{Bmatrix} \sin(\eta \ln r) \\ \cos(\eta \ln r) \end{Bmatrix}\right) \text{ for complex } \lambda = \xi + i\eta$$

$$\text{satisfying } D=0,\ 0<\Delta Re\lambda<1, \tag{7}$$

$$\mathbf{M} = 0(r^{\lambda-1} \ln r) \text{ for real } \lambda \text{ satisfying}$$

$$\frac{d^{n-m}D}{d\lambda^{n-m}} = 0,\ m<n,\ 0<\lambda\leq 1,$$

$$\mathbf{M} = 0\left(r^{\lambda-1} \ln r \begin{Bmatrix} \sin(\eta \ln r) \\ \cos(\eta \ln r) \end{Bmatrix}\right) \text{ for complex } \lambda = \xi + i\eta$$

$$\text{satisfying } \frac{d^{n-m}D}{d\lambda^{n-m}} = 0,\ m<n,\ 0<Re\lambda\leq 1,$$

as $r \rightarrow 0$ on R, where $\mathbf{M} = (M_\theta, M_r, M_{r\theta})$ is the vector of moment resultants. In (7), D is the determinant of the coefficient matrix stemming from the substitution of (6) into a set of edge conditions drawn from (4) and n is the order of this matrix, m its rank. For any particular combination of the edge conditions (4) for a wedge angle α, the values of λ in the ranges given in (7) may be regarded as the singular *eigenvalues* of the *eigenequation*, $D=0$. We now investigate the eigenequations resulting from such expansions.

Eigenequations

For the particular problem of the symmetric bending of a stress-free/stress-free wedge, substituting (6) with $b_1 = b_3 = b_5 = 0$ therein into the first of (4) and expanding the determinant of the resulting 3×3 coefficient matrix leads to

$$\sin(\lambda+1)\alpha/2(\lambda\sin\alpha + \sin\lambda\alpha) = 0(0<\alpha\leq 2\pi). \tag{8}$$

Equation (8) factors into two equations; however, the first of these, while not generating a completely trivial solution, does not give rise to any moment resultants and therefore contributes no singular fields. Consequently, it may be discarded leaving as our eigenequation for this case only the second factor. The eigenequations for the remaining combinations of the edge conditions each possess similar, non-singular, multiplicative factors. In Table 1, we suppress these and list only those parts of each eigenequation that have attendant singular fields.

Comparison of the first three cases in Table 1 with the corresponding extensional cases given in Williams [4] shows the eigenequations to be identical. Examining the fourth and fifth cases in Table 1 and noting that the conditions for the simply supported edge in (4) are the same as anti-symmetry requirements, we see these eigenequations are equivalent to the anti-symmetric parts of the first and second eigenequations, respectively, for a wedge of angle 2α. It follows that these two cases are also effectively contained in Williams' extensional analysis [4]. Finally, taking as the physical analogue of the simply supported/simply supported edge condition, the extensional anti-symmetry conditions, $u_r = 0$, $\sigma_\theta = 0$ where u_r is the radial displacement and σ_θ is the tangential normal stress, we find that the last case too has a corresponding extensional eigenequation. The significance of this correspondence is that discussions in the literature on the extensional eigenequations are directly applicable to elastic wedges generated by Reissner's bending theory.

Solutions for the dominant singular real part of λ are given for the first three cases in Table 1 by Williams [4] and are decomposed in effect into symmetric and anti-symmetric parts by Kalandiia [6], thus accounting for the first five cases. The roots for the last case, simply supported/simply supported, can be determined by inspection. Some of the eigenvalues given in [4, 6] are actually the real parts of complex solutions; however, no truly comprehensive search for complex roots appears to be available in the literature. Such a parameter study is outside the scope of the present work. Nonetheless, for any given application the determination of complex eigenvalues proceeds routinely on separating the pertinent eigenequation for the specific α-value into real and imaginary parts and solving the resulting, simple, simultaneous pair of transcendental equations. Likewise, logarithmic singularities have not been exhaustively searched for, but are straight forward to check for in any specific instance.

In conclusion we remark that the correspondence between the singular fields in Reissner's theory and those in extensional plate theory is not restricted merely to the singular eigenvalues, but carries over to the actual eigenfunctions which share the same r *and* θ dependences, as can be deduced from the solution (6) and its counterpart in Williams [4].

References

1 Williams, M. L., "Surface Stress Singularities Resulting From Various Boundary Conditions in Angular Corners of Plates Under Bending," *Proceedings of the First U.S. National Congress of Applied Mechanics*, 1951, pp. 325-329.

2 Reissner, E., "The Effect of Transverse Shear Deformation on the Bending of Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 12, 1945, pp. A-69-A-77.

3 Wang, N-M., "Effects of Plate Thickness on the Bending of an Elastic Plate Containing a Crack," *Journal of Mathematics and Physics*, Vol. 47, 1968, pp. 371-390.

4 Williams, M. L., "Stress Singularities Resulting from Various Boundary Conditions in Angular Corners of Plates in Extension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 19, 1952, pp. 526-528.

5 Dempsey, J. P., and Sinclair, G. B., "On the Stress Singularities in the Plane Elasticity of the Composite Wedge," *Journal of Elasticity*, Vol. 9, 1979, pp. 373-391.

6 Kalandiia, A. L., "Remarks on the Singularity of Elastic Solutions Near Corners," *Prikladnaia Matematika i Mekhanika*, Vol. 33, 1969, pp. 132-135.

Dynamic Behavior of Beam Structures Carrying Moving Masses

S. Saigal[1]

Introduction

The dynamic response of structures carrying moving masses is a problem of widespread practical significance. A detailed survey of research efforts in this field was given by Stanišić et al. [2]. The original problem is nonlinear in both local and convective derivatives [3] and is complicated by the presence of a Dirac-Delta function as a coefficient in the differential equation of motion. Previous methods [2] applied for the solution of this problem are approximate in nature and tedious in their hierarchy of mathematical operation. Recently, Stanišić [3] expressed the solution in terms of eigenfunctions satisfying the boundary, initial and transient conditions, for a heavy mass moving over a simply supported beam. However, in engineering practice there are problems that involve more complex boundary conditions and, therefore, it is of phenomenological interest to look into the physics of the dynamical behavior of a clamped and a cantilever beam under the action of heavy moving masses. The present study extends Stanišić's theory [3] to study the dynamic behavior of a

---

[1] School of Aeronautics and Astronautics, Purdue University, West Lafayette, Indiana 47907. Student Member, ASME.

fining attention to singular solutions, the resulting conditions may be summarized by

$\mathbf{M}=0(r^{\lambda-1})$ for real λ satisfying $D=0,\ 0<\lambda<1$,

$$\mathbf{M}=0\left(r^{\lambda-1}\begin{Bmatrix}\sin(\eta lnr)\\ \cos(\eta lnr)\end{Bmatrix}\right)\text{ for complex }\lambda=\xi+i\eta$$

$$\text{satisfying } D=0,\ 0<\Delta Re\lambda<1, \qquad (7)$$

$\mathbf{M}=0(r^{\lambda-1}lnr)$ for real λ satisfying

$$\frac{d^{n-m}D}{d\lambda^{n-m}}=0,\ m<n,\ 0<\lambda\leq 1,$$

$$\mathbf{M}=0\left(r^{\lambda-1}lnr\begin{Bmatrix}\sin(\eta lnr)\\ \cos(\eta lnr)\end{Bmatrix}\right)\text{ for complex }\lambda=\xi+i\eta$$

$$\text{satisfying }\frac{d^{n-m}D}{d\lambda^{n-m}}=0,\ m<n,\ 0<Re\lambda\leq 1,$$

as $r\to 0$ on R, where $\mathbf{M}=(M_\theta, M_r, M_{r\theta})$ is the vector of moment resultants. In (7), D is the determinant of the coefficient matrix stemming from the substitution of (6) into a set of edge conditions drawn from (4) and n is the order of this matrix, m its rank. For any particular combination of the edge conditions (4) for a wedge angle α, the values of λ in the ranges given in (7) may be regarded as the singular *eigenvalues* of the *eigenequation*, $D=0$. We now investigate the eigenequations resulting from such expansions.

Eigenequations

For the particular problem of the symmetric bending of a stress-free/stress-free wedge, substituting (6) with $b_1=b_3=b_5=0$ therein into the first of (4) and expanding the determinant of the resulting 3×3 coefficient matrix leads to

$$\sin(\lambda+1)\alpha/2(\lambda\sin\alpha+\sin\lambda\alpha)=0(0<\alpha\leq 2\pi). \qquad (8)$$

Equation (8) factors into two equations; however, the first of these, while not generating a completely trivial solution, does not give rise to any moment resultants and therefore contributes no singular fields. Consequently, it may be discarded leaving as our eigenequation for this case only the second factor. The eigenequations for the remaining combinations of the edge conditions each possess similar, non-singular, multiplicative factors. In Table 1, we suppress these and list only those parts of each eigenequation that have attendant singular fields.

Comparison of the first three cases in Table 1 with the corresponding extensional cases given in Williams [4] shows the eigenequations to be identical. Examining the fourth and fifth cases in Table 1 and noting that the conditions for the simply supported edge in (4) are the same as anti-symmetry requirements, we see these eigenequations are equivalent to the anti-symmetric parts of the first and second eigenequations, respectively, for a wedge of angle 2α. It follows that these two cases are also effectively contained in Williams' extensional analysis [4]. Finally, taking as the physical analogue of the simply supported/simply supported edge condition, the extensional anti-symmetry conditions, $u_r=0$, $\sigma_\theta=0$ where u_r is the radial displacement and σ_θ is the tangential normal stress, we find that the last case too has a corresponding extensional eigenequation. The significance of this correspondence is that discussions in the literature on the extensional eigenequations are directly applicable to elastic wedges generated by Reissner's bending theory.

Solutions for the dominant singular real part of λ are given for the first three cases in Table 1 by Williams [4] and are decomposed in effect into symmetric and anti-symmetric parts by Kalandiia [6], thus accounting for the first five cases. The roots for the last case, simply supported/simply supported, can be determined by inspection. Some of the eigenvalues given in [4, 6] are actually the real parts of complex solutions; however, no truly comprehensive search for complex roots appears to be available in the literature. Such a parameter study is outside the scope of the present work. Nonetheless, for any given application the determination of complex eigenvalues proceeds routinely on separating the pertinent eigenequation for the specific α-value into real and imaginary parts and solving the resulting, simple, simultaneous pair of transcendental equations. Likewise, logarithmic singularities have not been exhaustively searched for, but are straight forward to check for in any specific instance.

In conclusion we remark that the correspondence between the singular fields in Reissner's theory and those in extensional plate theory is not restricted merely to the singular eigenvalues, but carries over to the actual eigenfunctions which share the same *r and* θ dependences, as can be deduced from the solution (6) and its counterpart in Williams [4].

References

1 Williams, M. L., "Surface Stress Singularities Resulting From Various Boundary Conditions in Angular Corners of Plates Under Bending," *Proceedings of the First U.S. National Congress of Applied Mechanics*, 1951, pp. 325–329.

2 Reissner, E., "The Effect of Transverse Shear Deformation on the Bending of Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 12, 1945, pp. A-69–A-77.

3 Wang, N-M., "Effects of Plate Thickness on the Bending of an Elastic Plate Containing a Crack," *Journal of Mathematics and Physics*, Vol. 47, 1968, pp. 371–390.

4 Williams, M. L., "Stress Singularities Resulting from Various Boundary Conditions in Angular Corners of Plates in Extension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 19, 1952, pp. 526–528.

