Transactions of the ASME®

Technical Editor, L. B. FREUND

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

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.) for $115 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, $27.00, annually; to nonmembers, $115.00. Add $12.00 for postage to countries outside the United States and Canada.

INDEXED by Applied Mechanics Reviews and Engineering Information, Inc.

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 54 • NUMBER 4 • DECEMBER 1987

(Contents continued on Inside Back Cover)

CONTENTS (Continued)

DESIGN DATA AND METHODS

BRIEF NOTES

BOOK REVIEWS

Leon M. Keer Appointed New Technical Editor of the Journal of Applied Mechanics

Professor L. M. Keer has been appointed Technical Editor of the Journal of Applied Mechanics for a five-year term beginning January 1, 1988. Professor Keer succeeds Professor L. B. Freund, who completes his five-year term of service on December 31.

Ben Freund is a tough act to follow. His performance in running Journal operations, finding top candidates as Associate Editors, and interfacing with ASME's Publications Committee has been outstanding. Ben faced the same challenge five years ago when he succeeded C. S. Hsu, who served as Technical Editor during 1976–1982. The results of such leadership are evident in the current health and vitality of the Journal.

Leon's appointment is the culmination of a year-long search process. In accordance with the rules of the Applied Mechanics Division, the AMD Executive Committee established in the summer of 1986 a Search Committee to recommend candidates. The members were Jan Achenbach (Chairman), Richard Benson, Ben Freund, C.S. Hsu, Sid Leibovich, and Jim Rice. The committee forwarded its recommendations to the Executive Committee in June, and Leon was nominated and subsequently approved by the ASME Board on Communications. I should mention that the appointment requires a commitment from the Technical Editor's home institution; Northwestern University's willingness to provide the necessary office space is much appreciated. I should also mention that, having served on the previous search committee in 1981–1982, I can attest to the time-consuming nature of this work; we in the AMD are indebted to the search committee members for their fine service.

Leon M. Keer is Professor of Civil Engineering and Mechanical and Nuclear Engineering, as well as Associate Dean for Research and Graduate Study at Northwestern University. He holds B.S. and M.S. degrees from the California Institute of Technology and a Ph.D. from the University of Minnesota; also, he is a Registered Professional Engineer in the State of California. Leon has made important research contributions in the areas of stress analysis, fracture mechanics, geotechnology, mechanical contact, and tribology. He is a Fellow of the ASME and the American Academy of Mechanics.

Leon M. Keer

With this appointment, Leon joins a distinguished group of key figures in Applied Mechanics. The Journal of Applied Mechanics is a central activity of the AMD, and enjoys a prominent position in the mechanics literature. We are grateful to Ben Freund for his leadership in maintaining that position and extend to Leon our warm support as he begins his appointment as Technical Editor.

Thomas L. Geers
for the
AMD Executive Committee

J. D. Achenbach
Fellow ASME.

D. A. Sotiropoulos

H. Zhu

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

Characterization of Cracks From Ultrasonic Scattering Data

An inverse method for ultrasonic scattering data is proposed, to characterize a planar crack of general shape contained in an elastic solid. The method is based on an integral representation for the scattered field in the frequency domain. For a given scattered field the inverse problem has been formulated as a nonlinear optimization problem. At low frequencies its solution gives the location of the crack, the orientation of the crack-plane, and the crack-opening volumes. In addition the Mode I stress-intensity factor is obtained for a related static stress state corresponding to service loads.

Introduction

A crack in the interior of a solid body can often be detected nondestructively by observation of its effect on an externally applied ultrasonic field. The influence of a crack on the overall strength of a component generally depends, however, on its location, size, shape, and orientation. Hence, a useful nondestructive test should go beyond crack detection to crack characterization, with the ultimate aim of an efficient assessment of the crack's influence on the integrity of the body under service loads.

In this paper we propose an inverse method for ultrasonic crack-scattering data to characterize a crack of general shape in a homogeneous isotropic linearly elastic solid. The method is based on the well-known integral representation for the scattered field. Even though this representation is simplified for the far field and for large wavelengths as compared to a cross-sectional dimension of the crack, it retains validity for points closer to the crack than similar formulas presented in previous studies. For a given scattered field the inverse problem is subsequently formulated as a nonlinear optimization problem. Its solution gives the location of the crack, the crack's orientation, and the components of the crack-opening volume tensor induced by the probing ultrasonic field. In addition, for a related problem of quasi-static loading, the Mode I stress intensity factor and the strain energy induced by the presence of the crack have been computed directly from the results of the inverse problem.

In recent years numerous results have become available for fields generated by scattering of ultrasonic waves by cracks. Solutions for two-dimensional configurations in unbounded bodies have been discussed by Achenbach et al. (1982), who also listed earlier references, e.g., Mal (1970). The three-dimensional case of scattering of a plane wave by a penny-shaped crack has been investigated by Krenk and Schmidt (1982) and Martin (1982). The more difficult configurations of a surface-breaking and a subsurface crack has been considered by Mendelsohn et al. (1980) and Achenbach and Brind (1981), for the in-plane problem. Three-dimensional scattering by a surface-breaking crack has been analyzed by Angel and Achenbach (1984). For the surface breaking crack, comparisons between analytical results and experimental data have been presented by Yew et al. (1984), Dong and Adler (1984), and Vu and Kinra (1985). Experimental results for reflection of a surface wave by a subsurface crack oriented normal to the free surface were obtained by Khuri-Yakub et al. (1984), who observed very satisfactory agreement with theoretical results. There are, of course, also numerical studies, principally by the T-matrix method. Typical of these is the work of Visscher (1985).

Recent review papers of ultrasonic QNDE which include substantial discussions of scattering of ultrasonic waves by cracks are those of Fu (1982) and Thompson (1983). Interesting practical applications have been discussed by Coffey and Chapman (1983).

Whereas a substantial body of literature has become available for the direct problem of scattering by cracks, relatively little has been published on the inverse problem. A method based on inverse time-domain ray tracing, which has yielded some interesting results, has been discussed by Norris and Achenbach (1982). Papers by Teitel (1978) and Gubernatis and Domany (1979) have discussed the determination of orientation and size for cracks which are elliptical in shape, and whose location is known a priori. Their method is based on the quasi-static crack-opening displacements given by Eshelby (1957). The method of this paper, which is not based on Eshelby's result, offers two advantages over previous solutions. One advantage is that the crack may be of general shape (i.e., it is not assumed elliptical a priori). The other is that the solution is valid for frequencies higher than the "quasi-static" ones. As mentioned before, another new feature of this study is that the location of the crack does not have to be known a priori.

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 10, 1986; final revision June 24, 1987.

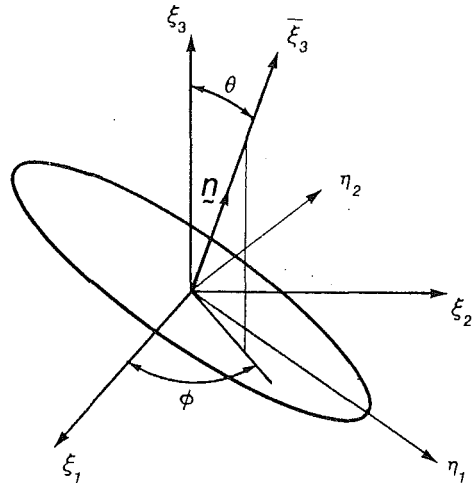


Fig. 1 Crack configuration in the x and ξ coordinate systems. The x system (not indicated) coincides with the ξ system.

Integral Representation for the Scattered Field

The faces of a perfect mathematical crack are smooth and infinitesimally close, but they do not interact with each other. From the analytical point of view a perfect mathematical crack is a surface in space which does not transmit tractions. The model is acceptable for a real crack, provided that the latter's faces are slightly separated and that the length characterizing crack-face roughness is much smaller than the dominant wavelengths of an incident pulse of ulstrasonic wave motion.

Figure 1 shows a flat crack of surface area A. Two coinciding coordinate systems (x_1, x_2, x_3) and (ξ_1, ξ_2, ξ_3) with origin at the "scattering center" (defined later) of A are defined. The unit normal, $\mathbf{n}$, to the crack-plane makes an angle θ with the x_3 axis, and the projection of $\mathbf{n}$ on the x_1x_2 plane makes an angle ϕ with the x_1 axis. The incidence of a pulse of elastic wave motion produces a scattered field. A representation for the scattered field can conveniently be obtained in the frequency domain. The incident pulse may be expressed as a Fourier superposition of harmonics, next the scattered field generated by a single incident harmonic can be calculated, and finally all scattered harmonic fields can be superimposed to yield the scattered pulse. Hence, the generic problem is one of scattering of an incident displacement wave of the general form

$$\mathbf{u}^{in}(\mathbf{x})\exp(-i\omega t), \tag{1}$$

where $\mathbf{u}^{in}(\mathbf{x})$ is understood to depend on frequency. The term $\exp(-i\omega t)$ will be omitted in the sequel. The scattered field is defined as the difference between the total field and the incident field, i.e.,

$$u_i(\mathbf{x}) = u_i^{in}(\mathbf{x}) + u_i^{sc}(\mathbf{x}) \tag{2}$$

The components of the corresponding stress tensors are related by

$$\tau_{ij}(\mathbf{x}) = \tau_{ij}^{in}(\mathbf{x}) + \tau_{ij}^{sc}(\mathbf{x}) \tag{3}$$

It is well known that the integral representation for the scattered displacement field may be expressed in the form

$$u_k^{sc}(\mathbf{x}) = \int_A C_{ijlm} D_{klm}(\mathbf{x}-\xi)\Delta u_i(\xi) n_j(\xi) dA(\xi) \tag{4}$$

where

$$D_{klm}(\mathbf{x}-\xi) = \frac{\partial}{\partial_{xm}} u^G_{k;l}(\mathbf{x}-\xi) \tag{5}$$

In these relations $u^G_{k;l}(\mathbf{x}-\xi)$ is the basic singular solution (Green's vector), i.e., the displacement at position $\mathbf{x}$ in direction x_k, due to a unit time-harmonic force applied in direction x_l at $\mathbf{x}=\xi$. Then $-D_{klm}(\mathbf{x}-\xi)$ is the displacement produced at $\mathbf{x}$ in the x_k direction by a double force applied at $\mathbf{x}=\xi$, with forces in the l direction and moment arm in the m direction. The displacement discontinuities $\Delta u_i(\xi) = \Delta u_i^{sc}(\xi)$ are related to the crack opening displacements $\Delta\bar{u}_i(\bar{\xi})$ by

$$\Delta u_i(\xi) = T_{ij}\Delta\bar{u}_j(\bar{\xi}) \tag{6}$$

where the matrix $[T]$ defines the rotation of the ξ system relative to the $\bar{\xi}$ system

$$[T] = \begin{bmatrix} -\cos\theta\cos\theta & \sin\theta & \cos\theta\sin\phi \\ -\sin\theta\cos\phi & -\cos\theta & \sin\theta\sin\phi \\ \cos\phi & 0 & \cos\phi \end{bmatrix} \tag{7}$$

and the $\bar{\xi}$ coordinate system is such that $\bar{\xi}_3$ is normal to the crack plane. The crack opening displacement $\Delta\bar{\mathbf{u}}(\bar{\xi})$ is defined in the usual manner as

$$\Delta\bar{\mathbf{u}}(\bar{\xi}) = \bar{\mathbf{u}}(\bar{\xi}+\epsilon\mathbf{n}) - \bar{\mathbf{u}}(\bar{\xi}-\epsilon\mathbf{n}) \tag{8}$$

where $\mathbf{n}$ is the outward normal to the crack plane as shown in Fig. 1.

For a homogeneous isotropic material C_{ijlm} has the form

$$C_{ijlm} = \lambda\delta_{ij}\delta_{lm} + \mu(\delta_{il}\delta_{jm} + \delta_{im}\delta_{jl}) \tag{9}$$

In the x system, the basic singular solution for an isotropic solid is

$$\mu u^G_{k;l}(\mathbf{x}-\xi) = \frac{1}{k_T^2}\frac{\partial^2}{\partial x_k \partial x_l}[-G(k_L R) + G(k_T R)] + G(k_T R)\delta_{kl} \tag{10}$$

where

$$G(k_\alpha R) = \frac{\exp(ik_\alpha R)}{4\pi R}, \quad \alpha = L, T \tag{11}$$

$$k_L = \frac{\omega}{c_L}, \quad c_L^2 = (\lambda+2\mu)/\rho, \tag{12}$$

$$k_T = \frac{\omega}{c_T}, \quad c_T^2 = \mu/\rho, \tag{13}$$

and

$$R = |\mathbf{x}-\xi| \tag{14}$$

The fields for the corresponding double forces are

$$\mu D_{klm}(\mathbf{x}-\xi) = \mu\frac{\partial}{\partial x_m} u^G_{k;l}(\mathbf{x}-\xi) = -\frac{c_T^2}{c_L^2} k_L G(k_L R) A^L_{klm} + k_T G(k_T R)\left[A^T_{klm} + \gamma_m\left(i - \frac{1}{k_T R}\right)\delta_{kl}\right], \tag{15}$$

where

$$A^\alpha_{klm} = -\gamma_k\gamma_l\gamma_m\left(i - \frac{1}{k_\alpha R}\right) - \frac{1}{k_\alpha R}\left[1 + \frac{3}{k_\alpha R}\left(i - \frac{1}{k_\alpha R}\right)\right]\Gamma_{klm} \tag{16}$$

$$\Gamma_{klm} = \gamma_k\delta_{lm} + \gamma_l\delta_{km} + \gamma_m\delta_{kl} - 5\gamma_k\gamma_l\gamma_m \tag{17}$$

$$\gamma_i = (x_i - \xi_i)/R \tag{18}$$

It is instructive to write the details of the scattered field in the $\bar{x}$ system, for which the $\bar{x}_3$ axis is along the normal to the crack plane. Then $j=3$ and equation (4) can be split into contributions from three basic sets of double forces, which are related to crack-opening displacements in the $\bar{x}_1$, $\bar{x}_2$, and $\bar{x}_3$ directions, respectively:

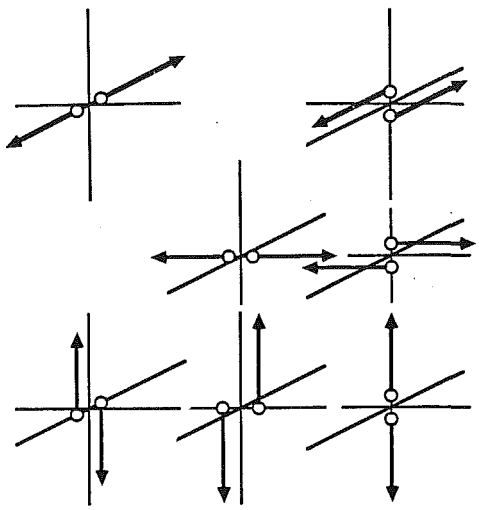


Fig. 2 Double forces for scattered fields

$$\bar{u}_k^{sc}(\bar{\mathbf{x}}) = \sum_{i=1}^{3} U_k^{(i)}(\bar{\mathbf{x}}) \tag{19}$$

where

$$U_k^{(1)}(\bar{\mathbf{x}}) = \mu U_k^{(1,13)}(\bar{\mathbf{x}}) + \mu U_k^{(1,31)}(\bar{\mathbf{x}}) \tag{20}$$

$$U_k^{(2)}(\bar{\mathbf{x}}) = \mu U_k^{(2,32)}(\bar{\mathbf{x}}) + \mu U_k^{(2,23)}(\bar{\mathbf{x}}) \tag{21}$$

$$U_k^{(3)}(\bar{\mathbf{x}}) = \lambda U_k^{(3,11)}(\bar{\mathbf{x}}) + \lambda U_k^{(3,22)}(\bar{\mathbf{x}}) + (\lambda + 2\mu) U_k^{(3,33)}(\bar{\mathbf{x}}) \tag{22}$$

In equations (19)–(22)

$$U_k^{(i,\ell m)}(\bar{\mathbf{x}}) = \frac{\partial}{\partial \bar{x}_m} \int_A \bar{u}_{k;\ell}^G(\bar{\mathbf{x}};\bar{\xi}) \Delta \bar{u}_i(\bar{\xi}) d\bar{\xi}_1 d\bar{\xi}_2 \tag{23}$$

Even though the results of this section have cast scattering theory in an elegant form, the actual direct scattering problem is far from solved, since the crack-opening displacements are not known a priori. To obtain a set of equations for $\Delta\bar{u}_i$, we must return to the representation given by equation (19). The stresses corresponding to equation (19) can be obtained by substitution into Hooke's law. Then by letting $\bar{\mathbf{x}}$ approach A, the stresses should satisfy $\bar{\tau}_{i3}^{sc}(\bar{\mathbf{x}}) = -\bar{\tau}_{i3}^{in}(\bar{\mathbf{x}})$. The resulting equations are integral equations for $\Delta\bar{u}_i(\xi)$. These equations require careful handling, because the integrands contain singularities. The general system of equations has been presented by Budiansky and Rice (1979).

The complexity of the exact method of solution has stimulated the development of approximate approaches. In two well-known approximate theories, which are valid at low and high frequencies, respectively, a form of $\Delta\bar{\mathbf{u}}(\bar{\xi})$ is postulated and $\bar{\mathbf{u}}^{sc}(\bar{\mathbf{x}})$ is subsequently computed by equation (19). At low frequencies the static crack-opening displacement can be substituted in the integral, to give the so-called quasi-static scattering theory. At very high frequencies the geometrical elastodynamic field on the illuminated crack-face can be used as the crack-opening displacement. Subsequent evaluation of equation (19) produces the "physical elastodynamics" or Kirchhoff approximation to the scattered field.

In many practical applications the distance from the observation point to the center of the crack is much larger than a characteristic dimension, a, of the crack. At point $\mathbf{x}$ outside the crack we may write

$$R = |\mathbf{x} - \xi| = r\left[1 - \frac{\xi \cdot \hat{\mathbf{x}}}{r} + \frac{1}{2}\frac{|\xi|^2}{r^2} - \frac{(\xi \cdot \hat{\mathbf{x}})^2}{2r^2} + \ldots\right] \tag{24}$$

where

$$r = (x_i x_i)^{1/2}, \quad \text{and} \quad \hat{x}_i = x_i / r \tag{25,26}$$

Since $r \gg a$, and $|\xi| = (\xi_1^2 + \xi_2^2 + \xi_3^2)^{1/2} \leq a$, equation (24) becomes

$$R = r \tag{27}$$

Furthermore, the assumption

$$(k_\alpha a)\frac{a}{r} \ll 1, \quad \alpha = L, T, \tag{28}$$

together with equation (24) gives

$$e^{ik_\alpha R} = e^{ik_\alpha r} e^{-ik_\alpha \xi \cdot \hat{\mathbf{x}}} \tag{29}$$

Equations (27) and (29), substituted in equations (10) and (4), will give rise to integrals of the general form

$$I_\alpha = \int_A e^{-ik_\alpha \xi \cdot \hat{\mathbf{x}}} \Delta u_i(\xi) n_j(\xi) dA(\xi) \tag{30}$$

In deriving equation (30) the assumptions $r \gg a$ and (28) were made. These assumptions include the case

$$k_\alpha a < 1 \quad \text{and} \quad r \gg a \tag{31,32}$$

which is under consideration in this study.

To determine the solution of the inverse problem, it is convenient to further reduce the integrals of equation (30). Use of the first mean value theorem in equation (30) gives

$$I_\alpha = e^{-ik_\alpha \xi^* \cdot \hat{\mathbf{x}}} V_{ij} \tag{33}$$

where

$$V_{ij} = \int_A \Delta u_i(\xi) n_j(\xi) dA(\xi) \tag{34}$$

In equation (33), ξ^* is a point on the crack plane. We call this point the scattering center. If $k_\alpha a \ll 1$ (quasistatic case), then $I_\alpha = V_{ij}$, provided that the origin of the coordinate systems $(\mathbf{x}, \xi)$ is any point on the crack plane. In this case, the scattered field given by equation (35) below, would be "independent" of the choice of the origin. In our study, we do not make the strict quasistatic assumption. The origin is chosen to coincide with the scattering center ξ^*. It should be mentioned that different observation points $\mathbf{x}$ would define different ξ^*. However, if the observation points are not very far from each other, it is expected that ξ^* would be approximately the same, and

$$I_\alpha \approx V_{ij} \tag{35}$$

Use of equation (35), reduces the scattered field given by equation (4) to

$$u_k^{sc}(\mathbf{x}) = C_{ij\ell m} D_{k\ell m}(\mathbf{x}) V_{ij} \tag{36}$$

where $D_{k\ell m}(\mathbf{x})$ follows from equation (15).

In the $\bar{x}$ coordinates $\bar{V}_{ij} = \bar{V}_{i3}\delta_{3j}$ are the crack-opening volumes corresponding to the crack-opening displacements defined by equation (8)

$$\bar{V}_{i3} = \int_A \Delta \bar{u}_i(\bar{\xi}) dA(\bar{\xi}) \tag{37}$$

The scattered field is now given by equation (19), where instead of equation (23) we have

$$U_k^{(i,\ell m)}(\bar{\mathbf{x}}) = \bar{D}_{k\ell m} \bar{V}_{i3} \tag{38}$$

The scattered field consists of fields generated by sets of double forces located at the scattering center of the crack, and whose configuration is summarized in Fig. 2.

A general theory for body-force equivalents of displacement discontinuities was first developed by Burridge and Knopoff (1964). It is not surprising that the scattering problem has been reduced to the same kind of source theory, which is well known from seismology and the theory of acoustic emission.

Inverse Problem

In terms of the known coordinate system, $x_1x_2x_3$, whose orientation is known but not the location of its origin, the scattered displacement field, equation (36), may be written in vector-matix form as

$$\{u^{sc}(\mathbf{x})\}_{6\times 1} = [L(\mathbf{x})]_{6\times 12}\{B\}_{12\times 1} \tag{39}$$

where the scattered displacement vector $\{u^{sc}\}$ has 6 components considering its real and imaginary parts. Similarly $\{B\}$ has 12 components. The ith equation of equation (39) in complex form is:

$$u_i^{sc}(\mathbf{x}) = [L_{i11}L_{i12}L_{i13}L_{i22}L_{i23}L_{i33}]\{B\} \tag{40}$$

where

$$\begin{aligned}
L_{i11} &= (\lambda+2\mu)D_{i11} + \lambda(D_{i22}+D_{i33})\\
L_{i12} &= 2\mu(D_{i12}+D_{i21})\\
L_{i13} &= 2\mu(D_{i13}+D_{i31})\\
L_{i22} &= (\lambda+2\mu)D_{i22} + \lambda(D_{i11}+D_{i33})\\
L_{i23} &= 2\mu(D_{i23}+D_{i32})\\
L_{i33} &= (\lambda+2\mu)D_{i33} + \lambda(D_{i11}+D_{i22})
\end{aligned} \tag{41}$$

and

$$\{B\}^T = [V_{11}V_{12}V_{13}V_{22}V_{23}V_{33}]^T \tag{42}$$

The terms D_{ilm} are given by equation (15). Let us suppose that the left-hand side of equation (39) is known. The equation then defines a set of 6 nonlinear equations for 15 unknowns (3 components of $\mathbf{x}$, and 12 components of $\mathbf{B}$). This is an underdetermined system of nonlinear equations. To obtain a unique solution the system has to be overdetermined. Thus, the data u_i^{sc} is needed in 3 observation points ($\mathbf{x}^1$, $\mathbf{x}^2$, $\mathbf{x}^3$), since the components of vector $\mathbf{B}$ are independent of $\mathbf{x}$. This way, there are 18 equations and 15 unknowns. To reduce the number of unknowns we first eliminate the unknown $\mathbf{B}$ and subsequently we solve for $\mathbf{x}$. However, an inspection of equation (39) reveals that the left inverse of $[L(\mathbf{x})]$ does not exist (more rows than columns). Thus to solve for $\mathbf{B}$ one needs to consider the data in 2 observation points simultaneously. This gives

$$\begin{bmatrix} u^{sc}(\mathbf{x}^1) \\ u^{sc}(\mathbf{x}^2) \end{bmatrix}_{12\times 1} = P(\mathbf{x}^1)_{12\times 12}\{B\}_{12\times 1}, \tag{43}$$

where it has been taken into account that $\mathbf{x}^1 = \mathbf{x}^2 + \mathbf{X}_{12}$, where $\mathbf{X}_{12}$ is known. Equation (43) yields

$$\{B\} = [P(\mathbf{x}^1)]^{-1}\begin{bmatrix} u^{sc}(\mathbf{x}^1) \\ u^{sc}(\mathbf{x}^2) \end{bmatrix} \tag{44}$$

Substitution of equation (44) into equation (39) for the 3rd observation point, $\mathbf{x}^3$, gives

$$\{u^{sc}(\mathbf{x}^3)\} = [L(\mathbf{x}^3)][P(\mathbf{x}^1)]^{-1}\begin{bmatrix} u^{sc}(\mathbf{x}^1) \\ u^{sc}(\mathbf{x}^2) \end{bmatrix} \tag{45}$$

The solution of the inverse problem has, therefore, been reduced to the solution of equation (45). This is a system of 6 nonlinear equations for 3 unknowns (x_1^1, x_2^1, x_3^1), since the relationship between $\mathbf{x}^1$, $\mathbf{x}^2$. $\mathbf{x}^3$, is known from the relative position of the instrument that measures the scattered data. To solve the nonlinear optimization problem defined by equation (45), the following six residuals, $g_i(\mathbf{x})$, are defined:

$$\{g(\mathbf{x}^1)\} = \{u^{sc}(\mathbf{x}^3)\} - [L(\mathbf{x}^3)][P(\mathbf{x}^1)]^{-1}\begin{bmatrix} u^{sc}(\mathbf{x}^1) \\ u^{sc}(\mathbf{x}^2) \end{bmatrix} \tag{46}$$

The residuals, $g_i(\mathbf{x}^1)$, are now minimized in the least squares sense with respect to $\mathbf{x}^1$, i.e., we seek

$$\min_{\mathbf{x}^1}\left\{\sum_{i=1}^{6} g_i^2(\mathbf{x}^1)\right\} \tag{47}$$

Equations (46) and (47) define a nonlinear least squares problem. Its solution is obtained through use of a modification of the Levenberg-Marquardt algorithm as outlined in the User-Guide for MinPack-1 (1980). The solution so obtained is the unknown $\mathbf{x}^1 = (x_1^1, x_2^1, x_3^1)$. Substitution of $\mathbf{x}^1$ in equation (44) gives the crack-opening volume tensor, V_{ij}, in the global $\mathbf{x}$ system.

Once V_{ij} has been obtained, the crack opening volumes in the $\bar{x}$ system may be obtained by the rule of transformation of a second order tensor

$$[\bar{V}] = [T]^{-1}[V][T], \tag{48}$$

where both $\bar{V}_{ij}$ and V_{ij} are symmetric,

$$[\bar{V}] = \begin{bmatrix} 0 & 0 & \frac{1}{2}\bar{V}_{13} \\ 0 & 0 & \frac{1}{2}\bar{V}_{23} \\ \frac{1}{2}\bar{V}_{31} & \frac{1}{2}\bar{V}_{32} & \bar{V}_{33} \end{bmatrix} \tag{49}$$

and $[T]$ is defined by equation (7). From the invariants of the two tensors we obtain

$$\bar{V}_{33} = V_{11} + V_{22} + V_{33} \tag{50}$$

as well as the relation

$$\begin{aligned}
\frac{1}{4}(\bar{V}_{13}^2 + \bar{V}_{23}^2) &= -(V_{11}V_{22} + V_{22}V_{33} + V_{11}V_{33})\\
&\quad + (V_{12}^2 + V_{23}^2 + V_{13}^2)
\end{aligned} \tag{51}$$

From the zero terms of $\bar{V}_{ij}$, equation (48) yields three nonlinear equations for the unknown angles θ and ϕ. The solution to this overdetermined system of equations is obtained in the same manner as the solution to equation (45). Subsequently $\bar{V}_{ij}$ follow from the remaining equations of (48). Equations (50) and (51) can be used as a check on the results.

The results of the inverse problem, namely, the crack's location, the orientation of the crack plane, and the crack-opening volumes, $\bar{V}_{ij}$, do not completely characterize the crack. In the next Section we will show that these results do, however, make it possible to estimate the maximum value of a stress-intensity factor under a service loading condition without further assumptions on the crack shape. As shown in the remainder of this Section, it is also possible to fully characterize the crack, provided that it may be assumed that the crack is elliptical in shape.

Suppose that the axes of the ellipse are defined by a system of coordinates η_1, η_2 where the major axis, η_1, makes an angle χ with the $\bar{\xi}_1$, axis, and where the ellipse would be defined by $(\eta_1/a)^2 + (\eta_2/b)^2 = 1$, $a > b$. As shown in this Section it is then possible to obtain a, b, and χ from the computed low frequency values of the crack-governing volumes $\bar{V}_{ij}$. If the crack would, in fact, not be elliptic in shape, then the computation will produce a, b, and χ for an equivalent crack of elliptical shape.

Let us define as $\bar{\tau}_{i3}$ the stress components on the crack plane at the centroid of the crack in the $\bar{x}$ coordinates. The components $\bar{\tau}_{i3}$ can be computed from the incident wave and the results of the inverse problem for crack location and orientation of the crack plane. Next we define $\bar{\tau}_{i3}^{\chi}$, where

$$\bar{\tau}^{\chi}_{i3} = \beta_{ij}\bar{\tau}_{j3} \tag{52}$$

$$\beta = \begin{bmatrix} \cos\chi & \sin\chi & 0 \\ -\sin\chi & \cos\chi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{53}$$

At low frequencies the tractions may be taken as uniform (and equal to τ^{χ}_{i3}) over the crack plane. The corresponding crack opening displacements then follow from Eshelby's results (1957) as

$$[\Delta\bar{u}^{\chi}_i] = C^{(i)}\bar{\tau}^{\chi}_{i3}\left(1 - \frac{\eta_1^2}{a^2} - \frac{\eta_2^2}{b^2}\right)^{1/2}, \tag{54}$$

where

$$C^{(3)} = \frac{2b}{E(\alpha)}\frac{1-\nu}{\mu}, \quad C^{(2)} = \frac{2a}{\mu\Gamma_2}, \quad C^{(1)} = \frac{2b}{\mu\Gamma_1} \tag{55a,b,c}$$

In these expressions

$$\Gamma_1 = E(\alpha) + \frac{\nu}{1-\nu}\frac{\beta^2}{\alpha^2}[K(\alpha) - E(\alpha)] \tag{56}$$

$$\Gamma_2 = \frac{E(\alpha)}{\beta} + \frac{\nu}{1-\nu}\frac{1}{\beta}\frac{1}{\alpha^2}[E(\alpha) - \beta^2 K(\alpha)] \tag{57}$$

$$\alpha = (1 - b^2/a^2)^{1/2}, \quad \beta = b/a \tag{58a,b}$$

Here ν is Poisson's ratio, and μ is the shear modulus, while $E(\alpha)$ and $K(\alpha)$ are complete elliptic integrals of the second and first kinds, respectively, with modulus α.

Integration of equation (54) over A yields the crack-opening volumes

$$\bar{V}^{\chi}_{i3} = \int_{A_E}[\Delta\bar{u}^{\chi}_i]dA = C^{(i)}\bar{\tau}^{\chi}_{i3}\frac{2}{3}\pi ab \tag{59}$$

These crack opening volumes are related to the crack opening volumes computed from the inverse problem by

$$\bar{V}^{\chi}_{i3} = \beta_{ij}\bar{V}_{j3} \tag{60}$$

where β_{ij} is defined by equation (53). Since $\bar{V}_{i3}$ are known, equation (60) gives 3 nonlinear equations for the 3 unknowns, a, b, and χ. The complicated forms of $C^{(i)}$ as given by equations (55a, b, c) may make these equations difficult to solve.

Numerical Tests

Several numerical tests were carried out to check the validity of the solution of the inverse problem. The results of one of the tests are given here. The crack considered is elliptical, with semi-axes defined by $b/a = 0.8$. The orientation of the crack is specified by $\theta = 1$, $\phi = 1$ (see Fig. 1). The crack is contained in an infinite elastic isotropic solid characterized by Poisson's ratio $\nu = 0.3$ and $c_L = 6{,}000$ m/s. The incident displacement field is a longitudinally polarized plane wave incident along the $-x_3$ direction, with amplitude u^o and wavenumber k_L.

To obtain the position of the crack and the crack-opening volumes $\bar{V}_{i3}$, the scattered displacement field is needed at 3 observation points. This data is synthesized by first solving the direct problem as defined by equation (36). For small $k_L a$, the crack-opening volumes are assumed to be given by equation (59). The real parts are zero. In general, this would not be true. Subsequent use of equation (36) gives the scattered displacement field. The three observation points chosen arbitrarily are $\mathbf{x}^1/a = (6, 7, 10)$, $\mathbf{x}^2/a = (10, 8.5, 10)$, $\mathbf{x}^3/a = (14, 10, 10)$. The scattered displacement was calculated at these 3 points, for $k_L a = 0.325, 0.35, 0.375, 0.4$. The scattered particle velocity components are listed in Table I in the global system, for $k_L a = 0.4$. Here $(\dot{\mathbf{u}}^{sc})_N = \dot{\mathbf{u}}^{sc}(\lambda + 2\mu)/|\sigma^{in}_{33}|$. Note that $\dot{\mathbf{u}}^{sc} = -ik_L c_L \mathbf{u}^{sc}$ and $|\sigma^{in}_{33}| = (\lambda + 2\mu)\, k_L u_o$.