5 Dempsey, J. P., and Sinclair, G. B., "On the Stress Singularities in the Plane Elasticity of the Composite Wedge," *Journal of Elasticity*, Vol. 9, 1979, pp. 373–391.

6 Kalandiia, A. L., "Remarks on the Singularity of Elastic Solutions Near Corners," *Prikladnaia Matematika i Mekhanika*, Vol. 33, 1969, pp. 132–135.

Dynamic Behavior of Beam Structures Carrying Moving Masses

S. Saigal[1]

Introduction

The dynamic response of structures carrying moving masses is a problem of widespread practical significance. A detailed survey of research efforts in this field was given by Stanišić et al. [2]. The original problem is nonlinear in both local and convective derivatives [3] and is complicated by the presence of a Dirac-Delta function as a coefficient in the differential equation of motion. Previous methods [2] applied for the solution of this problem are approximate in nature and tedious in their hierarchy of mathematical operation. Recently, Stanišić [3] expressed the solution in terms of eigenfunctions satisfying the boundary, initial and transient conditions, for a heavy mass moving over a simply supported beam. However, in engineering practice there are problems that involve more complex boundary conditions and, therefore, it is of phenomenological interest to look into the physics of the dynamical behavior of a clamped and a cantilever beam under the action of heavy moving masses. The present study extends Stanišić's theory [3] to study the dynamic behavior of a

---

[1] School of Aeronautics and Astronautics, Purdue University, West Lafayette, Indiana 47907. Student Member, ASME.

clamped and a cantilever beam under moving masses. The effects on response of the beam of magnitude and location of the moving mass on the structure and the convergence of amplitude coefficients of the series are studied.

Equation of Motion

The equation of motion of a beam on an elastic foundation subject to a moving mass M can be written in the nondimensional form, neglecting the convective part of acceleration, as [3]

$$\frac{\partial^4 Y(z,\tau)}{\partial z^4} + [1+\epsilon\delta(z-z_o)]\frac{\partial^2 Y(z,\tau)}{\partial \tau^2} + \kappa Y(z,\tau) = P\delta(z-z_o) \qquad (1)$$

where

$$z=\frac{x}{L};\ z_o(t)=\frac{x_o(t)}{L};\ \epsilon=\frac{M}{\rho AL};\ \tau=\frac{t}{\alpha};$$

$$\alpha^2=\frac{\rho AL^4}{EI};\ \kappa=\frac{kL^4}{EI}\ \ P=\frac{MgL^2}{EI};\ \ Y(x,t)=\frac{y(x,t)}{L};$$

EI = flexural rigidity of the beam; ρ = density of the material of the beam; A = area of cross-section of the beam; L = length of the beam; k = coefficient of elastic foundation; g = acceleration due to gravity; $y(x,t)$ = deflection of the beam at location x at time t; $\delta(\ )$ = Dirac-Delta function; $x_o(t)$ = position of the mass at time t. Also, we have,

a) Boundary conditions: $Y(0,\tau)=Y'(0,\tau)=Y(1,\tau)=Y'(1,\tau)=0$ for clamped beam and $Y(0,\tau)=Y'(0,\tau)=Y''(1,\tau)=Y'''(1,\tau)=0$ for cantilever beam.

b) Initial Conditions: $Y(z,0)=\dot{Y}(z,0)=0$, where primes (′) and dots ($\dot{}$) denote the derivatives with respect to z and τ, respectively.

Eigenfunctions

Assuming $Y(z,\tau)=Z(z)e^{i\Omega\tau}$, where Ω is the dimensionless frequency, the homogeneous part of (1) leads to

$$\frac{d^4Z(z)}{dz^4}-(\Omega^2-\kappa)Z(z)=\epsilon\Omega^2Z(z)\delta(z-z_o) \qquad (2)$$

The eigenfunctions are of the form of Green's function [3] and in the following two domains the homogeneous part of (2), i.e.,

$$\frac{d^4Z(z)}{dz^4}-(\Omega^2-\kappa)Z(z)=0$$

is considered

a) $0\le z<z_o,\quad Z(0)=Z'(0)=0$

b) $z_o<z\le 1$, $Z(1)=Z'(1)=0$ for clamped beam, and $Z''(1)=Z'''(1)=0$ for cantilever beam. The solutions can be written as:

a) For Clamped Beam

$$Z=A(z_o)\begin{cases}(\cos\lambda z-\cosh\lambda z)+a_1(z_o)(\sin\lambda z-\sinh\lambda z), & 0\le z<z_o\\ a_2(z_o)[\cos\lambda(z-1)-\cosh\lambda(z-1)]+a_3(z_o)[\sin\lambda(z-1)-\sinh\lambda(z-1)], & z_o<z\le 1\end{cases} \qquad (3)$$

b) For Cantilever Beam

$$Z=B(z_o)\begin{cases}(\cos\lambda z-\cosh\lambda z)+b_1(z_o)(\sin\lambda z-\sinh\lambda z), & 0\le z<z_o\\ b_2(z_o)[\cos\lambda(z-1)+\cosh\lambda(z-1)]+b_3(z_o)[\sin\lambda(z-1)+\sinh\lambda(z-1)], & z_o<z\le 1\end{cases} \qquad (4)$$

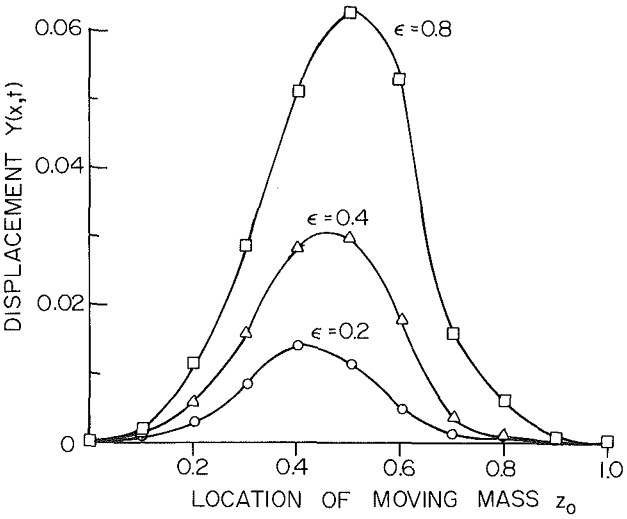


Fig. 1 Displacement of the clamped beam under the moving mass for various values of the mass ratio ϵ

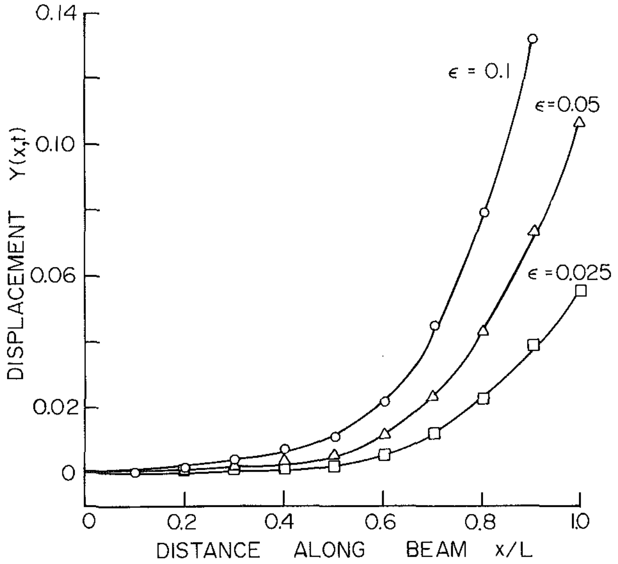


Fig. 2 Displacement of the cantilever beam under the moving mass for various values of the mass ratio ϵ

The constants a_1, a_2, and a_3 for the clamped beam and the constants b_1, b_2, and b_3 for the cantilever beam are determined from the conditions of continuity:

$$Z(z,z_o)\Big|_{z_o^-}^{z_o^+}=\frac{dZ(z,z_o)}{dz}\Big|_{z_o^-}^{z_o^+}=\frac{d^2Z(z,z_o)}{dz^2}\Big|_{z_o^-}^{z_o^+}=0 \qquad (5)$$

Integrating (2) and using the accompanying boundary conditions, the transient condition representing the jump of the shearing force at $z=z_o$ is obtained as

$$\frac{d^3Z(z,z_o)}{dz^3}\Big|_{z^-}^{z^+}=\epsilon\Omega^2Z(z_o,z_o) \qquad (6)$$

where $\Omega^2=\lambda^4-\kappa$. This condition is used in determining the dimensionless frequency Ω. Finally, the unknowns $A(z_o)$ and $B(z_o)$ for the clamped and cantilever beam, respectively, are determined by imposing the condition of orthonormality with the weight function $1+\epsilon\delta(z-z_o)$ [3] as

$$\int_0^{z_o} Z^2(z,z_o)dz+\int_{z_o}^{1} Z^2(z,z_o)dz+\epsilon Z^2(z_o,z_o)=1 \qquad (7)$$

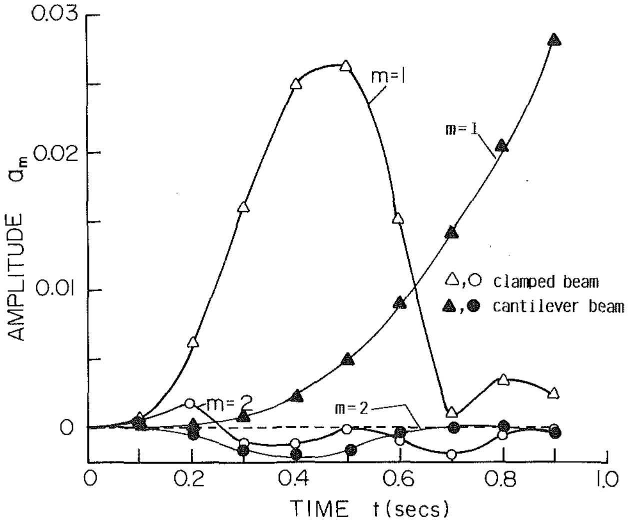


Fig. 3 Comparison of amplitude coefficients a_1 and a_2 for the first and second mode, respectively, at various times for mass ratio $\epsilon = 0.4$ (magnify vertical scale by 10 for the cantilever beam results)

Response of the Beam

The solution of (1) can be expressed as a series

$$Y(z,\tau) = \sum_{m=1}^{\infty} a_m(\tau,z_o) Z_m(z,z_o)$$

where $Z_m(z,z_o)$ are the orthonormal eigenfunctions described in the preceding section; $a_m(\tau, z_o)$ are the amplitude coefficients. Substituting in (1) and making the assumptions suggested by Stanišić based on numerical evidence [3] that $\partial Z_m(z,z_o)/\partial z_o$ and $\partial a_m(\tau,z_o)/\partial z_o$ together with their second order derivatives are very small so that their products can be neglected compared with $\partial^2 a_m(\tau,z_o)/\partial\tau^2 \cdot Z_m(z,z_o)$, the equation for coefficients a_m is obtained in the differential form as

$$\frac{\partial^2 a_m(\tau,z_o)}{\partial \tau^2} + \Omega_m^2(z_o) a_m(\tau,z_o) = P Z_m(z_o,z_o)$$

with

$$a_m(0,z_o) = \dot{a}_m(0,z_o) = 0$$

giving

$$a_m(\tau,z_o) = \int_o^\tau g(\tau;\theta) P Z_m(z_o,z_o) d\theta \tag{8}$$

where $g(\tau;\theta)$ is the Green's function such that

$$\frac{\partial^2 g(\tau;\theta)}{\partial \theta^2} + \Omega_m^2(z_o) g(\tau;\theta) = \delta(\tau-\theta)$$

and

$$g(\tau;\tau_f) = \frac{\partial g(\tau;\tau_f)}{\partial \theta} = 0.$$

$\tau_f = t_f/\alpha$; t_f = time taken by the moving mass to cross the entire beam. In addition, the Green's function has to satisfy the transient conditions

$$g(\tau;\theta)\Big|_{\tau^-}^{\tau^+} = 0 \text{ and } \frac{\partial g(\tau;\theta)}{\partial \theta}\Big|_{\tau^-}^{\tau^+} = 1.$$

The amplitude coefficients a_m can then be determined using (8).