The synthesized scattered data was used to solve the inverse problem. The position of one of the observation points was obtained as well as the crack-opening volumes, and the crack orientation. In Table 2, the actual parameters used in the solution of the direct problem (Actual) are compared with the solution of the inverse problem (Inverse 1) for all $k_L a$ considered. The match is excellent. The real parts are not shown even though the match is also excellent. Numerical tests were

Table 1 Real and imaginary parts of normalized velocity components for the scattered field at three positions: $\vec{x}^1/a$ = (6, 7, 10), $\vec{x}^2/a$ = (10, 8.5, 10), and $\vec{x}^3/a$ = (14, 10, 10). All components are in m/s.

$k_L a = 0.4$	$\vec{x}^1/a$	$\vec{x}^2/a$	$\vec{x}^3/a$
$(\dot{u}^R_1)_N$	3.60 m/s	−0.06 m/s	2.76 m/s
$(\dot{u}^I_1)_N$	−3.42	4.68	−7.80
$(\dot{u}^R_2)_N$	−4.08	−0.18	9.18
$(\dot{u}^I_2)_N$	0.60	−9.36	3.90
$(\dot{u}^R_3)_N$	0.42	−0.24	12.90
$(\dot{u}^I_3)_N$	−2.58	−13.56	8.76

Table 2 Comparison of actual parameters with results of the inverse problem, for use of data based on the actual parameters (Inverse 1) and use of modified data (Inverse 2), for four values of $k_L a$

		$\frac{x^1_1}{a}$	$\frac{x^1_2}{a}$	$\frac{x^1_3}{a}$	θ	ϕ	$\frac{i\bar{V}_{13}}{u^o A}$	$\frac{i\bar{V}_{23}}{u^o A}$	$\frac{i\bar{V}_{33}}{u^o A}$
	Actual	6	7	10	1.	1.	0.18	0	0.36
$k_L a = 0.325$	Inverse 1	6	7	10	1.	1.	0.18	0	0.36
	Inverse 2	6.48	7.35	10.80	0.83	0.98	0.17	−0.06	0.37
	Actual	6	7	10	1.	1.	0.20	0	0.39
$k_L a = 0.35$	Inverse 1	6	7	10	1.	1.	0.20	0	0.39
	Inverse 2	6.63	6.93	10.84	0.85	1.00	0.19	−0.07	0.39
	Actual	6	7	10	1.	1.	0.22	0	0.42
$k_L a = 0.375$	Inverse 1	6	7	10	1.	1.	0.22	0	0.42
	Inverse 2	6.61	7.69	10.38	0.87	0.98	0.20	−0.08	0.42
	Actual	6	7	10	1.	1.	0.24	0	0.45
$k_L a = 0.4$	Inverse 1	6	7	10	1.	1.	0.24	0	0.45
	Inverse 2	6.2	7.04	10.41	0.87	0.99	0.21	−0.08	0.44

also performed on modified data to test the stability of the solution. In Table 2, the results of one of the tests are shown (Inverse 2). For the modified data, the x_1 component of the scattered displacement field at $\mathbf{x}^1$ was taken as 0.92 times the exact displacement. The x_2 component of $\mathbf{u}^s$ at x^2 was 0.98 times the exact displacement, and the x_3 component of $\mathbf{u}^s$ at x^3 was 1.06 times the exact displacement. For all the other components of the scattered displacement the exact synthesized scattered data were used. The test displacement function used, in this particular example, has a frequency spectrum for each component that slightly deviates from that of the exact displacement. The results are satisfactory and it is concluded that errors in the scattered data cause errors of the same order in the crack location and orientation, and crack-opening volumes.

The success of the proposed inverse method in practical applications depends on the availability of suitable low frequency scattering data. Appropriate signal processing of experimental time-domain measurements will give frequency domain data. In the low-frequency range it may, however, be necessary to fit the displacement data on a curve of the general form constant $\times k_L^2$, to improve the accuracy. It is noted from Table 1 that an incident plane wave produced by a regular transducer will generate a very small scattered field. This is primarily due to the geometrical attenuation caused by the term a/r. The displacements tend to be particularly small in the plane perpendicular to the ray connecting the point of observation and the centroid of the inhomogeneity. The scattered displacement magnitude can, however, be amplified considerably by the use of a focused transducer. But even then only the radial displacement component may be useful. In that case the number of observation points would have to be tripled to make up for the lack of transverse displacement data. This would have to be done in any event if a water-bath configuration would be used.

Strength Considerations

The ultimate goal of quantitative nondestructive evaluation is to obtain information on the residual strength of materials and components. Information on the location, size, shape and orientation of cracks makes it possible to calculate critical strength parameters such as stress intensity factors. As pointed out by Budiansky and Rice (1978), it is, however, also feasible to compute stress intensity factors directly from low-frequency ultrasonic scattering data.

In this Section we show that the zero-frequency limits of the crack-opening volumes computed from the ultrasonic data, yield direct estimates of the stress intensity factors for a related service loading condition. For this purpose the incident ultrasonic wave is redefined as

$$u_3^I(\mathbf{x}) = \frac{u^o}{ik_L a} \exp(ik_L x_3) \tag{61}$$

in order that the zero-frequency limit will correspond to a nonzero stress of the form

$$\tau_{ij}^o = C_{ij33} \frac{u^o}{a} . \tag{62}$$

The corresponding crack-opening volumes follow from the computed $\bar{V}_{i3}$ as

$$\bar{V}_{i3}^o = \lim_{k_L \to 0} \frac{\bar{V}_{i3}}{ik_L a} \tag{63}$$

It is known (see, e.g., Budiansky and O'Connell, 1976) that a crack opening volume is related to a corresponding stress-intensity factor by

$$\bar{V}_{33}^o = \frac{1-\nu}{3\mu} \bar{\tau}_{33}^o \int_S \rho_c k_I^2 ds, \tag{64}$$

where S is the edge of the crack, ρ_c is a length, and $\bar{\tau}_{33}^o$ is the stress component normal to the crack. Furthermore, $k_I = K_I/\bar{\tau}_{33}^o$ is the reduced Mode I stress-intensity factor. According to Budiansky and Rice (1978) the right-hand-side of equation (66) can be approximated by an expression in terms of the maximum value of k_I. This results in

$$\bar{V}_{33}^o \simeq \frac{(1-\nu)\pi^3}{24\mu} \bar{\tau}_{33}^o [(k_I)_{\max}]^6 \tag{65}$$

Equation (65) gives $(k_I)_{\max}$, the maximum value of the reduced stress intensity factor, in terms of the results of the inverse problem:

$$(k_I^o)_{\max} = \left[\frac{24\mu}{(1-\nu)\pi^3} \frac{\bar{V}_{33}^o}{\bar{\tau}_{33}^o} \right]^{1/6} \tag{66}$$

Now suppose the body is subjected to a service load which gives rise to a static stress field T_{ij}^s. Let

$$T_{ij}^s = \kappa \tau_{ij}^o \tag{67}$$

It then immediately follows that the maximum Mode I stress intensity factor corresponding to the service load is

$$(K_I^s)_{\max} = \kappa \bar{\tau}_{33}^o (k_I^o)_{\max} \tag{68}$$

Analogous calculations can be carried out for combined Mode II and Mode III stress intensity factors.

Another quantity that can be computed from the ultrasonic test results is the total potential energy change ΔW of the solid body due to the presence of the crack. In the limit of zero frequency we obtain from the ultrasonic test results

$$\Delta W^o = \frac{1}{2} \bar{\tau}_{i3}^o \bar{V}_{i3}^o \tag{69}$$

The corresponding results for the static load immediately follows as

$$\Delta W^s = \kappa^2 \Delta W^o, \tag{70}$$

where equation (67) has been used.

Acknowledgment

This work was supported by the Office of Naval Research under Contract N00014-85-K-0401 with Northwestern University.

References

Achenbach, J. D., and Brind, R. J., 1981, "Scattering of Surface Waves by a Sub-Surface Crack," *Journal of Sound and Vibration*, Vol. 76, pp. 43–56.

Achenbach, J. D., Gautesen, A. K., and McMaken, H., 1982, "Ray Methods for Waves in Elastic Solids – With Applications to Scattering by Cracks," Pitman, Boston.

Angel, Y. C., and Achenbach, J. D., 1984, "Reflection and Transmission of Obliquely Incident Rayleigh Waves by a Surface-Breaking Crack," *Journal of the Acoustical Society of America*, Vol. 75, pp. 313–319.

Budiansky, B., and O'Connell, R. J., 1976, "Elastic Moduli of a Cracked Solid," *International Journal of Solids and Structures*, Vol. 12, pp. 81–97.

Budiansky, B., and Rice, J. R., 1978, "On the Estimation of a Crack Fracture Parameter by Long Wavelength Scattering," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 453–454.

Budiansky, B., and Rice, J. R., 1979, "An Integral Equation for Dynamic Elastic Response of an Isolated 3-D Crack," *Wave Motion*, Vol. 1, pp. 187–192.

Burridge, R., and Knopoff, L., 1964, "Body Force Equivalents for Seismic Dislocations," *Bulletin of the Seismological Society of America*, Vol. 54, pp. 1875–1888.

Coffey, J. M., and Chapman, R. K., 1983, "Application of Elastic Scattering Theory for Smooth Flat Cracks to the Quantitative Prediction of Ultrasonic Defect Detection and Sizing," *Nuclear Energy*, Vol. 22, pp. 319–333.

Dong, R., and Adler, L., 1984, "Measurements of Reflection and Transmission Coefficients of Rayleigh Waves From Cracks," *Journal of the Acoustical Society of America*, Vol. 76, pp. 1761–1763.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion and Related Problems," *Proceedings of the Royal Society of London*, Vol. A241. p. 376.

Fu, L. S., 1982, "Mechanics Aspects of NDE by Sound and Ultrasound," *Applied Mechanics Reviews*, Vol. 35, pp. 1047–1057.

Gubernatis, J. E., and Domany, E., 1979, "Rayleigh Scattering of Elastic Waves from Cracks," *Journal of Applied Physics*, Vol. 50, pp. 818–824.

Khuri-Yakub, B. T., Shui, Y., Kino, G., Marshall, D. B., and Evans, A. G., 1984, "Measurement of Surface Machining Damage in Ceramics," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 3, D. O. Thompson and D. E. Chimenti, eds., Plenum Press, New York, pp. 229–238.

Krenk, S., and Schmidt, H., 1982, "Elastic Wave Scattering by a Circular Crack," *Phil. Transactions of the Royal Society of London*, Vol. A308, Vol. 167–198.

Mal, A. K., 1970, "Interaction of Elastic Waves with a Penny-Shaped Crack," *International Journal of Engineering Sciences*, Vol. 8, pp. 381–388.

Martin, P. A., 1982, "Diffraction of Elastic Waves by a Penny-Shaped Crack," *Proceedings of the Royal Society of London*, Vol. A378, pp. 262–285.

Mendelsohn, D. A., Achenbach, J. D., and Keer, L. M., 1980, "Scattering of Elastic Waves by a Surface-Breaking Crack," *Wave Motion*, Vol. 2, pp. 277–292.

More, J., et al., 1980, "User Guide of MINPACK-1," Argonne National Laboratory Report ANL-80-74, Argonne, IL.

Norris, A. N., and Achenbach, J. D., 1982, "Mapping of a Crack Edge by Ultrasonic Methods," *Journal of the Acoustical Society of America*, Vol. 72, pp. 264–272.

Teitel, S., 1978, "Determination of Crack Characteristics from the Quasistatic Approximation for the Scattering of Elastic Waves," *Journal of Applied Physics*, Vol. 49, pp. 5763–5767.

Thompson, R. B., 1983, "Quantitative Ultrasound Nondestructive Evaluation Methods," ASME JOURNAL OF APPLIED MECHANICS, 50th Anniversary Issue, pp. 1191–1201.

Visscher, W. M., 1985, "Elastic Wave Scattering by a Surface-Breaking or Sub-Surface Planar Crack. II. Three-Dimensional Geometry," *Journal of Applied Physics*, Vol. 57, pp. 1538–1550.

Vu, B. Q., and Kinra, V. K., 1985, "Diffraction of Rayleigh Waves in a Half-Space. I. Normal Edge Crack," *Journal of the Acoustical Society of America*, Vol. 77, pp. 1425–1430.

Yew, C. H., Chen, K. G., and Wang, D. L., 1984, "An Experimental Study of Interaction Between Surface Waves and a Surface-Breaking Crack," *Journal of the Acoustical Society of America*, Vol. 75, p. 189.

A. H. Shah

Y. F. Chin

Department of Civil Engineering,
University of Manitoba,
Winnipeg, Canada

S. K. Datta

Department of Mechanical Engineering and
CIRES,
University of Colorado,
Boulder, CO 80309
Fellow ASME

Elastic Wave Scattering by Surface-Breaking Planar and Nonplanar Cracks

Scattering of elastic waves by surface-breaking planar and nonplanar (branched) cracks has been studied in this paper. Attention has been focused on the near-field surface displacements and the crack-tip(s) stress-intensity factors. For planar normal cracks the stress-intensity factors are shown to agree with earlier results. Numerical results showing normalized vertical surface displacements are presented for incident body and surface waves. It is shown that the results for planar and branched cracks can be significantly different in some instances.

Introduction

Problems of elastic wave scattering by surface-breaking and near-surface cracks are of considerable current interest for ultrasonic nondestructive evaluation. Ultrasonic scattering by planar cracks near or at the free surface of a semi-infinite elastic homogeneous medium has been studied theoretically by many authors. References to recent papers on this subject can be found in Shah et al. (1985), Achenbach et al. (1984), and Van der Hijden and Neerhoff (1984). Some experimental works on surface-breaking normal planar cracks have also been reported by Hirao et al. (1982), Yew et al. (1984), and Dong and Adler (1984).

Ultrasonic scattering by surface-breaking planar and branched cracks of arbitrary orientation is the subject of this investigation. To our knowledge this problem has not received much attention in the literature. An approximate solution that is valid at low frequencies was presented by Datta (1979) for *SH* wave diffraction by a canted surface-breaking planar crack. Subsequently, a hybrid finite element and eigenfunction technique was used by Datta et al. (1982) to study *SH* wave diffraction by a planar surface-breaking canted crack.

In this paper we use the same hybrid technique as in Datta et al. (1982) to study the scattering of in-plane body and surface waves by canted planar and normal surface-breaking branched cracks. We focus our attention to the near-field. Numerical results are presented for the vertical surface displacement amplitudes near the base of the crack and crack-tip(s) stress-intensity factors.

Formulation and Solution

Consider a homogeneous, isotropic, and linearly elastic medium with a surface-breaking crack of arbitrary orientation and shape as shown in Figs. 1(*a*) and 1(*b*). It is assumed that the displacement is independent of the z coordinate and its z component is zero. It is further assumed that the displacement $\mathbf{u}(x, y, t)$ at a point P to be time-harmonic of the form $\mathbf{u}(x, y)e^{-i\omega t}$ where ω is the circular frequency. Then $\mathbf{u}$ satisfies the equation of motion in $y > 0$ (at points not on the crack)

$$\mu \nabla^2 \mathbf{u} + (\lambda + \mu) \nabla \nabla \cdot \mathbf{u} = -\rho \omega^2 \mathbf{u} \tag{1}$$

where λ, μ are Lamé constants, ρ the mass density, and the factor $e^{-i\omega t}$ has been dropped.

The solution of equation (1) can be expressed in terms of longitudinal and shear wave potentials, ϕ and ψ, in the form

$$\mathbf{u} = \nabla \phi + \nabla \times (\psi \mathbf{e}_z) \tag{2}$$

Furthermore, in a homogeneous half-space, ϕ and ψ can be expressed in an infinite series of multipolar potentials as (Datta and El-Akily, 1978)

$$\phi = \sum_{n=-\infty}^{\infty} \left(a_n \phi_n^P + b_n \phi_n^S \right)$$
$$\psi = \sum_{n=-\infty}^{\infty} \left(a_n \psi_n^P + b_n \psi_n^S \right) \tag{3}$$

where expressions for ϕ_n^P, ψ_n^P, ϕ_n^S, and ψ_n^S, can be found in Datta and El-Akily (1978). The coefficients a_n, b_n are found by satisfying the appropriate boundary conditions. The series expansion is in terms of multipolar sources located at $0'$ (Figs. 1(*a*) and 1(*b*)) and is valid outside of a circle of radius large enough to enclose the crack inside (see Datta and El-Akily, 1978).

The representation (3) is not useful for satisfying the boundary conditions on the crack surface. For this reason, a different representation is needed in this near-field region. In this paper, the region inside the fictitious boundary B (Figs. 1(*a*) and 1(*b*)) is divided into finite elements having N_I number of interior nodes and N_B number of boundary nodes. In the following the region inside B is denoted by Region II.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August 22, 1986; final revision February 2, 1987.

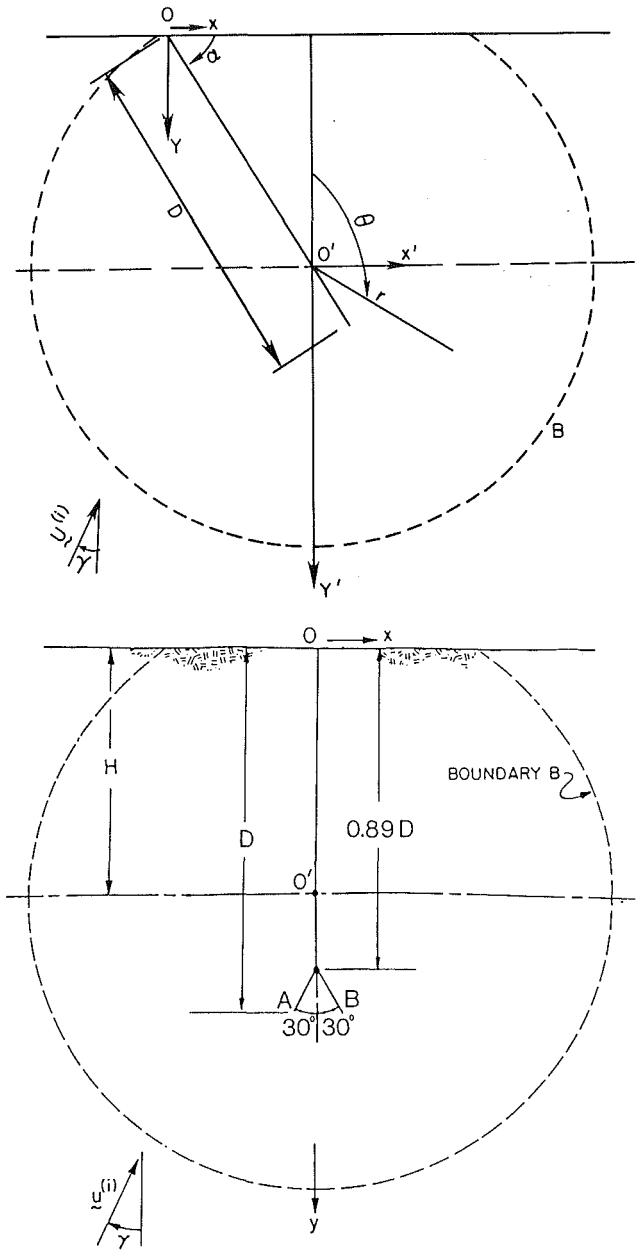


Fig. 1 *(a)* Geometry of a planar canted crack; *(b)* geometry of a branched Y crack

For the finite element representation in Region II, the energy functional is taken to be

$$F=\frac{1}{2}\iint_{\mathrm{II}}\left(\sigma\cdot\epsilon^{*}-\rho\omega^{2}\mathbf{u}\cdot\mathbf{u}^{*}\right)dxdy$$

$$-\frac{1}{2}\int_{B}\left(\mathbf{P}_{B}\cdot\mathbf{u}_{B}^{*}+\mathbf{P}_{B}^{*}\cdot\mathbf{u}_{B}\right)ds \tag{4}$$

where "*" indicates complex conjugate and σ and ϵ are column vectors defined as

$$\sigma=\{\sigma\}=\left(\sigma_{xx},\sigma_{yy},\sigma_{xy}\right)^{T} \tag{5}$$

$$\epsilon=\{\epsilon\}=\left(\epsilon_{xx},\epsilon_{yy},\epsilon_{xy}\right)^{T} \tag{6}$$

Superscript T denotes transpose. The $\mathbf{P}_B$ and $\mathbf{U}_B$ represent the traction and displacement on B, respectively.

It is assumed that the displacement field within the jth element is represented in terms of the shape functions $L_j(x, y)$ and elemental nodal displacements $\{\mathbf{q}_j^e\}$ as

$$\mathbf{u}^{e}=\sum_{j=1}^{N_e}L_j\mathbf{q}_j^{e} \tag{7}$$

where each $\mathbf{q}_j^e$ has two components, u_{xj} and u_{yj}, along the x and y directions, respectively. The number of nodes in each element is given by N_e.

The σ_{ij}^e and ϵ_{ij}^e are computed by substituting equation (7) into strain-displacement relations and these, in turn, into the stress-strain relations. Using these in equation (4), we get

$$F=\mathbf{q}_I^{*T}S_{II}\mathbf{q}_I+\mathbf{q}_I^{*T}S_{IB}\mathbf{q}_B+\mathbf{q}_B^{*T}S_{BI}\mathbf{q}_I+\mathbf{q}_B^{*T}S_{BB}\mathbf{q}_B-\mathbf{q}^{*T}\mathbf{P}_B^{(1)}-\mathbf{P}_B^{*T(1)}\mathbf{q}_B \tag{8}$$

in which $\mathbf{q}_I=\mathbf{q}_I^{(2)}$, $\mathbf{q}_B=\mathbf{q}_B^{(2)}$, $\mathbf{P}_B^{(1)}=\mathbf{P}_B^{(2)}$ and the elemental impedance matrices S_{ij} are defined as

$$\left[S^{e}\right]=\iint_{R_e}\left([B^{e}]^{T}[D][B^{e}]-\rho_e\omega^{2}[L]^{T}[L]\right)dxdy. \tag{9}$$

In equation (9),

$$[B^{e}]=\begin{bmatrix}\dfrac{\partial}{\partial x} & 0\\ 0 & \dfrac{\partial}{\partial y}\\ \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x}\end{bmatrix}\begin{bmatrix}L_1 & 0 & L_2 & \cdots\\ 0 & L_1 & 0 & \cdots\end{bmatrix}=[N][L].$$

Note that $[L]$ is a $2\times 2N_e$ matrix.

For an isotropic material $[D]$ is given by

$$[D]=\begin{bmatrix}\lambda_e+2\mu_e & \lambda_e & 0\\ \lambda_e & \lambda_e+2\mu_e & 0\\ 0 & 0 & \mu_e\end{bmatrix}$$

where λ_e and μ_e are the Lamé's constants.

In representing the inner region into finite elements it has been assumed that the crack faces do not come into contact. Also, in order to model the crack-tip singularity, quarter-point singularity elements have been used. The details of the singularity elements have been discussed by Barsoum (1977) and the extraction of stress intensity factors from the crack opening displacements (COD) has been discussed by Datta and Shah (1982) and Shah et al. (1986).

To find the constants a_n, b_n appearing in equation (3) and the nodal displacements in Region II, it is necessary to use the continuity of displacement and traction on B. This is discussed in the following.

The incident displacement fields will be assumed to arise from the incident plane P and SV waves, and their reflections from the free surface $y = 0$. The case of incident Rayleigh waves will also be considered.

Let us suppose that in the absence of the crack the free field is the sum of the incident and reflected fields, that is

$$u_j^{(0)}=u_j^{(i)}+u_j^{(r)}\quad(j=1,2). \tag{10}$$

For the Rayleigh wave $u_j^{(0)}$ is the associated displacement.

The total field outside B then is

$$u_j=u_j^{(s)}+u_j^{(0)}\quad(j=1,2) \tag{11}$$

where $u_j^{(s)}$ is given by equations (2) and (3).

Using equations (2) and (3), the displacements at the nodes on B can be written as

$$\{\mathbf{q}_B^{(s)}\}=[G]\{\mathbf{a}\} \tag{12}$$

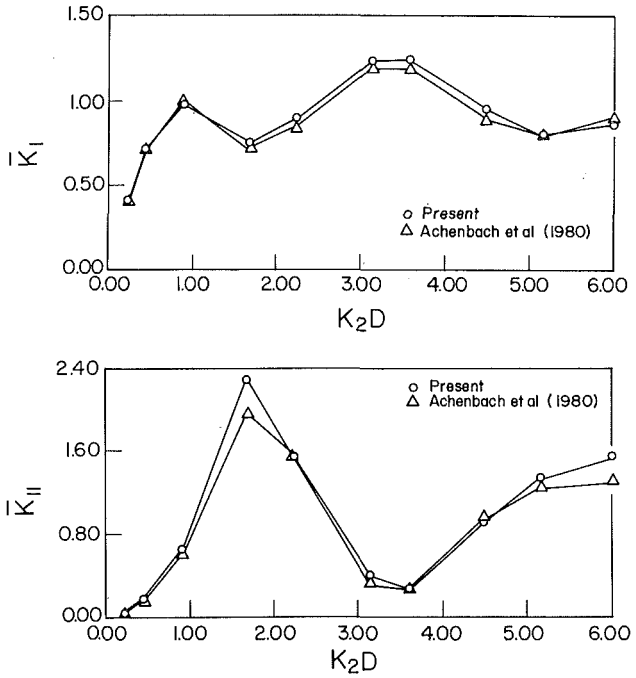


Fig. 2 Comparison of stress-intensity factors calculated by the present method and those given in Achenbach et al. (1980). Incident wave is a Rayleigh wave propagating in the x direction.

where $[G]$ is a $2N_B \times 2N_B$ matrix formulated in Appendix A and vector $\{\mathbf{a}\}$ is

$$\{\mathbf{a}\} = \left[a_1, \ldots, a_{N_B}, b_1, \ldots, b_{N_B}\right]^T$$

Similarly, using equations (2) and (3) in the stress-strain relation, the traction at the nodes on B can be expressed in the form

$$\{\sigma_B^{(s)}\} = [F]\{\mathbf{a}\} \tag{13}$$

where $[F]$ is also a $2N_B \times 2N_B$ matrix defined in Appendix A.

To express $\{\sigma_B^{(s)}\}$ in terms of $\{\mathbf{q}_B^{(s)}\}$, we use the expression for the virtual work done on the boundary B, which is

$$\delta\pi = \int_B \{\delta\mathbf{q}_B^{*(1)}\}^T \{\sigma_B^{(1)}\} d\Gamma \tag{14}$$

where superscript (1) denotes the total field in Region I (outside B).

Because of the continuity of displacements and traction on B, we have

$$\mathbf{q}_B^{(1)} = \mathbf{q}_B^{(2)} = \mathbf{q}_B^{(0)} + \mathbf{q}_B^{(s)} \tag{15}$$

$$\sigma_B^{(1)} = \sigma_B^{(2)} = \sigma_B^{(0)} + \sigma_B^{(s)} \tag{16}$$

where superscript (2) denotes the total field in Region II.

Substituting equations (12), (13), (15), and (16) in equation (14), and noting that $\delta\mathbf{q}_B^{(1)} = \delta\mathbf{q}_B^{(s)}$, we obtain from equation (14)

$$\delta\pi = \{\delta\mathbf{a}^*\}^T \{\mathbf{P}_B^{(1)}\} \tag{17}$$

where $\mathbf{P}_B^{(1)}$ is given by

$$\{\mathbf{P}_B^{(1)}\} = [\bar{R}]\{\mathbf{a}\} + \{\mathbf{P}_B^{(0)}\} \tag{18}$$

Here

$$[\bar{R}] = \int_B [G^*]^T [F] d\Gamma \tag{19}$$

and

$$\{\mathbf{P}_B^{(0)}\} = \int_B [G^*]^T \{\sigma_B^{(0)}\} d\Gamma. \tag{20}$$

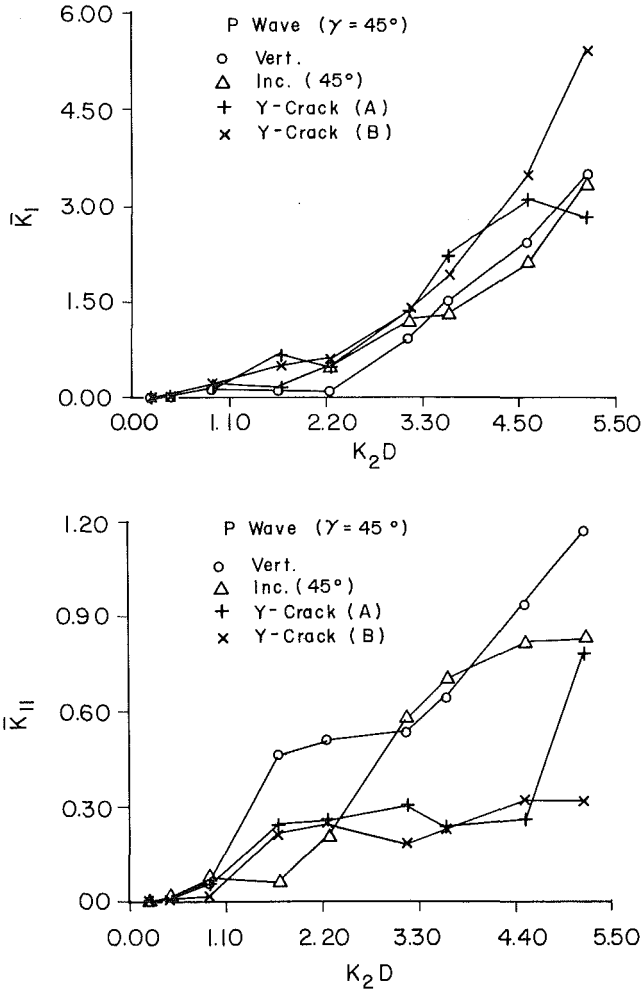


Fig. 3 Stress-intensity factors for various crack geometries due to incident P wave

Equations (19) and (20) are approximated by

$$[\bar{R}] = [G^*]^T [F] R\,\Delta\theta \tag{21}$$

and

$$\{\mathbf{P}_B^{(0)}\} = [G^*]^T \{\sigma_B^{(0)}\} R\,\Delta\theta \tag{22}$$

where $R\,\Delta\theta$ is the arc length between two adjacent boundary nodes on contour B. Note that the first two rows and last two rows of equations (21) and (22) are multiplied by $R\,\Delta\theta/2$ instead of $R\,\Delta\theta$, because they correspond to the first and last boundary nodes, respectively.

Substituting equation (12) in equation (8) and taking the variation, we obtain a set of simultaneous equations which may be written in matrix form as

$$\begin{bmatrix} S_{II} & S_{IB}G \\ G^{*T}S_{IB}^T & G^{*T}S_{BB}G \end{bmatrix} \begin{bmatrix} \mathbf{q}_I \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} -S_{IB}\mathbf{q}_B^{(0)} \\ -G^{*T}S_{BB}\mathbf{q}_B^{(0)} + \mathbf{P}_B^{(1)} \end{bmatrix}. \tag{23}$$

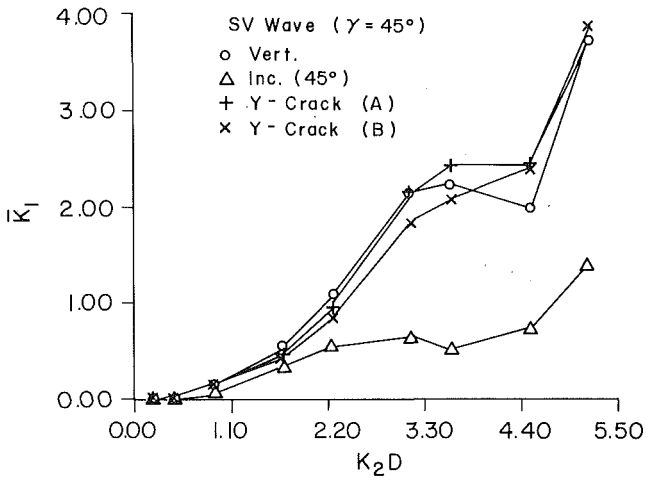
The first equation of equation (23) can be written as:

$$\mathbf{q}_I = -S_{II}^{-1}\left[S_{IB}G\mathbf{a} + S_{IB}\mathbf{q}_B^{(0)}\right]. \tag{24}$$

The second equation can be written as

$$G^{*T}S_{IB}^T\mathbf{q}_I + G^{*T}S_{BB}G\,\mathbf{a} = -G^{*T}S_{BB}\mathbf{q}_B^{(0)} + \mathbf{P}_B^{(1)}. \tag{25}$$