Numerical Results and Discussions

The formulations described above were used to determine the dynamic behavior of beams with $\alpha = 2.17$, $v = 6.096$ m/s (20 ft/s), L = 6.096 m (20 ft). The integrations in (7) and (8) were carried out numerically using Gauss integration of the tenth order [1]. The knowledge of the deflection of the structure under the moving mass as the mass travels across the structure is of interest to the structural engineer in highway and bridge construction. Such deflections are plotted in Figs. 1 and 2 for the clamped and the cantilever beam, respectively. The nonlinear nature of the dynamic response with increase in mass ratio is evident from these curves. The plots of the amplitude coefficients a_1 and a_2 corresponding to the first two modes in the series expansion for the clamped and the cantilever beam are given in Fig. 3. The drop of magnitude of a_2 compared to that of a_1 suggests a fast convergence of the series. In performing the above calculations, it was observed that the magnitude of the frequency measure Ω was affected largely by the position of the mass on the beam and changed only slightly with the magnitude of the mass.

References

1 Abramowitz, M., and Stegun, I. A., 1970, *Handbook of Mathematical Functions*, Dover Publications, New York, pp. 875–924.

2 Stanišić, M. M., Euler, J. A., and Montgomery, S. T., 1974, "On a Theory Concerning the Dynamical Behavior of Structures Carrying Moving Masses," *Ingenieur-Archiv*, Vol. 43, pp. 295–305.

3 Stanišić, M. M., 1985, "On a New Theory of the Dynamic Behavior of the Structures Carrying Moving Masses," *Ingenieur-Archiv*, to appear.

A More Direct and General Analysis of Moving Strong Discontinuity Surfaces in Quasi-Statically Deforming Elastic-Plastic Solids

W. J. Drugan[1]

1 Introduction

In a recent study, Drugan and Rice (1984) investigated what restrictions are placed on the possible existence of quasi-statically moving surfaces of strong discontinuity (across which components of stress, strain or velocity jump) by standard weak continuum mechanical assumptions coupled with *skeletal constitutive assumptions* believed to describe realistically a large class of elastic-plastic materials. Using a small displacement-gradient formulation, they proved that the standard set of assumptions examined requires all components of stress to be continuous across such propagating surfaces, and that only certain components of the *plastic* part of the strain tensor may jump provided specific conditions are met.

I develop here a more direct and general version of Drugan and Rice's (1984) main proof which lays bare its key features. This facilitates a demonstration that the severe restrictions deduced by Drugan and Rice do not hinge on their assumption of elastic linearity, but rather that the elastic component of material response can be arbitrary hyperelastic, with the key restrictions being that the elastic strain energy function is strictly convex and unaffected by plastic deformation.

2 Formulation

With reference to Fig. 1, let Σ denote a hypothesized surface of strong discontinuity that propagates with velocity $V > 0$ in the direction of the normal, x_1, through an elastic-plastic solid under general three-dimensional conditions. The Cartesian

---

[1] Department of Engineering Mechanics, University of Wisconsin, Madison, WI 53706. Assoc. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, February 21, 1985; final revision July 12, 1985.

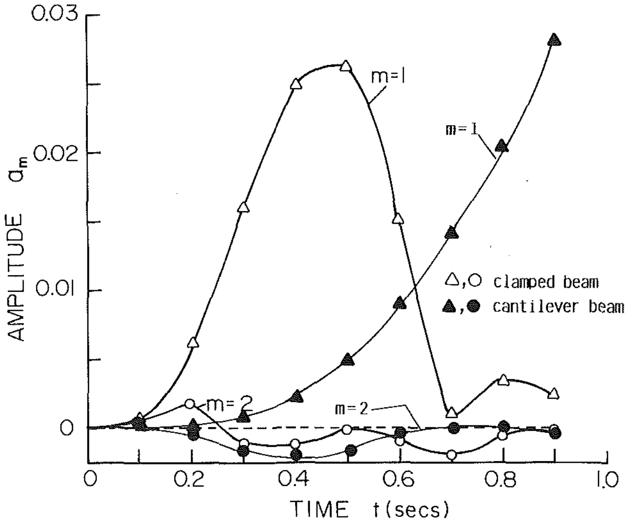


Fig. 3 **Comparison of amplitude coefficients a_1 and a_2 for the first and second mode, respectively, at various times for mass ratio $\epsilon = 0.4$ (magnify vertical scale by 10 for the cantilever beam results)**

Response of the Beam

The solution of (1) can be expressed as a series

$$Y(z,\tau) = \sum_{m=1}^{\infty} a_m(\tau,z_o) Z_m(z,z_o)$$

where $Z_m(z,z_o)$ are the orthonormal eigenfunctions described in the preceding section; $a_m(\tau, z_o)$ are the amplitude coefficients. Substituting in (1) and making the assumptions suggested by Stanišić based on numerical evidence [3] that $\partial Z_m(z,z_o)/\partial z_o$ and $\partial a_m(\tau,z_o)/\partial z_o$ together with their second order derivatives are very small so that their products can be neglected compared with $\partial^2 a_m(\tau,z_o)/\partial\tau^2 \cdot Z_m(z,z_o)$, the equation for coefficients a_m is obtained in the differential form as

$$\frac{\partial^2 a_m(\tau,z_o)}{\partial\tau^2} + \Omega_m^2(z_o) a_m(\tau,z_o) = P Z_m(z_o,z_o)$$

with

$$a_m(0,z_o) = \dot{a}_m(0,z_o) = 0$$

giving

$$a_m(\tau,z_o) = \int_o^{\tau} g(\tau;\theta) P Z_m(z_o,z_o) d\theta \tag{8}$$

where $g(\tau;\theta)$ is the Green's function such that

$$\frac{\partial^2 g(\tau;\theta)}{\partial\theta^2} + \Omega_m^2(z_o) g(\tau;\theta) = \delta(\tau-\theta)$$

and

$$g(\tau;\tau_f) = \frac{\partial g(\tau;\tau_f)}{\partial\theta} = 0.$$

$\tau_f = t_f/\alpha$; t_f = time taken by the moving mass to cross the entire beam. In addition, the Green's function has to satisfy the transient conditions

$$g(\tau;\theta)\Big|_{\tau^-}^{\tau^+} = 0 \text{ and } \frac{\partial g(\tau;\theta)}{\partial\theta}\Big|_{\tau^-}^{\tau^+} = 1.$$

The amplitude coefficients a_m can then be determined using (8).

Numerical Results and Discussions

The formulations described above were used to determine the dynamic behavior of beams with $\alpha = 2.17$, $v = 6.096$ m/s (20 ft/s), L = 6.096 m (20 ft). The integrations in (7) and (8) were carried out numerically using Gauss integration of the tenth order [1]. The knowledge of the deflection of the structure under the moving mass as the mass travels across the structure is of interest to the structural engineer in highway and bridge construction. Such deflections are plotted in Figs. 1 and 2 for the clamped and the cantilever beam, respectively. The nonlinear nature of the dynamic response with increase in mass ratio is evident from these curves. The plots of the amplitude coefficients a_1 and a_2 corresponding to the first two modes in the series expansion for the clamped and the cantilever beam are given in Fig. 3. The drop of magnitude of a_2 compared to that of a_1 suggests a fast convergence of the series. In performing the above calculations, it was observed that the magnitude of the frequency measure Ω was affected largely by the position of the mass on the beam and changed only slightly with the magnitude of the mass.

References

1 Abramowitz, M., and Stegun, I. A., 1970, *Handbook of Mathematical Functions*, Dover Publications, New York, pp. 875–924.

2 Stanišić, M. M., Euler, J. A., and Montgomery, S. T., 1974, "On a Theory Concerning the Dynamical Behavior of Structures Carrying Moving Masses," *Ingenieur-Archiv*, Vol. 43, pp. 295–305.

3 Stanišić, M. M., 1985, "On a New Theory of the Dynamic Behavior of the Structures Carrying Moving Masses," *Ingenieur-Archiv*, to appear.

A More Direct and General Analysis of Moving Strong Discontinuity Surfaces in Quasi-Statically Deforming Elastic-Plastic Solids

W. J. Drugan[1]

1 Introduction

In a recent study, Drugan and Rice (1984) investigated what restrictions are placed on the possible existence of quasi-statically moving surfaces of strong discontinuity (across which components of stress, strain or velocity jump) by standard weak continuum mechanical assumptions coupled with skeletal *constitutive assumptions* believed to describe realistically a large class of elastic-plastic materials. Using a small displacement-gradient formulation, they proved that the standard set of assumptions examined requires all components of stress to be continuous across such propagating surfaces, and that only certain components of the *plastic* part of the strain tensor may jump provided specific conditions are met.

I develop here a more direct and general version of Drugan and Rice's (1984) main proof which lays bare its key features. This facilitates a demonstration that the severe restrictions deduced by Drugan and Rice do not hinge on their assumption of elastic linearity, but rather that the elastic component of material response can be arbitrary hyperelastic, with the key restrictions being that the elastic strain energy function is strictly convex and unaffected by plastic deformation.

2 Formulation

With reference to Fig. 1, let Σ denote a hypothesized surface of strong discontinuity that propagates with velocity $V > 0$ in the direction of the normal, x_1, through an elastic-plastic solid under general three-dimensional conditions. The Cartesian

---

[1] Department of Engineering Mechanics, University of Wisconsin, Madison, WI 53706. Assoc. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, February 21, 1985; final revision July 12, 1985.

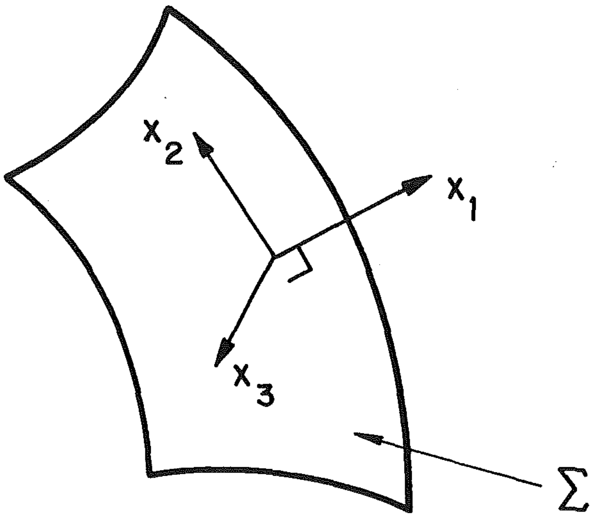


Fig. 1 Quasi-statically propagating surface of strong discontinuity

coordinate system x_1, x_2, x_3 moves with the surface Σ. Values of a field quantity, say $g(x_1, x_2, x_3, t)$, where t is time, directly ahead of and directly behind the moving surface Σ will be denoted as $g^{\pm} \equiv \lim_{\mu \downarrow 0} g(x_1, x_2, x_3, t_a \mp \mu)$, respectively, where t_a is the time at which Σ arrives at a particular material point. The jump in such a field quantity across Σ will be denoted as $[[g]] \equiv g^+ - g^-$.