Substituting equations (18) and (24) into equation (25), we obtain



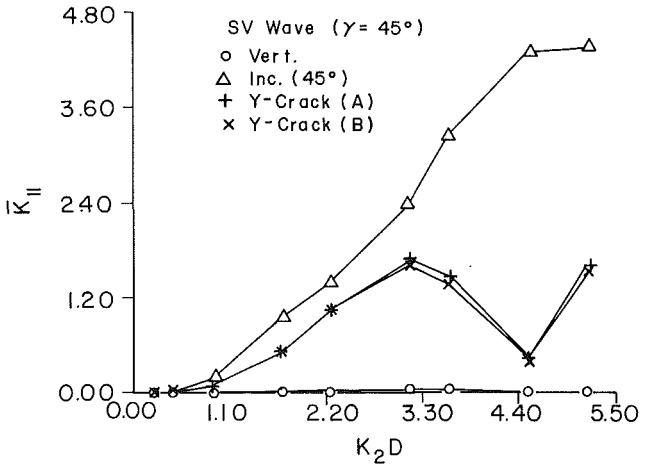
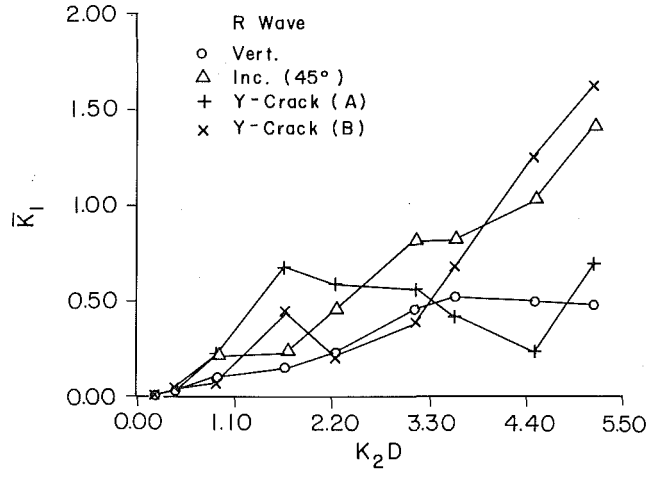


Fig. 4 Stress-intensity factor for various crack geometries due to incident *SV* wave



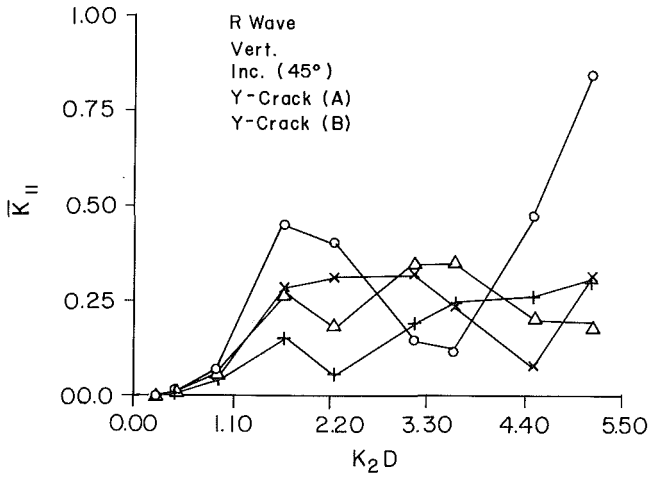


Fig. 5 Stress-intensity factor for various crack geometries due to incident Rayleigh wave propagating in the *x* direction

$$\left[G^{*T}\left(S_{BB}-S_{IB}^{T}S_{II}^{-1}S_{IB}\right)G-\bar{R}\right]\{\mathbf{a}\}=$$
$$-G^{*T}\left(S_{BB}-S_{IB}^{T}S_{II}^{-1}S_{IB}\right)\mathbf{q}_B^{(0)}+\mathbf{P}_B^{(0)} \tag{26}$$

In equation (26), the generalized coordiantes {**a**} are the only unknowns. Therefore, {**a**} can be evaluated. Once {**a**) are known, the near and far displacement and stress fields can be determined.

Numerical Results and Discussion

In this paper, the boundary B enclosing the interior region is not a complete circle, and so the potentials ϕ_n^P, ψ_n^P, and ϕ_n^S, ψ_n^S cannot be expanded in cylindrical wave functions as was done by Shah et al. (1985). Instead the integrals giving these potentials and their derivatives were evaluated numerically for every node on B. These integrals are of the same forms as the coefficients (P_{mn}, Q_{mn}, etc.) that appear in the wave function expansion used by Shah et al. (1985). Thus they are evaluated by using the same deformed path (El-Akily and Datta, 1981) that was used to calculate those coefficients. For more details the reader is referred to Datta and Sabina (1986) and Chin (1985). It was found that 24 terms were needed in the series expansions (3) for the range of frequencies ($0 < k_2 D < 5.5$) considered here.

The hybrid method is employed to study scattering by *P*, *SV*, and Rayleigh waves by three types of surface breaking cracks: a vertical crack (Fig. 1(*a*) with $\alpha = 90$ deg), a 45 deg inclined crack (Fig. 1(*a*), with $\alpha = 45$ deg), and a vertical branched (*Y*) crack (Fig. 1(*b*)). The Poisson's ratio of the material is taken to be 1/3. The method used here is applicable to a surface-breaking crack of arbitrary geometry. It can be used for a planar or a *Y* crack of arbitrary orientation as well as for a *Y* crack of arbitrary branch lengths and opening angles. Of course, if the branches are too close, then very fine meshes would be required. The examples chosen here are only for illustrative purposes.

Stress intensity factors at the tips of the cracks were calculated and for the particular case of a planar surface-breaking normal crack they were found to agree well with the results of Achenbach et al. (1980). These are shown in Fig. 2.

Next, normalized stress-intensity factors are shown in Figs. 3–5 for incident *P*, *SV*, and Rayleigh waves. It is seen that for *P* and Rayleigh waves the stress-intensity factors at the crack-tips of the three types of cracks are quite different, particularly at high frequencies. This is particularly significant for the branched crack, even though the branches are quite small.

The surface displacements at $y = 0$ are calculated by using equations (3) in (2) after {**a**} are calculated. Normalized values of $u_y^{(s)}$ are presented in Figs. 6–10. For each type of crack mentioned above, five cases of incident waves were considered: plane *P* wave incident at 0 deg and 45 deg, plane *SV* wave incident at 0 deg and 45 deg, and finally the Rayleigh wave. Some representative results are shown here. In these figures *NF* represents the normalization factor.

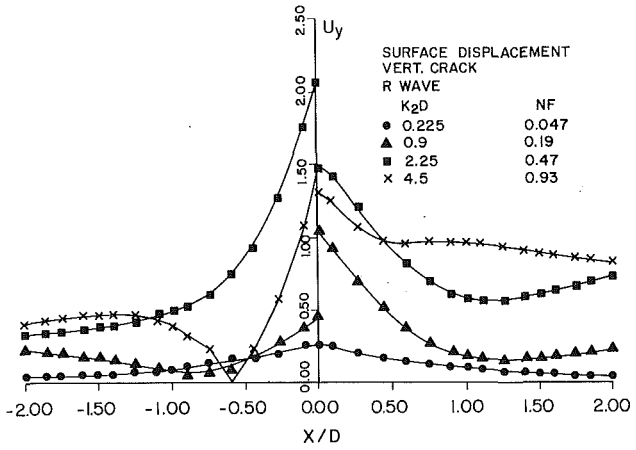


Fig. 6 Normalized scattered vertical surface displacement amplitude due to Rayleigh wave incident on a normal surface-breaking planar crack

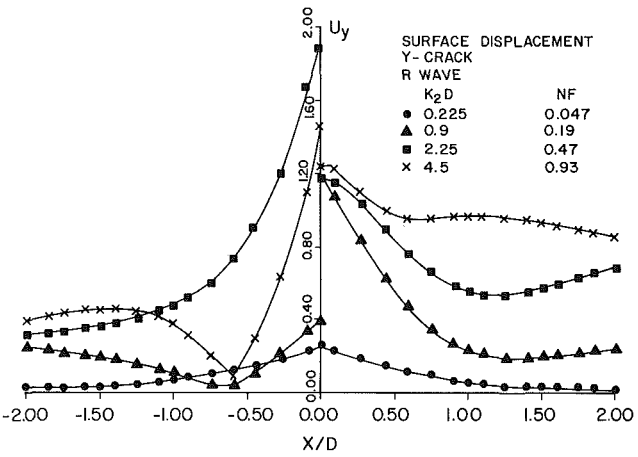


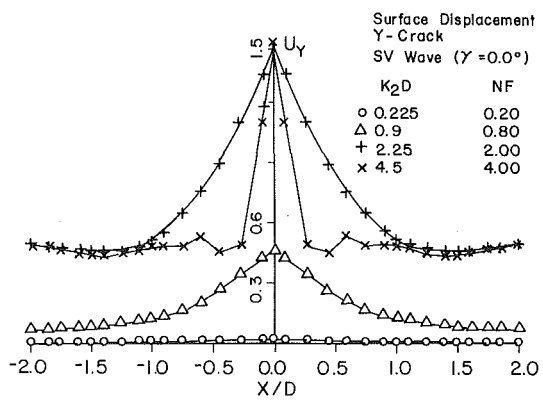
Fig. 7 Normalized scattered vertical surface displacement amplitude due to Rayleigh wave incident on a normal surface-breaking branched crack

Figures 6 and 7 show the scattered vertical surface displacement amplitudes for a Rayleigh wave incident from the left on a normal planar and branched crack. It is seen that there are large differences in the forward direction between the two cases as the frequency becomes large. In the backward direction, however, the differences are not very significant. Figures 8 and 9 show the results for an incident *SV* wave moving vertically as well as at 45 deg to the vertical. Large differences are found for vertical incidence, but not in the other case. Finally, Fig. 10 shows the case of a Rayleigh wave incident on a canted crack. This figure is to be contrasted with Fig. 6. The large differences shown clearly distinguish a canted crack from a normal crack.

Conclusion

Model calculations of elastic wave scattering by surface-breaking planar and nonplanar cracks have been presented. These calculations show that near-field surface displacements due to scattering by planar and branched cracks are quite different even when the branches are small. Also, it is found that signatures of normal and canted cracks are very dissimilar. These characteristic differences can be used to discriminate between the various cases.

In this paper we have confined our attention to near-field results. However, once the {**a**} are known they can be used to compute displacements at any point outside the circular arc B.



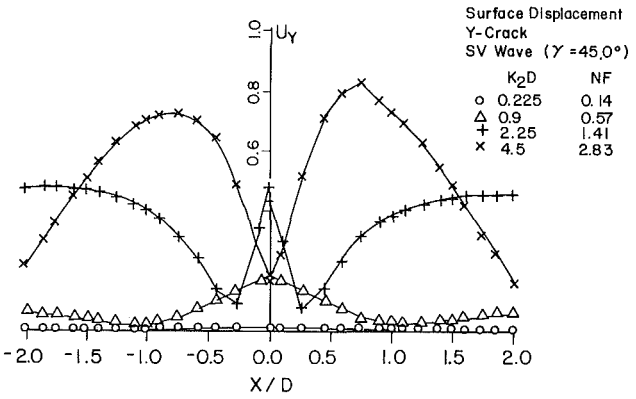
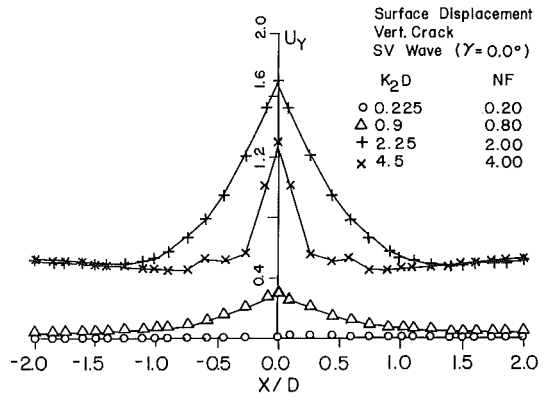


Fig. 8 Normalized scattered vertical surface displacement amplitude due to *SV* wave incident on a normal surface-breaking planar crack



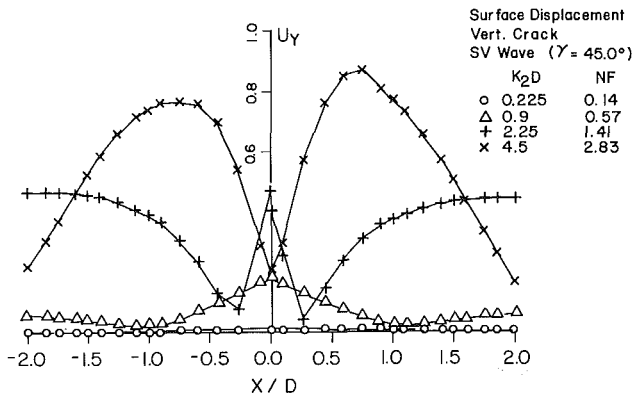


Fig. 9 Normalized scattered vertical surface displacement amplitude due to *SV* wave incident on a normal surface-breaking branched crack

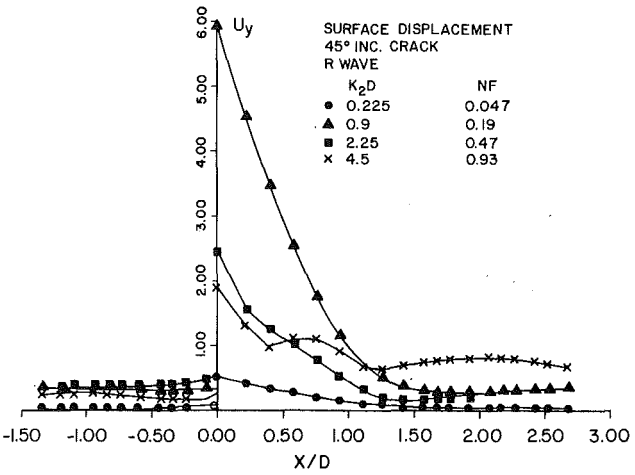


Fig. 10 Normalized scattered vertical surface displacement amplitude due to Rayleigh wave incident on a normal surface-breaking canted crack

Although the results presented here are for homogeneous medium, the technique can be generalized to study cracks in a composite medium. These are presently under investigation and will be reported later.

Acknowledgment

Results presented here were partly supported by grants from the Office of Naval Research (N00014-86-K-0280), the National Science Foundation (CEE81-20536), and the Natural Science and Engineering Research Council of Canada (A-7988).

References

Achenbach, J. D., Angel, Y. C., and Lin, W., 1984, "Scattering from Surface-Breaking and Near-Surface Cracks," *Wave Propagation in Homogeneous Media and Ultrasonic Nondestructive Evaluation*, Johnson, G. C., ed., AMD-Vol. 62, ASME, New York, pp. 93–109.

Achenbach, J. D., Keer, L. M., and Mendelsohn, D. A., 1980, "Elastodynamic Analysis of an Edge Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 551–556.

Barsoum, R. S., 1977, "Triangular Quarter-Point Elements as Elastic and Perfectly Plastic Crack Tip Elements," *International Journal of Numerical Methods in Engineering*, Vol. 11, pp. 85–98.

Chin, Y. F., 1985, *Scattering of Elastic Waves by Near-Surface Inhomogeneities*, M.S. Thesis, University of Manitoba.

Datta, S. K., 1979, "Diffraction of SH-Waves by Edge Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 101–106.

Datta, S. K., and El-Akily, N., 1978, "Diffraction of Elastic Waves by Cylindrical Cavity in a Half-Space," *Journal of the Acoustical Society of America*, Vol. 64, pp. 1692–1699.

Datta, S. K., and Sabina, F. J., 1986, "Matched Asymptotic Expansions Applied to Diffraction of Elastic Waves," *Acoustic, Electromagnetic and Elastic Wave Scattering, Vol. 2: Low and High Frequency Asymptotics*, Varadan, V. K., and Varadan, V. V., eds., North-Holland, Amsterdam, Chapter 2, pp. 71–264.

Datta, S. K., and Shah, A. H., 1982, "Scattering of SH Waves by Embedded Cavities," *Wave Motion*, Vol. 4, pp. 265–283.

Datta, S. K., Shah, A. H., and Fortunko, C. M., 1982, "Diffraction of Medium and Long Wavelength Horizontally Polarized Shear Waves by Edge Cracks," *Journal of Applied Physics*, Vol. 53, pp. 2895–2903.

Dong, R., and Adler, L., 1984, "Measurements of Reflection and Transmission Coefficients of Rayleigh Waves from Cracks," *Journal of the Acoustical Society of America*, Vol. 76, pp. 1761–1763.

El-Akily, N., and Datta, S. K., 1981, "Response of a Circular Cylindrical Shell to Disturbances in a Half-Space—Numerical Results," *Earthquake Engineering and Structural Dynamics*, Vol. 9, pp. 477–487.

Hirao, M., Fukuoka, H., and Miura, Y., 1982, "Scattering of Rayleigh Surface Waves by Edge Cracks: Numerical Simulation and Experiment," *Journal of the Acoustical Society of America*, Vol. 72, pp. 602–606.

Shah, A. H., Wong, K. C., and Datta, S. K., 1985, "Surface Displacements Due to Elastic Wave Scattering by Planar and Non-Planar Cracks," *Wave Motion*, Vol. 7, pp. 319–333.

Shah, A. H., Wong, K. C., and Datta S. K., 1986, "Dynamic Stress Intensity Factors for Buried Planar and Non-planar Cracks," *Internatinoal Journal of Solids and Structures*, Vol. 22, pp. 845–857.

van der Hijden, J. H. M. T., and Neerhoff, F. L., 1984, "Diffraction of Elastic Waves by a Sub-Surface Crack (in-Plane Motion)," *Journal of the Acoustical Society of America*, Vol. 75, pp. 1694–1704.

Yew, C. H., Chen, K. G., and Wang, D. L., 1984, "An Experimental Study of Interaction Between Surface Waves and a Surface Breaking Crack," *Journal of the Acoustical Society of America*, Vol. 75, pp. 189–196.

APPENDIX A

Formulation of Matrices [G] and [F]

As mentioned before, the scattered nodal displacement vector, $\{\mathbf{q}_B^{(s)}\}$, was formed by evaluating $u_x^{(s)}$ and $u_y^{(s)}$ at N_B number of points on contour B. Thus, we have:

$$\{\mathbf{q}_B^{(s)}\} = [G]\{\mathbf{a}\} \tag{A-1}$$

where

$$\{\mathbf{q}_B^{(s)}\} = \{u_{xB_1}, \ldots, u_{xB_{N_B}}, u_{yB_1}, \ldots, u_{yB_{N_B}}\}^T \tag{A-2}$$

$$\{\mathbf{a}\} = \{a_1, \ldots, a_{N_B}, b_1, \ldots, b_{N_B}\}^T \tag{A-3}$$

If $[G]$ is partitioned as,

$$\begin{bmatrix} GXA & GXB \\ \hline GYA & GYB \end{bmatrix}_{2N_B \times 2N_B} \tag{A-4}$$

then each of the $N_B \times N_B$ submatrices can be evaluated from equations (2) and (3) at (x_i, y_i) on B as

$$\begin{aligned} (GXA)_{in} &= \left(\phi_{n,x}^P + \psi_{n,y}^P\right) \\ (GXB)_{in} &= \left(\phi_{n,x}^S + \psi_{n,y}^S\right) \\ (GYA)_{in} &= \left(\phi_{n,y}^P - \psi_{n,x}^P\right) \\ (GYB)_{in} &= \left(\phi_{n,y}^S - \psi_{n,x}^S\right), \; i = 1 \text{ to } N_B \end{aligned} \tag{A-5}$$

The parameter n in the summation series of equation (3) is taken from $-(N_B/2 - 1)$ to $N_B/2$ for numerical purposes. Hence, n in the first column to the last column of each submatrix corresponds to $-(N_B/2 - 1)$ to $N_B/2$, respectively.

To formulate the matrix $[F]$, the components T_x and T_y of the traction vector T were calculated at each nodal point on B. If the radius vector of that point makes an angle θ with the x axes, then

$$\begin{aligned} T_x &= \sigma_{xx}\cos\theta + \sigma_{xy}\sin\theta \\ T_y &= \sigma_{xy}\cos\theta + \sigma_{yy}\sin\theta. \end{aligned} \tag{A-6}$$

Evaluating $\sigma_{xx}^{(s)}$, $\sigma_{yy}^{(s)}$, and $\sigma_{xy}^{(s)}$ at N_B number of points on contour B and substituting in equation (A-6), we have the scattered nodal stress vector, $\{\sigma_B^{(s)}\}$, as in equation (13),

$$\{\sigma_B^{(s)}\} = [F]\{\mathbf{a}\} \tag{A-7}$$

where

$$\{\sigma_B^{(s)}\} = \{T_{xB_1}, \ldots, T_{xB_{N_B}}, \ldots, T_{yB_{B_1}}, \ldots, T_{yB_{N_B}}\} \tag{A-8}$$

and $\{\mathbf{a}\}$ is defined in equation (A-3)

$[F]$ is partitioned as

$$\begin{bmatrix} FXA & FXB \\ \hline FYA & FYB \end{bmatrix}_{2N_B \times 2N_B}$$

Each of the $N_B \times N_B$ submatrices can be evaluated at (x_i, y_i) on B as,

$$(FXA)_{in} = \left[(\lambda + 2\mu)(\phi^P_{n,xx} + \psi^P_{n,xy}) + \lambda(\phi^P_{n,yy} - \psi^P_{n,xy}) \right] \cos\theta + \mu(2\phi^P_{n,xy} + \psi^P_{n,yy} - \psi^P_{n,xx}) \sin\theta$$

$$(FXB)_{in} = \left[(\lambda + 2\mu)(\phi^S_{n,xx} + \psi^S_{n,xy}) + \lambda(\phi^S_{n,yy} - \psi^S_{n,xy}) \right] \cos\theta + \mu(2\phi^S_{n,xy} + \psi^S_{n,yy} - \psi^S_{n,xx}) \sin\theta$$

$$(FYA)_{in} = \mu(2\phi^P_{n,xy} + \psi^P_{n,yy} - \psi^P_{n,xx}) \cos\theta + \left[(\lambda + 2\mu)(\phi^P_{n,yy} - \psi^P_{n,xy}) + \lambda(\phi^P_{n,xx} + \psi^P_{n,xy}) \right] \sin\theta$$

$$(FYB)_{in} = \mu(2\phi^S_{n,xy} + \psi^S_{n,yy} - \psi^S_{n,xx}) \cos\theta + \left[(\lambda + 2\mu)(\phi^S_{n,yy} - \psi^S_{n,xy}) + \lambda(\phi^S_{n,xx} + \psi^S_{n,xy}) \right] \sin\theta$$

Parameter n ranges from $-(N_B/2 - 1)$ to $N_B/2$ as discussed before, and $i = 1$ to N_B.

R. T. Shield
Department of Theoretical
and Applied Mechanics,
University of Illinois at Urbana-Champaign,
Urbana, IL 61801
Fellow ASME

Variational Principles for Some Nonstandard Elastic Problems

Variational principles are derived for some nonstandard problems involving elastic bodies in smooth contact. For these problems, the portions of the surfaces where one boundary condition holds rather than another must be determined as part of the solution to the problem. Cases considered include a body containing a crack or delamination, indentation by a rigid punch, and contact with an elastic foundation.

Introduction

The principles of minimum potential energy and minimum complementary energy for infinitesimal elastic deformations are well known for problems in which the tractions or the displacements are specified at each point of the surface of an elastic body (see Sokolnikoff, 1956, for example). The principles can lead to bounds on quantities of physical interest and can be used to obtain approximate analytical and numerical solutions. For some problems, the portions of the surface where one boundary condition holds rather than another must be determined as part of the solution to the problem. For example, in problems involving contact between elastic bodies, the shape of the contact area can vary with the loading (for references see Gladwell, 1980). Uniqueness of solution for typical problems of this type was considered by Shield (1982) for problems involving smooth contact between surfaces of elastic bodies. Here we again consider elastic problems involving smooth contact and develop variational principles for some typical situations: loading of a body containing a crack across which there is no cohesion, bodies in smooth contact, indentation by a smooth rigid punch, and contact with a smooth elastic foundation. The examples can be combined to treat more complex problems, such as the indentation of an elastic body containing a crack.

The principles rest on the positive-definiteness of the strain energy and they show that the potential and complementary energies attain absolute minimum values only when the trial functions generate the strains or stresses of the actual solution. Weaker stationary principles apply without the assumption of positive-definiteness of the strain energy. It is assumed that the integrals involved are convergent if they are improper and this requires the states considered to have finite total strain energy or total complementary energy. The elastic material can be inhomogeneous. At an interface between two materials in a composite, the displacements and traction are assumed to be continuous across the interface but delaminations can be included if they are modeled as cracks across which there is no cohesion.

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 2, 1987.

Elastic Body With a Crack

We assume that the strain-energy density W of the body is positive-definite and we write

$$2W(e) = c_{ijkl}e_{ij}e_{kl} \quad (c_{ijkl} = c_{klij} = c_{jikl}),$$

where e_{ij} are the infinitesimal strains referred to rectangular Cartesian axes x_i. The stresses t_{ij} are related to the strains through

$$t_{ij} = c_{ijkl}e_{kl}, \quad e_{ij} = C_{ijkl}t_{kl},$$

where C_{ijkl} have the same symmetries as c_{ijkl}. The strain-energy is also a positive-definite function W_C of the stresses,

$$2W_C(t) = C_{ijkl}t_{ij}t_{kl}.$$

In the unstressed reference state, the body occupies a region V with surface S and we suppose that it contains a crack across which the material has no cohesion. The crack is defined by a surface C in V, and we use $\mathbf{n}$ to denote the unit normal to one side of the surface C. We use $\pm$ signs to indicate values of quantities on the two sides of C, with the $+$ sign referring to the side of C with exterior normal $\mathbf{n}$. We shall also use square brackets to indicate the difference in the values of a quantity across C, so that for the displacement field $\mathbf{u}$, for example,

$$[\mathbf{u}] = \mathbf{u}^+ - \mathbf{u}^-.$$

Under loading of the body, we assume that at points of the crack surface C either (i) the crack opens with no traction transmitted across C or (ii) the two sides remain in smooth contact. Then at points of C we have

$$(\mathbf{u}^+ - \mathbf{u}^-)\cdot\mathbf{n} = [\mathbf{u}]\cdot\mathbf{n} \leq 0, \tag{1}$$

and we require on C

$$\left.\begin{aligned} &\text{either } (i)\ \mathbf{T}^+ = -\mathbf{T}^- = 0 \quad \text{when } [\mathbf{u}]\cdot\mathbf{n} < 0 \\ &\text{or } (ii)\ \mathbf{T}^+ = -\mathbf{T}^+ = -p\mathbf{n} \quad \text{when } [\mathbf{u}] = 0, \end{aligned}\right\} \tag{2}$$

where $\mathbf{T}$ is the surface traction and $p(x)$ is the (nonnegative) pressure transmitted across C. We set $p = 0$ at points of C where (i) holds.

The loading of the body is assumed to be caused by a body force $\mathbf{F}$, prescribed surface tractions $\mathbf{T}^G$ on a portion S_T of S and prescribed displacements $\mathbf{u}^G$ on the remainder S_U of S. The displacement field $\mathbf{u}$ is unique except possibly for a rigid displacement, depending on the conditions on S_U (Shield, 1982).

We define the potential energy P as the functional

$$P\{u'\} = \int_V W(e')dV - \int_{S_T} \mathbf{T}^G \cdot \mathbf{u}' dS - \int_V \mathbf{F} \cdot \mathbf{u}' dV \quad (3)$$

for fields $\mathbf{u}'$ such that

$$\mathbf{u}' = \mathbf{u}^G \text{ on } S_U, \quad [\mathbf{u}'] \cdot \mathbf{n} \leq 0 \text{ on } C. \quad (4)$$

We set

$$\mathbf{u}' = \mathbf{u} + \Delta\mathbf{u}.$$

Then $\Delta\mathbf{u}$ is zero on S_U and we have

$$W(e') = W(e) + t_{ij}\Delta e_{ij} + W(\Delta e). \quad (5)$$

The stresses t_{ij} satisfy equilibrium with body force $\mathbf{F}$ and with the divergence theorem and boundary conditions we get

$$P\{u'\} - P\{u\} = \int_V W(\Delta e)dV + \int_C \mathbf{T} \cdot [\Delta\mathbf{u}]dS,$$

where contributions from both sides of C have been included in the integral over C. For the actual solution $\mathbf{T} \cdot [\mathbf{u}]$ is zero on C in view of equations (2) and the integral over C has the value

$$\int_C \mathbf{T} \cdot [\mathbf{u}']dS = -\int_C p\mathbf{n} \cdot [\mathbf{u}']dS.$$

From equation (4) this is seen to be greater than or equal to zero and with W positive-definite, we have

$$P\{u'\} \geq P\{u\}$$

with equality if and only if $\mathbf{u}$ and $\mathbf{u}'$ have the same strains and $\mathbf{n} \cdot [\mathbf{u}']$ is zero where p is nonzero.

Thus we have: *For displacements which satisfy the displacement boundary conditions and have no interpenetration of material across the crack, the potential energy* P *is least for the displacements of the actual solution.*

The complementary energy Q is defined as

$$Q\{t'\} = \int_V W_C(t')dV - \int_{S_U} \mathbf{T}' \cdot \mathbf{u}^G dS \quad (6)$$

for stresses t'_{ij} in equilibrium with $\mathbf{F}$ and such that

$$t_{ij}n_j = T_i^G \text{ on } S_T, \quad \mathbf{T}'^{+} = -\mathbf{T}'^{-} = -p'\mathbf{n} \text{ on } C,$$

where $p' \geq 0$. We set

$$t'_{ij} = t_{ij} + \Delta t_{ij},$$

so that the stresses Δt_{ij} satisfy equilibrium with no body force and have zero traction on S_T. We have

$$W_C(t') = W_C(t) + \Delta t_{ij}e_{ij} + W(\Delta t), \quad (7)$$

and with the divergence theorem we obtain

$$Q\{t'\} - Q\{t\} = \int_V W_C(\Delta t)dV + \int_C \Delta T \cdot [\mathbf{u}]dS.$$

As before $\mathbf{T} \cdot [\mathbf{u}]$ is zero on C and the integral over C becomes

$$\int_C \mathbf{T}' \cdot [\mathbf{u}]dS = -\int_C p'\mathbf{n} \cdot [\mathbf{u}]dS \geq 0.$$

It follows that

$$Q\{t'\} \geq Q\{t\}$$

with equality if and only if $t'_{ij} = t_{ij}$ in V, and we have: *For stress fields with the given surface tractions and in equilibrium with the given body force and which transmit at most pressure across the crack, the complementary energy* Q *is least for the stresses of the actual solution.*

For the actual solution we have

$$P\{u\} + Q\{t\} = 2\int_V WdV - \int_S \mathbf{T} \cdot \mathbf{u}dS - \int_V \mathbf{F} \cdot \mathbf{u}dV = 0, \quad (8)$$

using the divergence theorem and $\mathbf{T} \cdot [\mathbf{u}] = 0$ on the crack (Clapeyron's theorem). We then have

$$P\{u'\} \geq P\{u\} \geq -Q\{t'\}, \quad (9)$$

and the principles provide upper and lower bounds for the potential energy.

Smooth Contact Between Elastic Bodies

For simplicity we consider contact between two elastic bodies occurring over parts of the surfaces of the bodies which are nearly plane, but the approach is easily generalized. In the reference configuration the bodies touch at the origin 0 of coordinates and the plane $x_3 = 0$ is tangent to both bodies at 0. The bodies occupy regions V_1 and V_2 with the x_3 axis pointing into V_2. Under loading, contact may occur over surfaces C_1 and C_2 of the bodies, defined as the nearly plane surfaces

$$C_1: x_3 = f(x_1, x_2), \quad C_2: x_3 = g(x_1, x_2),$$

where x_1, x_2 lie in a region C of the $x_1 - x_2$ plane enclosing the origin and

$$f(x_1, x_2) \leq g(x_1, x_2).$$

The contact is smooth and we denote the pressure between the two bodies by the nonnegative function $p(x_1, x_2)$, defined over C. Then

$$T_3^2 = -T_3^1 = p(x_1, x_2) \text{ on } C, \quad (10)$$

where the superscripts indicate values for the two bodies. The bodies do not penetrate each so that the displacement component u_3 satisfies

$$u_3^2 - u_3^1 \geq f - g \text{ on } C. \quad (11)$$

At each point of C we require

$$p = 0 \text{ when } u_3^2 - u_3^1 > f - g, \; u_3^2 - u_3^1 = f - g \text{ when } p > 0. \quad (12)$$

On the remaining portions S_1 and S_2 of the surfaces of the two bodies, we suppose that tractions $\mathbf{T}^G$ are specified on parts S_{1T}, S_{2T} and displacements $\mathbf{u}^G$ are specified on parts S_{1U}, S_{2U} of S_1, S_2, respectively. We use S_T and S_U to denote $S_{1T} + S_{2T}$ and $S_{1U} + S_{2U}$. The body force $\mathbf{F}$ is assumed known in $V = V_1 + V_2$. The solution is unique except possibly for a rigid body displacement (Shield, 1982).