In the sequel, components of tensors with respect to the Cartesian coordinate system of Fig. 1 are indicated either by the Latin indices i, j, k, l which have range 1, 2, 3, or by the Greek indices α, β which have range 2, 3 only and thus refer to tensor components in planes parallel to planes that are tangent to Σ. Both types of index follow the summation convention.

Referring the reader interested in a more detailed and justified development to Drugan and Rice (1984), I here summarize the jump conditions implied by standard continuum-mechanical postulates, and the weak constitutive assumptions to be employed.

Equilibrium requires traction continuity across Σ, so that

$$[[\sigma_{1i}]] = 0 \tag{1}$$

where $\sigma_{ij} = \sigma_{ji}$ is the stress tensor. I assume continuity of the displacement vector across Σ, $[[u_i]] = 0$. In terms of the small-strain tensor, $\epsilon_{ij} \equiv (1/2)\,(\partial u_i/\partial x_j + \partial u_j/\partial x_i)$, this requires

$$[[\epsilon_{1\alpha}]] = \frac{1}{2}[[\partial u_\alpha/\partial x_1]] \tag{2}$$

$$[[\epsilon_{\alpha\beta}]] = 0. \tag{3}$$

The constitutive assumptions to be employed are identical to those of Drugan and Rice (1984), except that instead of requiring the elastic part of a total strain increment to be *linearly* related to a stress increment, this relationship can be arbitrarily nonlinear so long as the response is hyperelastic with a strictly convex elastic strain energy function that is unaffected by plastic deformation. Thus, I assume that a total strain increment can be additively decomposed into elastic and plastic parts

$$d\epsilon_{ij} = d\epsilon^e_{ij} + d\epsilon^p_{ij}, \tag{4}$$

where the elastic part is related to the stress increment by

$$d\sigma_{ij} = \frac{\partial^2 \phi(\epsilon^e)}{\partial \epsilon^e_{ij}\,\partial \epsilon^e_{kl}}\, d\epsilon^e_{kl} \tag{5}$$

with $\phi(\epsilon^e)$ being the positive-definite, single-valued elastic strain energy function.

The plastic behavior of the materials considered is assumed to be in accord with the maximum plastic work inequality

$$(\sigma_{ij} - \sigma^o_{ij})\, d\epsilon^p_{ij} \geq 0 \tag{6}$$

where σ_{ij} is the stress state (at yield) corresponding to the plastic strain increment $d\epsilon^p_{ij}$, and σ^o_{ij} is any other stress state that is at or below yield. I note that assumptions (4) through (6) permit very general elastic and plastic anisotropy.

3 Analysis

A key step in the Drugan and Rice (1984) analysis is the realization that integration of the maximum plastic work inequality (6) at a material point just during the passage of Σ can provide an explicit restriction on jumps in components of stress and strain. Thus, we examine

$$\int_{\epsilon^{p+}}^{\epsilon^{p-}} (\sigma_{ij} - \sigma^+_{ij})\, d\epsilon^p_{ij} \geq 0, \tag{7}$$

which follows from (6) whenever $\boldsymbol{\sigma}^+$ remains on or inside the yield surface for all states along the strain path from $\boldsymbol{\epsilon}^+$ to $\boldsymbol{\epsilon}^-$. This will always be true, e.g., for materials whose current yield locus incorporates all prior yield loci, so (7) clearly permits isotropic hardening as well as many types of anisotropic hardening including many cases of yield surface vertex formation.

Employ (4) to rewrite (7) in the instructive form

$$\int_{\epsilon^+}^{\epsilon^-} (\sigma_{ij} - \sigma^+_{ij})\, d\epsilon_{ij} - \int_{\epsilon^{e+}}^{\epsilon^{e-}} (\sigma_{ij} - \sigma^+_{ij})\, d\epsilon^e_{ij} \geq 0. \tag{8}$$

The clarity of the present derivation lies in the fact that the first integral in (8) vanishes identically because I restrict possible paths from $\boldsymbol{\epsilon}^+$ to $\boldsymbol{\epsilon}^-$ to those for which (1) and (3) are satisfied by all states traversed along the path. The vanishing of this integral is easily verified by noting that the integrand simplifies as

$$(\sigma_{ij} - \sigma^+_{ij})\, d\epsilon_{ij} = (\sigma_{\alpha\beta} - \sigma^+_{\alpha\beta})\, d\epsilon_{\alpha\beta} \tag{9}$$

since σ_{1i} are continuous across Σ via (1), and finally that (9) vanishes identically since (3) requires $d\epsilon_{\alpha\beta} = 0$ across Σ.

Now, (8) has reduced to the illuminating restriction

$$-\int_{\epsilon^{e+}}^{\epsilon^{e-}} (\sigma_{ij} - \sigma^+_{ij})\, d\epsilon^e_{ij} \geq 0. \tag{10}$$

At this point one realizes that the assumption of hyperelastic behavior will permit direct evaluation of this integral; employing (5), we have

$$\phi(\epsilon^{e-}) - \phi(\epsilon^{e+}) - \sigma^+_{ij}(\epsilon^{e-}_{ij} - \epsilon^{e+}_{ij}) \leq 0. \tag{11}$$

The form of this expression implies immediately that if strict convexity of $\phi(\epsilon^e)$ is required, then

$$[[\epsilon^e_{ij}]] = 0. \tag{12}$$

This is true because strict convexity of $\phi(\epsilon^e)$ means that it must satisfy (11) with the inequality sign reversed, the equality holding only when $\epsilon^{e-}_{ij} \equiv \epsilon^{e+}_{ij}$. Given (12), (5) requires

$$[[\sigma_{ij}]] = 0. \tag{13}$$

Thus I have proven two of the main results of Drugan and Rice (1984), namely that all components of stress and elastic strain must be continuous across a quasi-statically moving surface in an elastic-plastic solid, under the more general conditions that the elastic part of the response may be nonlinear hyperelastic with a strictly convex elastic strain energy function. The interested reader is referred to Drugan and Rice (1984) for derivations of restrictions on jumps in components of velocity and plastic strain for some important material models; I observe that all of these additional results remain valid for the more general constitutive assumptions made here, given the above proofs of (12) and (13).

Acknowledgment

Support of this work by the US National Science Foundation, Solid Mechanics Program, under Grant MEA-8312348 is gratefully acknowledged.

References

Drugan, W. J., and Rice, J. R., 1984 "Restrictions on Quasi-Statically Moving Surfaces of Strong Discontinuity in Elastic-Plastic Solids," in *Mechanics of Material Behavior: The Daniel C. Drucker Anniversary Volume*, G. J. Dvorak and R. T. Shield, eds., Elsevier, Amsterdam, pp. 59–73.

A Note on the Obtainment of Instantaneous Penetration Information From Final Penetration Data

O. E. R. Heimdahl[1] and J. C. Schulz[1]

A simple, practical technique is described for generating instantaneous penetration depth versus velocity information for a given projectile-target combination from experimentally obtained final penetration depth versus impact velocity data. The technique applies to penetrations where the resistive pressure is a function of velocity only.

Introduction

Final penetration depth versus impact velocity data for a projectile-target combination are as a rule easily obtained experimentally. Measurements of instantaneous penetration depth as a function of instantaneous velocity tend to be more difficult. (See, however, [1, 2, 3].) This note describes a technique, applicable to a fairly broad class of penetrations, whereby final penetration depth versus impact velocity data can be used to generate instantaneous penetration depth versus velocity at any desired impact velocity for a projectile-target combination. To the authors' knowledge this very simple technique, which appears to have practical application, has not been described previously in the literature.

Description of Technique

The equation of motion for a rigid body penetrating a material stably and without yaw is often assumed to be of the form

$$-m\,dv/dt = -mv\,dv/ds = A\,g(v) \tag{1}$$

where m = projectile mass
A = presented area
s = instantaneous penetration depth
v = velocity
t = time
$g(v)$ = resistive pressure.

The assumption is that the resistive pressure is a function of velocity only. The Poncelet, Petry and Young equations [4] are all of this form, differing only in the particular expression used for the resistive pressure.

Integration of (1) yields

$$s = m/A[G(V) - G(v)] \tag{2}$$

where

V = impact velocity (velocity at $s = 0$)
$G(v) = \int v/g(v)\,dv$.

The final penetration depth, P, obtained by setting $v = 0$ in (2), is

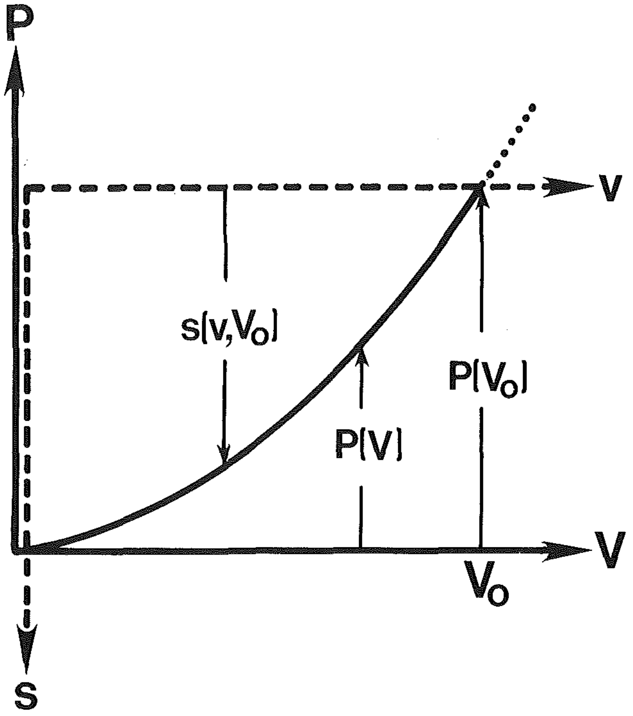


Fig. 1 Relationship between $P-V$ and $s-v$ axes

$$P = m/A[G(V) - G(0)]. \tag{3}$$

Combination of (2) and (3) yields the important result

$$s(v,V) = P(V) - P(v). \tag{4}$$

In words, (4) says that the instantaneous penetration depth equals the final penetration depth corresponding to the impact velocity minus the final penetration depth corresponding to the instantaneous velocity. This result may be arrived at heuristically by observing that at any time during a penetration obeying (1) the distance of penetration remaining depends only on the current velocity.

Interpreted graphically (4) indicates that the $s-v$ curve is a reflection and translation of the $P-V$ curve. Thus, the instantaneous penetration depth as a function of velocity can be read directly from a plot of final penetration depth versus impact velocity simply by repositioning the axes. This is illustrated in Fig. 1 where the $P-V$ axes are shown solid and the $s-v$ axes are shown dashed. The amount of translation, $P(V_o)$, depends on the particular impact velocity, V_o.

The resistive pressure can be determined from numerical or graphical differentiation of experimental $P-V$ data through the relation

$$g(v) = mv/A/(dP/dv). \tag{5}$$

The time corresponding to a given velocity can be estimated as

$$t = -m/A\int_V^v 1/g(v)\,dv. \tag{6}$$

Errors associated with numerical or graphical integration limit accuracy of the time estimate.

Discussion

A technique has been described which allows determination of "instantaneous" penetration information from more easily obtained "final" penetration results. That is to say, if final penetration depth versus impact velocity data is available, then instantaneous penetration depth as a function of velocity for a given impact velocity can be generated. In addition, the resistive pressure and the time as functions of velocity can also

---

[1] Naval Weapons Center, China Lake, Calif. 93555

Manuscript received by ASME Applied Mechanics Division, July 2, 1985; final revision August 26, 1985.

tion, Solid Mechanics Program, under Grant MEA-8312348 is gratefully acknowledged.

References

Drugan, W. J., and Rice, J. R., 1984 "Restrictions on Quasi-Statically Moving Surfaces of Strong Discontinuity in Elastic-Plastic Solids," in *Mechanics of Material Behavior: The Daniel C. Drucker Anniversary Volume*, G. J. Dvorak and R. T. Shield, eds., Elsevier, Amsterdam, pp. 59–73.

A Note on the Obtainment of Instantaneous Penetration Information From Final Penetration Data

O. E. R. Heimdahl[1] and J. C. Schulz[1]

A simple, practical technique is described for generating instantaneous penetration depth versus velocity information for a given projectile-target combination from experimentally obtained final penetration depth versus impact velocity data. The technique applies to penetrations where the resistive pressure is a function of velocity only.

Introduction

Final penetration depth versus impact velocity data for a projectile-target combination are as a rule easily obtained experimentally. Measurements of instantaneous penetration depth as a function of instantaneous velocity tend to be more difficult. (See, however, [1, 2, 3].) This note describes a technique, applicable to a fairly broad class of penetrations, whereby final penetration depth versus impact velocity data can be used to generate instantaneous penetration depth versus velocity at any desired impact velocity for a projectile-target combination. To the authors' knowledge this very simple technique, which appears to have practical application, has not been described previously in the literature.

Description of Technique

The equation of motion for a rigid body penetrating a material stably and without yaw is often assumed to be of the form

$$-m\,dv/dt = -mv\,dv/ds = A\,g(v) \qquad (1)$$

where m = projectile mass
A = presented area
s = instantaneous penetration depth
v = velocity
t = time
$g(v)$ = resistive pressure.

The assumption is that the resistive pressure is a function of velocity only. The Poncelet, Petry and Young equations [4] are all of this form, differing only in the particular expression used for the resistive pressure.

Integration of (1) yields

$$s = m/A[G(V) - G(v)] \qquad (2)$$

where

V = impact velocity (velocity at $s = 0$)
$G(v) = \int v/g(v)\,dv$.

The final penetration depth, P, obtained by setting $v = 0$ in (2), is

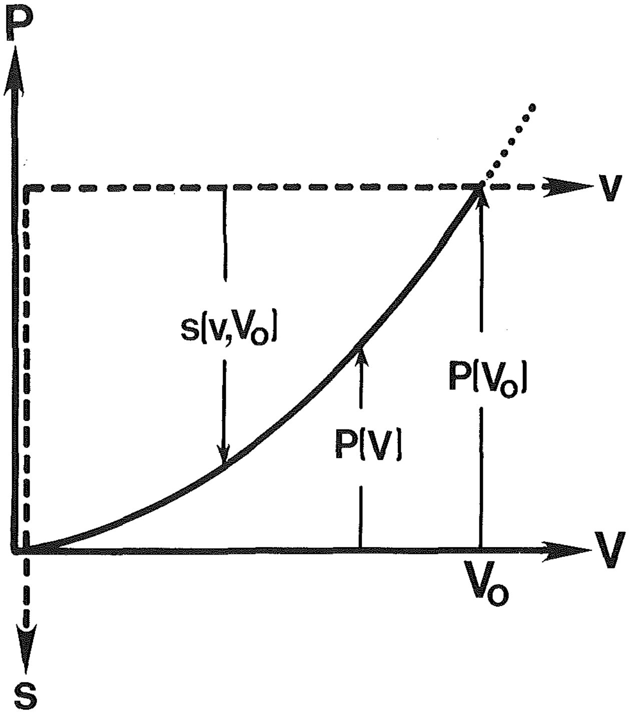


Fig. 1 Relationship between $P-V$ and $s-v$ axes

$$P = m/A[G(V) - G(0)]. \qquad (3)$$

Combination of (2) and (3) yields the important result

$$s(v,V) = P(V) - P(v). \qquad (4)$$

In words, (4) says that the instantaneous penetration depth equals the final penetration depth corresponding to the impact velocity minus the final penetration depth corresponding to the instantaneous velocity. This result may be arrived at heuristically by observing that at any time during a penetration obeying (1) the distance of penetration remaining depends only on the current velocity.

Interpreted graphically (4) indicates that the $s-v$ curve is a reflection and translation of the $P-V$ curve. Thus, the instantaneous penetration depth as a function of velocity can be read directly from a plot of final penetration depth versus impact velocity simply by repositioning the axes. This is illustrated in Fig. 1 where the $P-V$ axes are shown solid and the $s-v$ axes are shown dashed. The amount of translation, $P(V_o)$, depends on the particular impact velocity, V_o.

The resistive pressure can be determined from numerical or graphical differentiation of experimental $P-V$ data through the relation

$$g(v) = mv/A/(dP/dv). \qquad (5)$$

The time corresponding to a given velocity can be estimated as

$$t = -m/A \int_V^v 1/g(v)\,dv. \qquad (6)$$

Errors associated with numerical or graphical integration limit accuracy of the time estimate.

Discussion

A technique has been described which allows determination of "instantaneous" penetration information from more easily obtained "final" penetration results. That is to say, if final penetration depth versus impact velocity data is available, then instantaneous penetration depth as a function of velocity for a given impact velocity can be generated. In addition, the resistive pressure and the time as functions of velocity can also

---

[1] Naval Weapons Center, China Lake, Calif. 93555

Manuscript received by ASME Applied Mechanics Division, July 2, 1985; final revision August 26, 1985.

be determined. The technique is amenable to either numerical or graphical application.

The assumption is made that the resistive pressure is a function of velocity only. More generally, this pressure will depend on time and/or depth in addition to velocity. A material with a memory (viscoelastic, viscoplastic, etc.) or a material with properties (density, strength, etc.) that vary in the penetration direction will violate this assumption. However, for penetration events for which the assumption applies, this technique allows extraction of additional useful information from test data.

References

1 Murff, J. C., and Coyle, H. M., "Low Velocity Penetration of Kaolin Clay," *Journal of Soil Mechanics and Foundations Division of the ASCE*, Vol. 99, 1973, pp. 375-389.

2 Longcope, D. B., and Forrestal, M. J., "Penetration of Targets Described by a Mohr-Coulomb Failure Criterion with a Tension Cutoff," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, 1983, pp. 327-333.

3 Forrestal, M. J., Lee, L. M., Jenrette, B. D., and Setchell, R. E., "Gas-Gun Experiments Determine Forces on Penetrators into Geological Targets," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, 1984, pp. 602-607.

4 Backman, M. E., and Goldsmith, W., "The Mechanics of Penetration of Projectiles into Targets," *International Journal of Engineering Science*, Vol. 16, 1978, pp. 38-39.

Journal of Applied Mechanics

Discussion

Stability of Short Columns Subjected to Distributed Axial Loads[1]

J. Rondal[2] and R. Maquoi.[3] The authors are to be congratulated for their valuable contribution to the stability problem of columns subjected to distributed axial forces.

Though of limited practical interest in civil engineering, the case of uniformly distributed axial forces is much more encountered in connection with offshore activities.

The aim of the present discussion is, on the one hand, to bring additional theoretical reference values and, on the other hand, to comment briefly on the limitations of the results at the light of elastoplastic behavior of steel material. First, the last column of Table 1 in the author's paper can be completed by means of results published originally in reference [1].

In addition, it must be emphasized that the critical buckling load is meaningful as long as it does not exceed the compression yield load [2, 3]. Thus, the range of low slenderness ratios is no more governed by the critical buckling stress and the author's conclusions are, therefore, questionable. The deviation between these values of λ and those for slender columns [1] is expressed in percents as:

$$\Delta\lambda(\text{percent}) = 100\,\frac{\lambda(\text{slender column}) - \lambda(L/r)}{\lambda(\text{slender column})} \tag{1}$$

and is plotted in Fig. 1. It decreases rapidly with L/r and tends to zero. Furthermore, for mild steel, which has f_y equal to 235 MPa and is commonly used for steel construction in civil engineering, the transition between the squash load and the Eulerian critical load occurs for following slenderness ratio

$$L/r(\text{mild steel}) = \pi\,\sqrt{210{,}000/235} = 93.91 \tag{2}$$

For any kind of boundary conditions, the maximum discrepancy reaches only 1 percent, and can be disregarded

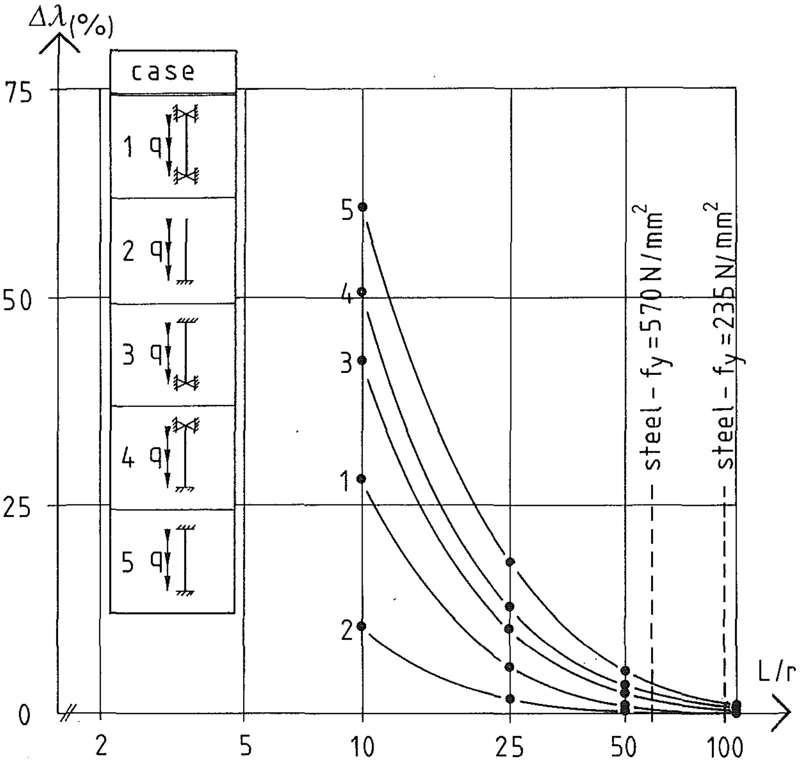


Fig. 1 Relative discrepancy between author's values and values for slender columns

---

[1] By Shastry, B. P., and Venkateswara Rao, G., and published in the March 1985 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 229–230.

[2] Associate Professor, Institute of Civil Engineering, University of Liège, Quai Banning, 6, 4000 Liège, Belgium.

[3] Associate Professor, Institute of Civil Engineering, University of Liège, Quai Banning, 6, 4000 Liège, Belgium.

(Fig. 1). For higher steel grades, for instance with a yield stress of 570 MPa, that is commonly used for casing and tubing, the discrepancy is about 3.5 percent and is still negligible.

References

1 Tseng Huang, and Dareing, D. W., "Buckling and Frequencies of Long Vertical Pipes," ASCE, *Journal of the Engineering Mechanics Division*, Vol. 95, 1969, pp. 167-181.

2 Bjorhovde, R., "The Safety of Steel Columns," ASCE, *Journal of the Structural Division*, Vol. 104, 1978, pp. 463-477.

3 Rondal, J., and Maquoi, R., "Formulations d'AYRTON-PERRY pour le flambement des barres métalliques," *Construction Métallique*, No. 4, 1979, pp. 41-53.