The potential energy P is again defined by equation (3) for displacements which satisfy the boundary conditions on S_U and which satisfy (11) on C. As in the previous section, we look at the difference between $P\{u'\}$ and $P\{u\}$, use equation (5) and the divergence theorem, and find that the difference involves the integral

$$\int_{C_1 + C_2} \mathbf{T} \cdot (\mathbf{u}' - \mathbf{u})dS.$$

In view of equations (10) and (12), the integral becomes

$$\int_C p\{u_3'^2 - u_3'^1 - (f - g)\}dS,$$

and this is nonnegative because u_3' satisfies (11). Thus we can show that: *For displacements which satisfy the displacement boundary conditions and which satisfy (11) on* C, *the potential energy is least for the displacements of the actual solution.*

The complementary energy Q is defined by

$$Q\{t'\} = \int_V W_C(t')dV - \int_{S_U} \mathbf{T}' \cdot \mathbf{u}^G dS - \int_C p'(f - g)dS \quad (13)$$

for stresses t'_{ij} in equilibrium with the given body force and given tractions and which involve at most a pressure p' between the bodies across C. The difference between $Q\{t'\}$ and $Q\{t\}$ is transformed as in the previous section and we are led to consider the sign of

$$\int_{C_1+C_2} (\mathbf{T}'-\mathbf{T})\cdot\mathbf{u}dS - \int_C (p'-p)(f-g)dS.$$

This can be written as

$$\int_C (p'-p)\{u_3^2-u_3^1-(f-g)\}dS.$$

and with equations (11) and (12) we see that the integral is nonnegative. Thus we can show that: *For stress fields in equilibrium with the given body force and given tractions and which involve at most a pressure between the bodies across* C, *the complementary energy is least for the stresses of the actual solution.*

For the actual solution we again have $P\{u\} = -Q\{t\}$ and the principles lead to upper and lower bounds on $P\{u\}$.

Indentation by a Smooth Rigid Punch

In order to illustrate problems in which an elastic body can come into contact with a rigid body of known shape, we consider indentation by a smooth rigid punch when the possible area of contact is a region C of the x_1-x_2 plane enclosing the origin 0. The exterior normal to the body at 0 is along the x_3 axis, and the remainder of the surface of the body is denoted by S.

When the movement of the punch is known, we will have

$$u_3 \le g(x_1, x_2) \text{ on } C, \tag{14}$$

where g is a known function, and we require at each point of C

$$p=0 \text{ when } u_3 < g,\ u_3 = g \text{ when } p > 0, \tag{15}$$

where p is the contact pressure. Tractions are prescribed on a part S_T of S and displacements on the remainder S_U of S, with a known body force in the region V occupied by the body.

The problem can be considered as a limiting case of contact between two elastic bodies. The potential energy is defined to be the functional (3) for displacements which satisfy the conditions on S_U and (14) on C. We can then proceed as in the previous section to show that: *The potential energy is least for the displacements of the actual solution.*

The complementary energy Q is defined to be

$$Q\{t'\} = \int_V W_C(t')dV - \int_{S_U} \mathbf{T}'\cdot\mathbf{u}^G dS + \int_C p'g dS$$

for stresses in equilibrium with the given body force and surface tractions and which involve a pressure p' in the contact area C. Then: Q *will be an absolute minimum for the stresses of the actual solution.*

Instead of prescribing the movement of the punch, we may prescribe the downward force L on the punch and the moments M_1, M_2 of the force about the x_1, x_2 axes, with prescribed loading on S as before. The contact pressure p must then satisfy

$$L = \int_C p dS, \quad M_1 = -\int_C p x_2 dS, \quad M_2 = \int_C p x_1 dS. \tag{16}$$

For a known punch shape $g(x_1, x_2)$ the solution will satisfy

$$u_3 \le g(x_1, x_2) - d + ax_2 - bx_1 \text{ on } C, \tag{17}$$

with equality where p is nonzero. The constants d, a, b are determined as part of the solution, which will be unique except possibly for a rigid body displacement (Shield, 1982). (Other problems may be treated; for example, we may require the punch to indent without tilting and then a, b are zero and M_1, M_2 are not prescribed.)

The potential energy is defined as

$$P\{u'\} = \int_V W(e')dV - \int_{S_T} \mathbf{T}^G\cdot\mathbf{u}'dS$$

$$-\int_V \mathbf{F}\cdot\mathbf{u}'dV - Ld' - M_1a' - M_2b' \tag{18}$$

for fields $\mathbf{u}'$ which satisfy the displacement conditions on S_U and which are such that

$$\mathbf{u}'_3 \le g(x_1, x_2) - d' + a'x_2 - b'x_1 \text{ on } C, \tag{19}$$

where d', a', b' are constants. After transformation, we find that

$$P\{\mathbf{u}'\} - P\{u\} = \int_V W(\Delta e)dV$$

$$-\int_C p(\mathbf{u}'_3 - u_3)dS - L(d'-d) - M_1(a'-a) - M_2(b'-b).$$

If we set

$$v = u_3 + d - ax_2 + bx_1,\ v' = u'_3 + d' - a'x_2 + b'x_1$$

and use equations (16), then apart from the strain-energy term the right-hand side becomes

$$-\int_C p(v'-v)dS.$$

Now where $p > 0$, equality holds in (17) and $v = g$. Because $v' \le g$ from (19), we then see that the integrand is nonpositive. Thus: *The potential energy is least for the displacements of the actual solution.*

The complementary energy is defined to be

$$Q\{t'\} = \int_V W_C(t')dV - \int_{S_U} \mathbf{T}'\cdot\mathbf{u}^G dS + \int_C p'g dS$$

for stresses in equilibrium with the given body force and surface tractions and which involve a pressure p' in the contact area C satisfying the loading conditions (16). We can then show that

$$Q\{t'\} - Q\{t\} = \int_V W_C(\Delta t)dV - \int_C (p'-p)(u_3 - g)dS.$$

Because p' and p apply the same resultant force and moment over C, the integral over C can be written as

$$\int_C (p'-p)(u_3 - g + d - ax_2 + bx_1)dS.$$

Equality holds in (17) where $p > 0$ so that there is no contribution to the integral from p, and the contribution from p' is nonpositive in view of (17). It follows that: *The complementary energy is least for the stresses of the actual solution.*

For the actual solution we have $P\{u\} = -Q\{t\}$ for both punch problems of this section.

Smooth Contact With an Elastic Foundation

Variational principles also hold for elastic bodies which can receive support from a foundation of the Winkler type. For simplicity we assume here that the foundation has a plane surface and lies below the plane $x_3 = 0$. The portion of the surface of an elastic body that can come into contact with the foundation is the nearly plane surface

$$x_3 = f(x_1, x_2)$$

touching the plane $x_3 = 0$ but lying entirely above it, where x_1, x_2 lie in a region C of the x_1-x_2 plane. The reactive

pressure p of the foundation is proportional to the downward displacement of the surface, so that on C we require

$$p=0 \text{ when } u_3 \geq -f, p=-K(u_3+f) \text{ when } u_3 < -f, \quad (20)$$

where $K \geq 0$ is the stiffness of the foundation (K may vary with x_1, x_2). If we define $q(u_3)$ by

$$q(u_3)=0 \text{ when } u_3 \geq -f, q(u_3)=1 \text{ when } u_3 < -f,$$

then we can write

$$T_3=p=-K(u_3+f)\ q(u_3) \text{ on } C, \quad (21)$$

while the tangential tractions are zero on C. Boundary conditions on the remaining surface S of the body and a body force field are prescribed as before.

The potential energy is defined to be

$$P\{u'\}=\int_V W(e')dV-\int_{S_T}\mathbf{T}^G\cdot\mathbf{u}'dS$$

$$-\int_V \mathbf{F}\cdot\mathbf{u}'dV+\frac{1}{2}\int_C K(u_3'+f)^2q'dS,$$

where $q'=q(u_3')$, for displacement fields u_3' which satisfy the displacement boundary conditions on S_U. Using equation (5) and the divergence theorem, we find that

$$P\{u'\}-P\{u\}=\int_V W(\Delta e)dV$$

$$+\frac{1}{2}\int_C K(v'^2q'+v^2q-2vqv')dS \quad (22)$$

in which

$$v=u_3+f, \quad v'=u_3'+f.$$

By considering the various possibilities for the signs of v and v' together with the corresponding values for the step functions q and q', we find that the integrand of the integral over C in equation (22) is nonnegative. Thus we again have: *The potential energy is an absolute minimum for the displacements of the actual solution.*

For the complementary energy we take

$$Q\{t'\}=\int_V W_C(t')dV-\int_{S_U}\mathbf{T}'\cdot\mathbf{u}^G dS$$

$$+\int_C (p'^2/2K+p'f)dS$$

for stresses in equilibrium with the given body force and surface tractions and which involve a pressure p' on the foundation interface C. Using equation (7) and the divergence theorem, we can show that

$$Q\{t'\}-Q\{t\}=\int_V W_C(\Delta t)dV$$

$$+\int_C \{(p'^2-p^2)/2K+(p'-p)(u_3+f)\}dS.$$

From equations (20), the integrand of the integral over C is found to be nonnegative and we see that: *The complementary energy is an absolute minimum for the stresses of the actual solution.*

We can also show that $P\{u\} = -Q\{t\}$, so that the principles can be used to bound $P\{u\}$.

Acknowledgment

This work was supported in part by the Office of Naval Research through the National Center for Composite Materials Research, College of Engineering, University of Illinois at Urbana-Champaign.

References

Gladwell, G. L. M., 1980, *Contact Problems in the Classical Theory of Elasticity*, Sijthoff and Noordhoff, Alpen aan den Rijn.

Shield, R. T., 1982, "Uniqueness for Elastic Crack and Punch Problems," ASME Journal of Applied Mechanics, Vol. 49, pp. 516–518.

Sokolnikoff, I. S., 1956, *Mathematical Theory of Elasticity*, Second Ed., McGraw-Hill, New York.

Sufficient Symmetry Conditions for Isotropy of the Elastic Moduli Tensor

R. M. Christensen
Lawrence Livermore National Laboratory,
Livermore, CA 94550
Fellow ASME

Symmetry conditions are found that assure isotropy of the fourth rank tensor of elastic moduli. Crystallography provides the answer to this problem in the two-dimensional context, namely one axis of three-fold symmetry assures the isotropy of properties in the plane normal to the axis. The present work provides the answer in the three-dimensional problem: 6 axes of five-fold symmetry are sufficient to give isotropy of the elastic moduli. An important restriction must accompany the present result. The derivation is given in the special form appropriate to low density materials which have a microstructure that transmits load according to the axial deformation of a space network of material distributed into micro-struts. The corresponding fiber composite idealization is that of a fiber dominated system, it therefore follows that if the fibers take the 6 specific orientations in three-space then isotropy is obtained.

Introduction

One of the basic questions in materials science concerns the determination of the symmetry conditions that assure isotropy. Of course isotropy has meaning only when specified relative to some particular property. For example, thermal conductivity is isotropic for the case of cubic symmetry. However, elastic moduli are not isotropic for cubic symmetry. The crucial distinction in this particular example is that heat conduction is characterized by a second rank tensor while elastic moduli are of a fourth rank tensor. In general, the number of independent constants associated with a particular property are directly determined by and known for the various types of symmetry encountered in the field of crystallography. However, none of the 32 crystal classes has symmetry sufficient to assure isotropy of the tensor of elastic moduli. This important property is the subject of the present work and sufficient conditions on material symmetry will be determined in order to assure isotropy of elastic moduli.

The simplest of the crystal classes is that involving cubic symmetry. This type of symmetry involves 4 axes of three-fold symmetry and it is sufficient to give isotropy of any second rank property tensor, thus involving only a single property (see for example Nye, 1957). The fourth rank elasticity tensor for cubic symmetry involves three independent constants, one more than is necessary for isotropy. Thus the symmetry type to give isotropy of the elastic moduli or compliances must be of higher order than that of cubic symmetry. In crystollography treatments, isotropy of moduli is introduced by merely stating it without defining a minimum symmetry class. From a group theory point of view this corresponds to saying that the symmetry class has an infinite number of elements. It is conceivable that this is the only possible solution to the problem of associating symmetry conditions with isotropy of moduli, however, this can easily be reasoned to not be likely in the following sense. Since 4 axes of three-fold symmetry assures isotropy of second rank tensors, it is likely that there is some higher order symmetry type, but with a finite number of elements, that assures isotropy of fourth rank property tensors. This will be found to be the case.

Another piece of evidence supports the contention that a rather simple symmetry class should assure isotropy of moduli. First, some terminology must be standardized. Henceforth the term isotropy will be taken to mean isotropy with respect to the tensor of elastic properties. Also the term isotropy means isotropy with respect to three-dimensional Euclidean space. The latter terminology immediately raises the question of corresponding behavior in two-space. That is, what are the conditions that give isotropy in two-space? The answer to this question has been known for many years and was first determined in the context of crystallography. Specifically, close-packed hexagonal symmetry implies elastic properties characterization through the 5 constants of transverse isotropy. In the plane of the hexagonal symmetry, the material is isotropic. Thus, in a planar, two-dimensional context isotropy is assured by close pack hexagonal symmetry which is mathematically characterized by 1 axis of six-fold symmetry. So then 1 axis of six-fold symmetry gives isotropy in two-space; although it is not obvious, so too does three-fold symmetry. It is reasonable to expect that some rather simple symmetry form would also give isotropy in three-space. Incidentally, the planar use of six-fold symmetry to give isotropy in two dimensions is widely employed in composite material laminate construction whereby lamina stacked at 60 deg angles with respect to each other assure "quasi-isotropy," meaning isotropy in the plane.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 24, 1986.

Paper No. 87-WA/APM-18.

The present work will prove that 6 axes of five-fold symmetry are sufficient to provide isotropy. However, some qualifications must be placed upon the derivation. The present interest in isotropy arose from project work upon very low density materials (see Gibson and Ashby, 1982, for a discussion of the many features of low density material microstructure). In some low density materials with appropriate microstructure, load can be transferred through a network of uniaxially load-bearing struts in the case of an open celled material. The material micro-structure involves various orientations of these struts. Previous theoretical work (Christensen, 1986) on these types of materials involved the assumption of conditions of isotropy through the device of letting the micro-material load-bearing members assume a random orientation in three-space. This then avoided the question of determining the minimum number of spatial orientations needed for isotropy of these types of materials. In the following work the manner of proving that 6 axes of five-fold symmetry gives isotropy is restricted to the low density material type involving load transfer through uniaxial deformation of micro-material members.[1] This restriction corresponds to other common idealizations. For example, the present results would also apply to fiber composite materials with the fibers oriented in three-space with the corresponding restriction being that the fiber-matrix system is fiber dominated. The present results also apply to a material whose stiffness is modeled by atomic interaction, but subject to the proviso that these interactions are only of the central force type. Primarily though, the present proof is effected through use of rigorous results from the analysis of low density material mechanical behavior. It is probable, but not proved here, that the solution to the stated isotropy problem, involving 6 axes of five-fold symmetry, applies to more general material types as well.

Finally one last clarification should be mentioned. These symmetry characterizations should not be confused with features of cell architecture in low density materials. Cell structure and type is a completely different area from questions of material symmetry, even though such cell structures do admit symmetry characterizations. To say this another way, symmetry types can be specified with no restriction or implication to the cell geometry. In fact, a cell type microstructure need not even exist.

Conceptually, the manner of proving that 6 axes of five-fold symmetry gives isotropy involves showing that for this type of symmetry the effective modulus of deformation with respect to a particular direction of uniaxial strain is in fact independent of direction. Before getting to this proof, it is necessary to obtain some preliminary results by examining the vastly simpler two-dimensional problem, and then by examining cubic symmetry in the three-dimensional case. The starting point for the present work is to recall some results from the analysis of isotropic, very low density materials since the proof relies upon special forms of tensor transformations with respect to these types of materials.

Low Density, Open Cell Material Isotropic Properties

The stress-strain relations for linear, isotropic elasticity have the common form

$$\sigma_{ij} = \lambda \delta_{ij} \epsilon_{kk} + 2\mu \epsilon_{ij} \tag{1}$$

where λ and μ are the Lamé constants and the Cartesian tensor notation is used. In terms of the alternate two mechanical properties involving the bulk modulus, k, and μ, then equation (1) has the alternate form

$$\sigma_{ij} = k \delta_{ij} \epsilon_{kk} + 2\mu \left(\epsilon_{ij} - \frac{\delta_{ij} \epsilon_{kk}}{3} \right) \tag{2}$$

[1] Budiansky and Kimmel (1986) have demonstrated isotropy for this type of symmetry in the context of modeling biological materials. See also the acknowledgment.

In the context of low density materials, the problem is to determine the two effective isotropic properties, μ and k, in terms of the amount of material in the space network of struts and the elastic modulus, E_m, of the strut material itself. This problem was first solved by Gent and Thomas (1963) in the context of cellular materials and by Christensen and Waals (1976) in the context of fiber reinforced materials having randomly oriented fibers. The final results are

$$\mu = \frac{cE_m}{15}$$

and (3)

$$k = \frac{cE_m}{9}$$

where c is the volume fraction of material, of modulus E_m. These are completely rigorous results from elasticity theory subject only to the idealization of load transfer by uniaxial deformation of the micro-structure. (See Christensen, 1986, for elaboration upon this condition.) The other isotropic properties corresponding to equations (3) are given by

$$\lambda = \frac{cE_m}{15}$$

$$E = \frac{cE_m}{6} \tag{4}$$

$$\nu = \frac{1}{4}$$

The value of Poisson's ratio of 1/4 corresponds to the original one constant elasticity theory of Poisson and Cauchy whereby atoms were viewed as material points subject to central force attraction or repulsion. Obviously this corresponds to a macroscopic idealization of load transfer through a network of struts.