Authors' Closure

The authors appreciate Professors J. Rondal and R. Maquois' interest in their paper and additional references cited and agree that yield criterion is to be considered while dealing with the stability of very short columns. However, as suggested by the discussers, materials having high yield strength may represent columns with low L/r. For example, the higher grade steel quoted in the discussion represents a cantilever with L/r as low as 30. The theme of the paper was to provide stability parameters for a wide range of short columns and hence considered the range of L/r from 10 to 500 for the numerical results. The authors thank the discussers for the graphical representation of the numerical results for practical considerations.

Sufficient Conditions for Small Particles to Hold Together Because of Adhesion Forces[1]

J. Reed.[2] In a recent article Fichman and Pnueli (1985) presented a particle impact-adhesion model for elastic-plastic impacts. There are some similarities between this model and another recent model (Rogers and Reed, 1984). The Fichman and Pnueli approach appears to be an improvement over Rogers and Reed since they calculate the limiting elastic velocity using a pressure distribution including adhesive forces whereas in Rogers and Reed adhesive forces are neglected. Unfortunately there is a serious error in their equations.

Fichman and Pneuli's equation (7) has been incorrectly quoted from Johnson, Kendall and Roberts (1971). Consequently, some of their subsequent equations are also incorrect. Equation (9) in Fichman and Pneuli is

$$\delta = \frac{2a^2}{3R} - \frac{2}{3}\left(\frac{6\pi\gamma a}{K}\right)^{1/2} \tag{1}$$

where δ is the distance of approach of the centre of the two bodies, a, is the contact radius, R is the geometric radius of the two bodies in contact, γ is the adhesive energy per unit area of both bodies and

$$K = \frac{4}{3}\left[\frac{(1-\nu_1{}^2)}{E_1} + \frac{(1-\nu_2{}^2)}{E_2}\right]^{-1}$$

where ν is Poisson's ratio and E Young's modulus.

In the limit where the adhesive forces are negligible the above equation should reduce to the classical Hertz equation (Timoshenko and Goodier, 1971) which is

$$\delta = \frac{a^2}{R} \tag{2}$$

Obviously as $\gamma \to 0$ the Fichman and Pnueli equation gives $\delta \to 2a^2/3R$ which is not the correct limit.

The Fichman and Pnueli equation (9) ought to be (as seen from Rogers and Reed, 1984, equation (13) or Barquins, 1983 equation (7))

$$\delta = \frac{a^2}{R} - \frac{2}{3}\frac{(P_1 - P_0)}{Ka} \tag{3}$$

$$= \frac{a^2}{3R}\left(1 + \frac{2P_o}{P_1}\right) \tag{4}$$

where P_o is the applied load, and P_1 is given by

$$P_1 = P_o + 3\pi\gamma R + (6\pi\gamma R P_o + (3\pi\gamma R)^2) \tag{5}$$

If the adhesive energy is negligible then $P_1 = P_o$. Therefore, equation (4) approaches the Hertz solution as expected.

The correct equations linking the critical elastic impact velocity, V_{e1}, and the elastic yield limit, p^*, following the same method as Fichman and Pnueli, are

$$\frac{1}{2}M\,V_{e1}^2$$

$$= \frac{1}{9K^{2/3}R^{1/3}}\int_o^{P_{1_{e1}}} \frac{[P_1 - (6\pi\gamma R P_1)^{1/2}][6P_1 - (6\pi\gamma R P_1)^{1/2}]}{P_1{}^{4/3}}\,dP_1$$

$$= \frac{1}{9K^{2/3}R^{1/3}}\left\{\frac{18}{5}P_{1_{e1}}^{5/3} - 6(6\pi\gamma R)^{1/2}P_{1_{e1}}^{7/6} + 9\ \pi\gamma R\ P_{1_e}^{2/3}\right\} \tag{6}$$

and

$$p^* = \frac{1}{2\pi}\frac{K^{2/3}}{R^{2/3}}(3P_{1_{e1}}^{1/3} - (6\pi\gamma R)^{1/2}P_{1_{e1}}^{-1/6}) \tag{7}$$

Equations (6) and (7) above ought to be used to link V_{e1} and p^* rather than equations (12) and (20) of Fichman and Pnueli.

Because their equation (9) was incorrect this also means that the analysis used to calculate the plastic work in the annulus at the edge of the contact area ϵ_{pl_2} (equations (23) to (32)) is also incorrect.

Finally, the two energy balance equations ((21) and (22)) are misleading. We feel equation (22) has a typographical error in it. It would make better sense if it read

$$\epsilon_{KIN} \leq \epsilon_{pl_2} + H(\epsilon_{KIN} - \epsilon_{el})\cdot\epsilon_{pl_1}$$

Unfortunately, it is not clear anywhere in the article how ϵ_{pl_1} is calculated. The energy balance criterion in equation (21) serves no purpose other than giving a sticking condition in an artificial situation (i.e. when the maximum pressure in the centre of the contact zone just equals the elastic yield limit).

References

Barquins, M., 1983, "Adhesive Contact and Kinetics of Adherence Between a Rigid Sphere and an Elastomeric Solid," International Journal of Adhesion and Adhesives, Vol. 3, pp. 72-84.

Fichman, M., and Pnueli, D., 1985, "Sufficient Conditions for Small Particles to Hold Together because of Adhesion Forces," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 105-108.

Johnson, K. A., Kendall, K., and Roberts, A. D., 1971, "Surface Energy of the Contact of Elastic Solids," *Proceedings of the Royal Society*, Series A, Vol. 324, pp. 301-313.

Rogers, L. N., and Reed, J., 1984, "The Adhesion of Particles Undergoing an Elastic Plastic Impact with a Surface," *Journal Physics D.*, Vol. 17, pp. 667-689.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw Hill, New York, p. 410.

---

[1] By M. Fichman and D. Pnueli, and published in the March 1985 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 105-108.

[2] Berkeley Nuclear Laboratories, CEGB, Berkeley, Gloucestershire GL13 9PB, England.

(Fig. 1). For higher steel grades, for instance with a yield stress of 570 MPa, that is commonly used for casing and tubing, the discrepancy is about 3.5 percent and is still negligible.

References

1 Tseng Huang, and Dareing, D. W., "Buckling and Frequencies of Long Vertical Pipes," ASCE, *Journal of the Engineering Mechanics Division*, Vol. 95, 1969, pp. 167-181.

2 Bjorhovde, R., "The Safety of Steel Columns," ASCE, *Journal of the Structural Division*, Vol. 104, 1978, pp. 463-477.

3 Rondal, J., and Maquoi, R., "Formulations d'AYRTON-PERRY pour le flambement des barres métalliques," *Construction Métallique*, No. 4, 1979, pp. 41-53.

Authors' Closure

The authors appreciate Professors J. Rondal and R. Maquois' interest in their paper and additional references cited and agree that yield criterion is to be considered while dealing with the stability of very short columns. However, as suggested by the discussers, materials having high yield strength may represent columns with low L/r. For example, the higher grade steel quoted in the discussion represents a cantilever with L/r as low as 30. The theme of the paper was to provide stability parameters for a wide range of short columns and hence considered the range of L/r from 10 to 500 for the numerical results. The authors thank the discussers for the graphical representation of the numerical results for practical considerations.

Sufficient Conditions for Small Particles to Hold Together Because of Adhesion Forces[1]

J. Reed.[2] In a recent article Fichman and Pnueli (1985) presented a particle impact-adhesion model for elastic-plastic impacts. There are some similarities between this model and another recent model (Rogers and Reed, 1984). The Fichman and Pnueli approach appears to be an improvement over Rogers and Reed since they calculate the limiting elastic velocity using a pressure distribution including adhesive forces whereas in Rogers and Reed adhesive forces are neglected. Unfortunately there is a serious error in their equations.

Fichman and Pneuli's equation (7) has been incorrectly quoted from Johnson, Kendall and Roberts (1971). Consequently, some of their subsequent equations are also incorrect. Equation (9) in Fichman and Pneuli is

$$\delta = \frac{2a^2}{3R} - \frac{2}{3}\left(\frac{6\pi\gamma a}{K}\right)^{1/2} \tag{1}$$

where δ is the distance of approach of the centre of the two bodies, a, is the contact radius, R is the geometric radius of the two bodies in contact, γ is the adhesive energy per unit area of both bodies and

$$K = \frac{4}{3}\left[\frac{(1-\nu_1{}^2)}{E_1} + \frac{(1-\nu_2{}^2)}{E_2}\right]^{-1}$$

where ν is Poisson's ratio and E Young's modulus.

In the limit where the adhesive forces are negligible the above equation should reduce to the classical Hertz equation (Timoshenko and Goodier, 1971) which is

$$\delta = \frac{a^2}{R} \tag{2}$$

Obviously as $\gamma \to 0$ the Fichman and Pnueli equation gives $\delta \to 2a^2/3R$ which is not the correct limit.

The Fichman and Pnueli equation (9) ought to be (as seen from Rogers and Reed, 1984, equation (13) or Barquins, 1983 equation (7))

$$\delta = \frac{a^2}{R} - \frac{2}{3}\frac{(P_1 - P_0)}{Ka} \tag{3}$$

$$= \frac{a^2}{3R}\left(1 + \frac{2P_o}{P_1}\right) \tag{4}$$

where P_o is the applied load, and P_1 is given by

$$P_1 = P_o + 3\pi\gamma R + (6\pi\gamma R P_o + (3\pi\gamma R)^2) \tag{5}$$

If the adhesive energy is negligible then $P_1 = P_o$. Therefore, equation (4) approaches the Hertz solution as expected.

The correct equations linking the critical elastic impact velocity, V_{e1}, and the elastic yield limit, p^*, following the same method as Fichman and Pnueli, are

$$\frac{1}{2}M\,V_{e1}^2$$

$$= \frac{1}{9K^{2/3}R^{1/3}}\int_o^{P_{1_{e1}}} \frac{[P_1 - (6\pi\gamma R P_1)^{1/2}][6P_1 - (6\pi\gamma R P_1)^{1/2}]}{P_1{}^{4/3}}\,dP_1$$

$$= \frac{1}{9K^{2/3}R^{1/3}}\left\{\frac{18}{5}P_{1_{e1}}^{5/3} - 6(6\pi\gamma R)^{1/2}P_{1_{e1}}^{7/6} + 9\,\pi\gamma R\,P_{1_e}^{2/3}\right\} \tag{6}$$

and

$$p^* = \frac{1}{2\pi}\frac{K^{2/3}}{R^{2/3}}(3P_{1_{e1}}^{1/3} - (6\pi\gamma R)^{1/2}P_{1_{e1}}^{-1/6}) \tag{7}$$

Equations (6) and (7) above ought to be used to link V_{e1} and p^* rather than equations (12) and (20) of Fichman and Pnueli.

Because their equation (9) was incorrect this also means that the analysis used to calculate the plastic work in the annulus at the edge of the contact area ϵ_{pl_2} (equations (23) to (32)) is also incorrect.

Finally, the two energy balance equations ((21) and (22)) are misleading. We feel equation (22) has a typographical error in it. It would make better sense if it read

$$\epsilon_{KIN} \le \epsilon_{pl_2} + H(\epsilon_{KIN} - \epsilon_{el}) \cdot \epsilon_{pl_1}$$

Unfortunately, it is not clear anywhere in the article how ϵ_{pl_1} is calculated. The energy balance criterion in equation (21) serves no purpose other than giving a sticking condition in an artificial situation (i.e. when the maximum pressure in the centre of the contact zone just equals the elastic yield limit).

References

Barquins, M., 1983, "Adhesive Contact and Kinetics of Adherence Between a Rigid Sphere and an Elastomeric Solid," International Journal of Adhesion and Adhesives, Vol. 3, pp. 72-84.

Fichman, M., and Pnueli, D., 1985, "Sufficient Conditions for Small Particles to Hold Together because of Adhesion Forces," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 105-108.

Johnson, K. A., Kendall, K., and Roberts, A. D., 1971, "Surface Energy of the Contact of Elastic Solids," *Proceedings of the Royal Society*, Series A, Vol. 324, pp. 301-313.

Rogers, L. N., and Reed, J., 1984, "The Adhesion of Particles Undergoing an Elastic Plastic Impact with a Surface," *Journal Physics D.*, Vol. 17, pp. 667-689.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw Hill, New York, p. 410.

---

[1] By M. Fichman and D. Pnueli, and published in the March 1985 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 105-108.

[2] Berkeley Nuclear Laboratories, CEGB, Berkeley, Gloucestershire GL13 9PB, England.

Authors' Closure

a) As pointed out by Reed, Fichman and Pnueli's (1985) work is an improvement[1] on Rogers and Reed (1984). Indeed, in addition to the elastic-plastic deformation at the center of the circle of contact, we have considered the plastic deformation at the edge of that circle. This plastic deformation always exists during impact, unlike plastic dissipation at the center of the circle, which appears only after the yield point there has been reached. We do not show how to calculate ϵ_{pl_1}, the dissipation work at the center of the circle, because it is not needed to obtain the condition for the two particles sticking together. Once the yield point at the center has been reached – any surplus kinetic energy is dissipated. Thus, only the energy dissipated at the edge of the circle, ϵ_{pl_2}, is required. This result has a very interesting implication: particles with diameters smaller than some critical values always stick together, whatever their velocities might be.

b) We regret that there is a type-setting error and the second term on the right-hand side of our equation (7) has been omitted. This equation (7) should read:

$$p_o = \frac{3}{2} k\delta a_1 - \frac{P_1}{2} \tag{1}$$

This equation agrees with Reed's comments.

c) Another typographical error has caused a 2/3 constant to appear in the first term of our equation (9), which corrected should read (the whole equation):

$$\delta = \frac{a_1^2}{R} \pm \frac{2}{3}\sqrt{\frac{6\gamma\pi}{k}}\, a_1^{\frac{1}{2}} \tag{2}$$

Unfortunately, this error was made before the paper was submitted, and influenced other expressions. R in equations (11), (12), (29) ÷ (32) must be corrected to:

$$R = \frac{2}{3} \bullet \frac{R_1 R_2}{R_1 + R_2} \tag{3}$$

d) Finally, Rogers and Reed (1984) does not have an equation (13) in it, and equation (5) in Reed's Comments should be:

$$P_1 = P_o + 3\pi\gamma R + \sqrt{6\pi R P_o + (3\pi\gamma R)^2}$$

and not as written; which just shows that mistakes will happen. In conclusion, we are very grateful to Dr. Reed for his comments, which did bring about these necessary corrections.

References

Same as in Reed's Comments.

A New Rate Principle Suitable for Analysis of Inelastic Deformation of Plates and Shells[2]

J. N. Reddy[3]. Much of the paper is a review of well-known variational principles of elasticity which can be found in a number of books (e.g., see Oden and Reddy, 1976). The 'new variational principle' presented in the paper is not new, and can be found in the monograph by Oden and Reddy (1976). More specifically, the functional Π_M in equation (11) of the paper is exactly the same as that in equation (4.115) on page 115 of this reference. The monograph also contains a number of related (fourteen in total) functionals for linear elasticity (on pp. 114–119) (Oden and Reddy, 1974), and variational principles for viscoelasticity (on pp. 143–169) and nonlinear elasticity and inelasticity (on pp. 173–189). Also, the title of the paper is not justified because the authors have not presented any analytical discussion of the specialization of the elasticity principles to plates and shells. Only a qualitative discussion of a possible extention is presented.

References

Oden, J. T., and Reddy, J. N., "On Dual Complementary Variational Principles in Mathematical Physics," *Int. J. Engng. Sci.*, Vol. 12, 1974, pp. 1–29.

Oden, J. T., and Reddy, J. N., *Variational Methods in Theoretical Mechanics*, Springer-Verlag, Berlin, 1976 (2nd Ed., 1982).

Authors' Closure

We thank Professor Reddy for his interest in our work [1]. We regret that the work of Oden and Reddy did not come to our attention prior to publication of our paper and that we ended up rediscovering their variational principle. The review of existing variational principles, which occupied half a page of our paper, was given in order to set the stage for what was to follow. These, of course, were clearly referenced ([5] and [6] of our paper). Further, we purposely restricted ourselves, in this paper, to a qualitative discussion of the application of our principle to inelastic shells. We have indicated in the paper that "a strictly two-dimensional formulation containing vector and tensor fields referred to the base vectors of the undeformed shell midsurface" would be published elsewhere ([8] of our paper). This paper [2] has just been published. Professor Reddy had communicated his concern about our paper, in a private letter to one of us, soon after it was published last September. We immediately replied to him and sent him a preprint of our Acta Mechanica paper. In view of this, we are really quite surprised to find his continuing concern regarding an analytical treatment of this variational principle for inelastic shells, as voiced in the last two sentences of his discussion.

References

1 Mukherjee, S., and Kollmann, F. G., 1985, "A New Rate Principle Suitable for Analysis of Inelastic Deformation of Plates and Shells," ASME Journal of Applied Mechancis, Vol. 52, pp. 533–535.

2 Kollmann, F. G., and Mukherjee, S., 1985, "A General Geometrically Linear Theory of Inelastic Thin Shells," *Acta Mechanica*, Vol. 57, pp. 41–67.

The Problem of Minimizing Stress Concentration at a Rigid Inclusion[1]

G. S. Bjorkman, Jr.[2] **and R. Richards Jr.**[3] The objective of this paper, as stated by Professor Wheeler, "is an optimization problem aimed at finding the best shape for a rigid inclusion imbedded in an elastic matrix of infinite extent, if the stress concentration is to be minimized." It should be emphasized that the solution to this problem had already been obtained earlier by the writers (Bjorkman and Richards, 1979; Richards and Bjorkman, 1980). In these two works the writers found that the rigid-inclusion shape which satisfied the harmonic field condition (i.e., the condition that the first invariant of the original stress (or strain) field remain unperturbed everywhere in the field) in a biaxial field is an ellipse whose axes are inversely proportional to the principal normal strains of the original field (i.e., $a/b = \epsilon_2/\epsilon_1$) irrespective of plane stress or plane strain. This is precisely the result Professor Wheeler obtains in equation (4.1) in less-

---

[1] But do not rely on it. Their paper was submitted long before Rogers and Reed's was published.

[2] By S. Mukherjee and F. G. Kollmann, and published in September 1985 issue of ASME Journal of Applied Mechanics, Vol. 52, pp. 533–535.

[3] Clifton C. Garvin Professor of Engineering Science and Mechanics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

---

[1] By L. Wheeler and published in the March 1985 issue of the ASME Journal of Applied Mechanics, Vol. 52, pp. 83–86.

[2] Senior Consultant, Cygna Energy Services, Boston, MA 02210

[3] Professor of Civil Engineering, State University of New York at Buffalo, Buffalo, N.Y. 14260

Authors' Closure

a) As pointed out by Reed, Fichman and Pnueli's (1985) work is an improvement[1] on Rogers and Reed (1984). Indeed, in addition to the elastic-plastic deformation at the center of the circle of contact, we have considered the plastic deformation at the edge of that circle. This plastic deformation always exists during impact, unlike plastic dissipation at the center of the circle, which appears only after the yield point there has been reached. We do not show how to calculate ϵ_{pl_1}, the dissipation work at the center of the circle, because it is not needed to obtain the condition for the two particles sticking together. Once the yield point at the center has been reached – any surplus kinetic energy is dissipated. Thus, only the energy dissipated at the edge of the circle, ϵ_{pl_2}, is required. This result has a very interesting implication: particles with diameters smaller than some critical values always stick together, whatever their velocities might be.

b) We regret that there is a type-setting error and the second term on the right-hand side of our equation (7) has been omitted. This equation (7) should read:

$$p_o = \frac{3}{2} k\delta a_1 - \frac{P_1}{2} \tag{1}$$

This equation agrees with Reed's comments.

c) Another typographical error has caused a 2/3 constant to appear in the first term of our equation (9), which corrected should read (the whole equation):

$$\delta = \frac{a_1^2}{R} \pm \frac{2}{3}\sqrt{\frac{6\gamma\pi}{k}}\, a_1^{\frac{1}{2}} \tag{2}$$

Unfortunately, this error was made before the paper was submitted, and influenced other expressions. R in equations (11), (12), (29) ÷ (32) must be corrected to:

$$R = \frac{2}{3} \cdot \frac{R_1 R_2}{R_1 + R_2} \tag{3}$$

d) Finally, Rogers and Reed (1984) does not have an equation (13) in it, and equation (5) in Reed's Comments should be:

$$P_1 = P_o + 3\pi\gamma R + \sqrt{6\pi R P_o + (3\pi\gamma R)^2}$$

and not as written; which just shows that mistakes will happen. In conclusion, we are very grateful to Dr. Reed for his comments, which did bring about these necessary corrections.

References

Same as in Reed's Comments.

A New Rate Principle Suitable for Analysis of Inelastic Deformation of Plates and Shells[2]

J. N. Reddy[3]. Much of the paper is a review of well-known variational principles of elasticity which can be found in a number of books (e.g., see Oden and Reddy, 1976). The 'new variational principle' presented in the paper is not new, and can be found in the monograph by Oden and Reddy (1976). More specifically, the functional Π_M in equation (11) of the paper is exactly the same as that in equation (4.115) on page 115 of this reference. The monograph also contains a number of related (fourteen in total) functionals for linear elasticity (on pp. 114–119) (Oden and Reddy, 1974), and variational principles for viscoelasticity (on pp. 143–169) and nonlinear elasticity and inelasticity (on pp. 173–189). Also, the title of the paper is not justified because the authors have not presented any analytical discussion of the specialization of the elasticity principles to plates and shells. Only a qualitative discussion of a possible extention is presented.

References

Oden, J. T., and Reddy, J. N., "On Dual Complementary Variational Principles in Mathematical Physics," *Int. J. Engng. Sci.*, Vol. 12, 1974, pp. 1–29.

Oden, J. T., and Reddy, J. N., *Variational Methods in Theoretical Mechanics*, Springer-Verlag, Berlin, 1976 (2nd Ed., 1982).

Authors' Closure

We thank Professor Reddy for his interest in our work [1]. We regret that the work of Oden and Reddy did not come to our attention prior to publication of our paper and that we ended up rediscovering their variational principle. The review of existing variational principles, which occupied half a page of our paper, was given in order to set the stage for what was to follow. These, of course, were clearly referenced ([5] and [6] of our paper). Further, we purposely restricted ourselves, in this paper, to a qualitative discussion of the application of our principle to inelastic shells. We have indicated in the paper that "a strictly two-dimensional formulation containing vector and tensor fields referred to the base vectors of the undeformed shell midsurface" would be published elsewhere ([8] of our paper). This paper [2] has just been published. Professor Reddy had communicated his concern about our paper, in a private letter to one of us, soon after it was published last September. We immediately replied to him and sent him a preprint of our Acta Mechanica paper. In view of this, we are really quite surprised to find his continuing concern regarding an analytical treatment of this variational principle for inelastic shells, as voiced in the last two sentences of his discussion.

References

1 Mukherjee, S., and Kollmann, F. G., 1985, "A New Rate Principle Suitable for Analysis of Inelastic Deformation of Plates and Shells," ASME Journal of Applied Mechancis, Vol. 52, pp. 533–535.

2 Kollmann, F. G., and Mukherjee, S., 1985, "A General Geometrically Linear Theory of Inelastic Thin Shells," *Acta Mechanica*, Vol. 57, pp. 41–67.

The Problem of Minimizing Stress Concentration at a Rigid Inclusion[1]

G. S. Bjorkman, Jr.[2] **and R. Richards Jr.**[3] The objective of this paper, as stated by Professor Wheeler, "is an optimization problem aimed at finding the best shape for a rigid inclusion imbedded in an elastic matrix of infinite extent, if the stress concentration is to be minimized." It should be emphasized that the solution to this problem had already been obtained earlier by the writers (Bjorkman and Richards, 1979; Richards and Bjorkman, 1980). In these two works the writers found that the rigid-inclusion shape which satisfied the harmonic field condition (i.e., the condition that the first invariant of the original stress (or strain) field remain unperturbed everywhere in the field) in a biaxial field is an ellipse whose axes are inversely proportional to the principal normal strains of the original field (i.e., $a/b = \epsilon_2/\epsilon_1$) irrespective of plane stress or plane strain. This is precisely the result Professor Wheeler obtains in equation (4.1) in less-

---

[1] But do not rely on it. Their paper was submitted long before Rogers and Reed's was published.

[2] By S. Mukherjee and F. G. Kollmann, and published in September 1985 issue of ASME Journal of Applied Mechanics, Vol. 52, pp. 533–535.

[3] Clifton C. Garvin Professor of Engineering Science and Mechanics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

[1] By L. Wheeler and published in the March 1985 issue of the ASME Journal of Applied Mechanics, Vol. 52, pp. 83–86.

[2] Senior Consultant, Cygna Energy Services, Boston, MA 02210

[3] Professor of Civil Engineering, State University of New York at Buffalo, Buffalo, N.Y. 14260

Journal of Applied Mechanics

Book Reviews

Review of Progress in Quantitative Nondestructive Evaluation. Vol. 3A and 3B. Edited by D. O. Thompson and D. E. Chimenti. Plenum Publishing Corp., New York, 1984. 1489 pages. Price $225.

REVIEWED BY S. K. DATTA[1]

This volume contains the texts of presentations at the annual Review of Progress in Quantitative NDE held at the University of California, Santa Cruz, Aug. 7-12, 1983. There are 142 papers and the transcript of the panel discussion on inversion and reconstruction. The editors currently noted in the Preface: "This Review, possibly the most correctly comprehensive annual symposium in NDE, emphasizes both basic research and early engineering applications; it provides a valuable forum for the transfer of technical information." In this volume the reader will learn about the latest developments in NDE modeling and applications. Because the emphasis here is on the current state of the art, the papers tend to be sketchy. However, overall they provide a valuable source of information for those involved in NDE research and applications.

The Volume is organized into 9 chapters dealing with the following topics: NDE Reliability, Ultrasonics (Section A: Probability of Detection; Section B: Scattering; Section C: Sizing; Section D: Transducers; Section E: Signal Processing; Section F: Imaging and Reconstruction), Eddy Currents (Section A: Probability of Detection; Section B: Modeling; Section C: Sizing; Section D: Probes), Acoustic Emission, Thermal Wave Imaging, and Optical Methods, Inverse Methods, Composite Materials, Material Properties, Acoustoelasticity, and Residual Stress. In addition to these there are two chapters dealing with new techniques and new NDE Systems.

The reviewer believes that this volume will be a useful addition to the libraries of the practitioners in NDE.

Handbook of Heat Transfer Fundamentals. Second Edition. By W. M. Rohsenow, J. P. Hartnett and E. N. Ganic. McGraw-Hill, New York, 1985. 1440 pages. Price $95.

REVIEWED BY P. D. RICHARDSON[2]

Heat transfer plays an important role in almost every technological process. Experiment and analysis in this subject flourished even in the 19th century. Having personally entered the subject in the 1950's, and having prepared very recently a chapter for another Handbook, I approached this review with relish.

As subjects grow and develop they go through phases. In heat transfer the use of dimensional analysis led to simplifying approaches – parametric representation of differential equations for theoretical work, and use of dimensionless numbers for plotting experimental data – which found their summaries in McAdams' "Heat Transmission" (1933). This work and its later editions served as handbooks as well as demonstrations of the applicability of analytical methods to practical problems. The period 1945-1965 brought new technological challenges – supersonic flight, space flight and nuclear reactors prominent among them – and a new tool, the digital computer. This latter was a foil for additional experiments of improved accuracy because comparison with more accurate and reliable calculations was possible. This applied particularly to laminar forced convection. (Turbulent forced convection also benefited from introduction of computational approaches, but accuracy and broad reliability, while improved, do not match that for laminar flows.) The 197-page Chapter 8 (Forced Convection, External Flows) illustrates this time frame in its references: for the 1950's there are 8 citations per year, rising in the early 1960's to about 15 per year and then falling progressively to 1.5 per year in the 1980's. By contrast, the surge in Numerical Methods in Heat Transfer (Chapter 5) is more recent; citations for the early 1960's run close to 1 per year but accelerate in the later 1960's and run close to 10 per year in the early 1970's. A major factor in this was the introduction of finite element methods.

In 1973 the first edition (and single volume) of the Handbook of Heat Transfer was published. Research has progressed since then. The volume under review is nearly 1500 pages long, 2 1/2 inches (63 mm) thick, in English and SI units, with 14 Chapters, and is the first of two volumes for the second edition. It has almost 700 illustrations. A fine example of how well subject matter has been updated is given by Warren Rohsenow in his 94-page Chapter 12 on Boiling. The studies of the 1960's (which included major advances in phenomenological understanding of boiling) are represented by over 10 citations per year, and even in the latter half of the 1970's the citations run about 8 per year. Chapters not mentioned already cover Mathematical Methods, Conduction, Thermophysical Properties, Natural Convection, Forced Convection (Internal Flows), Rarefied Gases, Electric and Magnetic Fields, Condensation, Two-Phase Flow (each of the latter four topics being covered is less than 50 pages), and Radiation. Chapter authors are all distinguished researchers.

In a volume of this scope and purpose authors need to provide perspective, physical understanding, analytical representation, results for important cases or situations (including equations, graphs, or tables which can be used directly by the reader) and an entry to the literature. The authors have done this well. However, it is a handbook and not an encyclopedia: There are problems within the scope of the chapters for which solutions have been published but which are not represented here. The subject index of roughly 2000 line entries is thin; the most useful indexes often run to 2–2.5 entries per page of text, at which level this volume would have 3000 or more line entries. There is no author index. A reader-friendly feature is the almost completely uniform nomenclature throughout.

---

[1] Professor, Department of Mechanical Engineering, University of Colorado, Boulder, CO 80309. Fellow ASME.

[2] Professor of Engineering, Brown University, Providence, RI 02912. Fellow ASME.

The chapter on thermophysical properties, curtailed from that in the first edition, should be different by giving uniformly the *latest* internationally-agreed property data and by giving information for modern materials: e.g., Table 65, reproduced from the 1950's, includes in its "selected nonmetals" none of the polymers in widespread use in the 1980's. One may perhaps allow use of outdated material properties in an elementary text where they will be used only in exercises: but in a handbook I would hope to find the currently most accurate values. The page of references on Fickian interdiffusion coefficients could have been saved by citing T. R. Marrero and E. A. Mason, *J. Phys. Chem. Ref. Data*, Vol. 1, 1972, pp. 3-118, which reviews them.

Into whose hands would I place this Handbook? Certainly students and colleagues who might want a direct introduction to specific topics; also research engineers, consultants and other technically adept engineers. For some problems it is enough; for others, it helps provide an entry to the literature, especially from the 1955-1980 era.

Continuum Theory of the Mechanics of Fibre-Reinforced Composites. Edited by A. J. M. Spencer. Springer-Verlag, New York, 1984. 284 pages.

REVIEWED BY Z. HASHIN[1]

This book contains lectures given by A. H. England, D. F. Parker, A. C. Pipkin, T. G. Rolgers and A. J. M. Spencer on various aspects of the title subject matter at the International Centre for Mechanical Sciences (CISM) in Udine, Italy in 1981. The subject may be appropriately defined as mechanics of idealized strongly anisotropic materials. A basic assumed characteristic of such materials is inextensibility, i.e., zero strain, in one or more directions (fiber directions) (although Pipkin considers a case where the tensile strain but not the compressive strain vanishes). Another often assumed characteristic is incompressibility.

The first two chapters by Spencer, and Rogers are concerned with inextensible and incompressible materials. Spencer discusses constitutive relations for elasticity and plasticity and Rogers is concerned with finite deformations and the intrinsic stress singularities and discontinuities which arise in these kinds of idealized materials. The third chapter by England is concerned with plane problems for inextensible and incompressible linearly elastic strongly anisotropic solids. Pipkin in the fourth chapter discusses stress channelling and boundary layers in plane linear elastic deformation on the basis of inextensibility but not incompressibility, demonstrating that the stress singularities and discontinuities encountered in the idealized material are limiting cases of high stress gradients in strongly anisotropic materials. The fifth chapter by Rogers is concerned with mechanics of helically wound fiber reinforced cylinders when the material is inextensible in fiber directions and incompressible. This is a subject of engineering significance and it appears that the incompressibility assumption may introduce significant inaccuracies for the elastic behavior of actual fiber composites while the inextensibility assumption would be acceptable only for high modulus graphite/epoxy composites. In Chapter 6 Pipkin discusses fracture mechanics for inextensible materials. I believe this to be of particular interest in view of the relative simplicity of the theory in comparison with usual fracture mechanics of anisotropic materials. The test of the theory is of course experimental verification for high modulus fiber composites.

In Chapter 7 Spencer discusses reinforcement of holes in plates by fiber reinforced disks and Chapter 8, written by Parker, is concerned with wave propagation in inextensible and incompressible materials. Spencer in chapter 9 discusses dynamics of rigid-plastic beams and plates. He presents a number of simple solutions to important problems and it would seen that the simplified theory for idealized materials should be of particular value for this subject matter. Finally Pipkin in Chapter 10 generalizes Rivlin's theory of inextensible networks to the case when the fibers are inextensible in tension but not in compression thus taking into account in simple and elegant fashion the microbuckling of stiff fibers within a soft matrix.

The editor and the authors are to be commended for having succeeded in joining the various chapters organically together and they have thus been able to present a coherent and logical account of this interesting and important area of the mechanics of solids.

Finally, a semantic comment: The subject presented is mechanics of idealized strongly anisotropic materials and its description as mechanics of fiber reinforced materials is not always appropriate. The assumption of inextensibility would be quite inaccurate for glass/epoxy and metal matrix fiber composites and should probably be restricted to high modulus graphite/epoxy and fiber reinforced rubber. The assumption of incompressibility would not be adequate for linear elastic behavior and is much more appropriate for plastic strains and for fiber reinforced rubber.

---

[1] Professor, Department of Solid Mechanics, Materials and Structures, Tel-Aviv University, Tel-Aviv, Israel. Fellow ASME.