It will be of value to have the proper form of the stress strain relation (1) for uniaxial strain, as specialized to low density, open cell materials. Combining equations (1), (3), and (4) gives

$$\sigma_{11} = \eta \epsilon_{11} \tag{5}$$

where

$$\eta = \lambda + 2\mu = \frac{cE_m}{5} \tag{6}$$

with ϵ_{11} being the only nonzero component of strain. The symbol η will be used throughout as the modulus of one-dimensional applied strain, i.e., uniaxial strain.

It also will be useful to have forms corresponding to those just given but specified for two-dimensional conditions of plane stress. The appropriate stress-strain forms are

$$\sigma_{11} = \frac{E}{1-\nu^2}(\epsilon_{11} + \nu \epsilon_{22})$$

$$\sigma_{22} = \frac{E}{1-\nu^2}(\epsilon_{22} + \nu \epsilon_{11}) \tag{7}$$

$$\sigma_{12} = 2\mu \epsilon_{12}$$

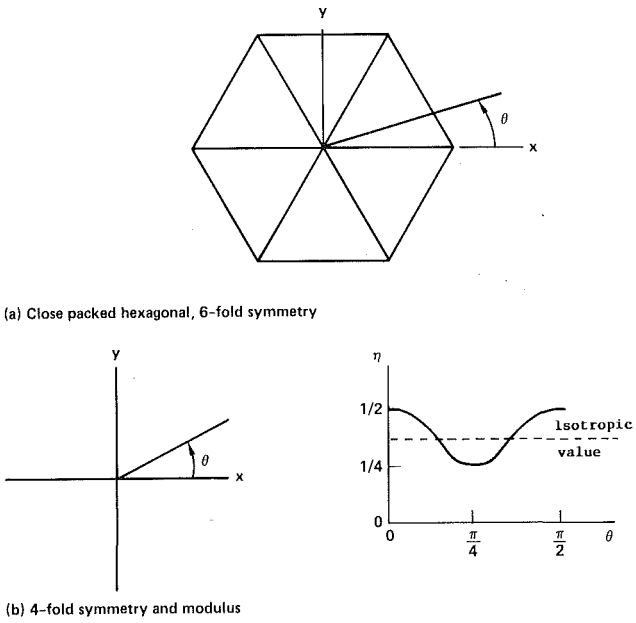
where from Christensen and Waals (1976) the appropriate properties are given by

$$E = \frac{cE_m}{3}$$

$$\nu = \frac{1}{3} \tag{8}$$

and

$$\mu = \frac{cE_m}{8}$$



(a) Close packed hexagonal, 6-fold symmetry

(b) 4-fold symmetry and modulus

Fig. 1 Two-dimensional forms

For one-dimensional strain, in the context of plane stress, relations (7) and (8) give

$$\sigma_{11} = \frac{3}{8} cE_m \epsilon_{11} \tag{9}$$

where $\epsilon_{22} = \epsilon_{12} = 0$.

The factors 3/8 in equation (9) and 1/5 in equation (6) represent the geometric effect of having the load bearing micro-structure members be randomly oriented in two-space and in three-space, respectively. These results will be applied in the next sections to determine conditions on symmetry to give isotropy.

Two-Dimensional Symmetry Requirements

It will be instructive to have displayed the relationship between isotropy and symmetry in the two-dimensional case before proceeding to the three-dimensional situation. As has already been mentioned, the two-dimensional case is fully understood. Following standard crystallography forms, as in Nye (1957), close packed, hexagonal symmetry implies transversely isotropic properties, which in the plane of the hexagonal symmetry are isotropic. Relative to an axis normal to this plane, the symmetry type is specified as one of six-fold symmetry.

Figure 1(*a*) shows the six-fold symmetry form with 3 axes at 120 deg angular intervals. Now stiffness is a fourth rank tensor with the tensor transformation

$$G_{mnpq} = G_{ijk\ell} \ell_{mi} \ell_{nj} \ell_{pk} \ell_{q\ell} \tag{10}$$

where ℓ_{mi} are the direction cosines. For the present interest in one-dimensional load carrying members, as the struts in the micro-scale view of low density materials, only one term in $G_{ijk\ell}$ is nonzero, thus

$$\sigma_{ij} = G_{ijk\ell} \epsilon_{k\ell} \tag{11}$$

where

$$G_{ijk\ell} = E_m A \text{ for } i, j, k, \ell = 1$$
$$= 0 \text{ otherwise} \tag{12}$$

with axis 1 taken in the direction of the member, and A the cross-sectional area of the member. With this restriction of $G_{ijk\ell}$ the tensor transformation assures an especially simple form involving trigonometric functions to the fourth power.

With respect to a one-dimensional strain state, as in the preceding section, the appropriate characterization of stiffness in the direction θ in Fig. 1(*a*) is given by

$$\eta = \frac{1}{3}\cos^4\theta + \frac{1}{3}\cos^4\left(\frac{\pi}{3} - \theta\right) + \frac{1}{3}\cos^4\left(\frac{2\pi}{3} - \theta\right) \tag{13}$$

where each of the three terms is the appropriately tensor transformed value for the three material load bearing directions in Fig. 1(*a*), and the coefficients of 1/3 are normalizing terms such that if the three directions were coincident the coefficient would be unity, with η non-dimensionalized. Expanding the last two terms in equation (13) and using trigonometric identities reduces (13) to a form independent of θ, proving isotropy, and giving simply

$$\eta = \frac{3}{8} \tag{14}$$

Comparing the result (14) with the known isotropic result from the previous section, namely equation (9), shows coincidence, as it must. Note that in equations (13) and (14) the term cE_m appearing in equation (9) has been absorbed into η. The simple proof of two-dimensional isotropy just given appeals to the especially simple case of uniaxial load transfer specified by equation (12). However, this latter restriction is not necessary, because in crystallography it is well known that three-fold or six-fold symmetry gives two-dimensional isotropy, with no further restriction.

Next, the two-dimensionally anisotropic case of Fig. 1(*b*) is treated, since it will reveal some features which reappear in the three-dimensional case. For two orthogonal material direction axes the symmetry is four-fold and the one-dimensional strain modulus is given by

$$\eta = \frac{1}{2}\cos^4\theta + \frac{1}{2}\cos^4(90 + \theta) \tag{15}$$

where again the coefficients of 1/2 are normalizing terms. Expanding the last term in equation (15) gives

$$\eta = \frac{1}{2}\cos^4\theta + \frac{1}{2}\sin^4\theta \tag{16}$$

The dependence upon angle θ precludes isotropy. It is obvious that the maximum and minimum values of η are given by

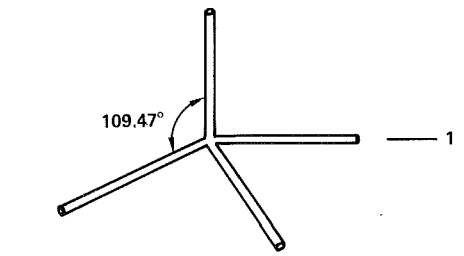
$$\eta_{\max} = \eta_{\theta=0} = \frac{1}{2}$$
$$\eta_{\min} = \eta_{\theta=\frac{\pi}{4}} = \frac{1}{4} \tag{17}$$

These maximum and minimum values of modulus under uniaxial strain conditions bracket the isotropic value of 3/8, as shown in Fig. 1(*b*). The maximum modulus direction is seen to be coincident with one of the material axes while the minimum modulus direction is, in an angular sense, equidistant from the two directions of material structure. These two-dimensional results are almost trivial, nevertheless they will provide guidance in the three-dimensional case where the manner for proceeding is far from trivial.

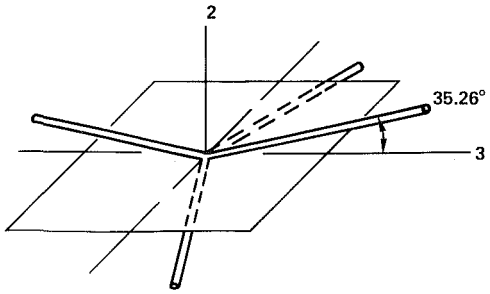
Three-Dimensional Symmetry Requirements

Two cases of three-dimensional symmetry will be studied, 4 axes of three-fold symmetry and 6 axes of five-fold symmetry. The latter case is the one to be proven to provide isotropy. The former case, that of cubic symmetry, is studied first since it is much simpler and it will be helpful to see its features of anisotropy.

4 Axes of Three-Fold Symmetry. The arrangement of 4



(a) Tetragonal axes arrangement



(b) Properties directions

Fig. 2 Cubic symmetry

axes of three-fold symmetry is obtained from the standard geometric construction of a tetrahedron or cube. The system of cubic symmetry is as shown in Fig. 2 where each axis makes an angle of 109.47 deg with each other axis. In this arrangement the material is distributed into four load bearing directions, having the stated symmetry.

The modulus due to one-dimensional strain in the direction of any one axis, as in Fig. 2(*a*), is given by the nondimensional form

$$\eta_1 = \frac{1}{4}(1) + \frac{1}{4}(3)\cos^4 109.47\text{deg} \tag{18}$$

where the 1/4 factors refer to a normalizing coefficient for four directions of symmetry, and each of the four directions transforms according to the special uniaxial forms in equations (11) and (12). Equation (18) becomes

$$\eta_1 = 0.259 \tag{19}$$

In Fig. 2(*b*) the plane is taken such that it makes equal angles with all four axes. Axis 2 is normal to the plane and also makes the same angle with all four axes. In direction 2 the one-dimensional strain modulus is given by

$$\eta_2 = \frac{(4)}{4}\cos^4 54.735\text{deg} \tag{20}$$

where there are 4 directions with the 1/4 normalizing coefficient. Equation (20) becomes

$$\eta_2 = 0.111 \tag{21}$$

A third direction, labeled 3 in Fig. 2(*b*), gives the corresponding modulus as

$$\eta_3 = \frac{(2)}{4}\cos^4 35.265\text{deg} \tag{22}$$

becoming

$$\eta_3 = 0.222 \tag{23}$$

The following characteristics emerge from this exercise. It is no surprise that the modulus is direction dependent since the symmetry is cubic with three independent constants. As discussed by Love (1944) and Nye (1957), Neumann's Principle asserts that the property under consideration, in this case the one-dimensional strain modulus, η, must have the same or

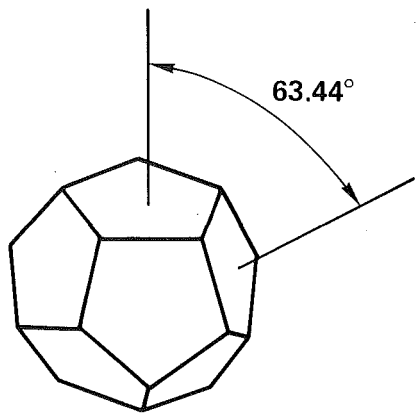


Fig. 3 Pentagonal dodecahedron

a higher degree of symmetry as that of the basic geometry. This then guarantees that the directional dependence of η will have 4 axes of three-fold symmetry and 3 axes of four-fold symmetry. That is, the shape of the surface for η as a function of direction will appear cubical in nature, having the form of a distorted cube. This then is sufficient to establish the extremum character for η in the directions of the material members and in the directions that make equal angles with them, that is in the directions of the diagonals of the cube and in the directions of the cubical axes, respectively. These extrema could either be maxima or minima. The results (19) and (21), however, show that

$$\eta_{\max} = n_1 = 0.259$$

and (24)

$$\eta_{\min} = \eta_2 = 0.110$$

Thus this type of symmetry is quite anisotropic with the ratio of maximum to minimum one-dimensional strain modulus being a little over a factor of 2, (actually 7/3), and bracketing the isotropic value of 1/5. Furthermore, it is seen that the direction of maximum modulus is coincident with one of the material direction axes while the direction of minimum modulus makes equal angles with all four axes. This behavior is the same as was found in the preceding two-dimensional cases, and this characteristic is of importance in the next section.

6 Axes of 5-Fold Symmetry. With the background provided in the previous sections there is now enough information to proceed to the main case, the symmetry condition that assures isotropy in three-space. As already examined, cubic symmetry involves 4 axes of three-fold symmetry. Perhaps the next case to consider would be 5 axes of four-fold symmetry, however such a condition does not exist. Proceeding further suggests 6 axes of five-fold symmetry. This is a standard form and it is associated with two types of regular solids. These are the pentagonal dodecahedron and the icosahedron. The former involves twelve faces of pentagonal form while the later is made up of 20 faces of equilateral triangles. The pentagonal dodecahedron has 20 vertices while the icosahedron has 12 vertices. In the case of the pentagonal dodecahedron, 6 axes of five-fold symmetry are established by erecting normals to the faces at their centers, while in the case of the icosahedron the geometric center of the solid is connected to the vertices. In the following work the geometric features of 6 axes of five-fold symmetry will be taken from the corresponding geometric characteristics of a pentagonal dodecahedron.

In the preceding case of cubic symmetry, it was sufficient to use characteristic angles with a specified number of significant figures and perform the moduli calculations at that level of ac-

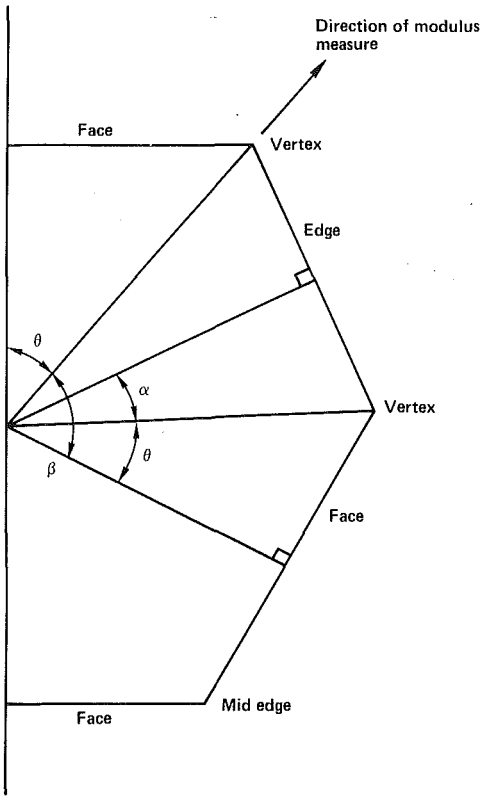


Fig. 4 Directions relative to a vertex of a pentagonal dodecahedron

curacy to demonstrate anisotropy. In the present case to prove isotropy it is not permissible to do numerical calculations to some specified degree of accuracy. Rather, only closed-form analytical expressions can be used to obtain the proof. The obvious way to attempt the proof would be to express the property of interest as a function of angular orientation in a spherical coordinate system. With the effects of all six axes combined, an independence of angular orientation for the property would prove isotropy. While this simple procedure was followed in the two-dimensional case given earlier, it would be extremely complicated to do so here in the three-dimensional case. Fortunately there is a much more direct path to the proof of isotropy, to be described next.

A pentagonal dodecahedron is shown schematically in Fig. 3. As already mentioned the 6 axes of five-fold symmetry are established by the normals to each face at their centers. Thus in the present problem the one-dimensional material members are distributed into these six directions. In terms of angular orientation, it is seen that the direction coincident with one axis is an extremum orientation insofar as modulus is concerned. This follows from Neumann's principle which was just discussed in connection with the cubic symmetry case. It is not clear whether it is a maximum or a minimum. Likewise a direction taken through a vertex is an extremum orientation because of the following symmetry pattern. Relative to a direction through a vertex, three of the symmetry axes have an equal angle with it, while the other three axes also have an equal (greater or lesser) angle with it. There is sufficient symmetry in this to assure an extremum characteristic. Again it is not known whether it would constitute a maximum or a minimum. Therefore, what can be said is that the direction coincident with one material axis and the direction that makes equal angles with the material axes constitute extremum orientations with respect to modulus, one being a maximum, the other a minimum. The preceding work on cubic symmetry and two-dimensional orthotropy suggests that the direction coincident with one axis provides the maximum modulus while the other direction provides the minimum modulus. The proof of isotropy, to be given next, involves showing that the one-dimensional strain moduli in these two directions are identically the same, and coincide with the known isotropic value, and since these are extremum orientations, the fourth-rank tensor of moduli must be isotropic.

Begin by considering the stiffness in the direction of a vertex in the dodecahedron of Fig. 3. The appropriate geometry needed to determine the direction cosines for all six axes are as shown in Fig. 4. The appropriate nondimensional modulus for uniaxial strain is given by

$$\eta = 3\left(\frac{1}{6}\right)\cos^4\theta + 3\left(\frac{1}{6}\right)\cos^4\beta \tag{25}$$

where the three nearest material directions are of angle θ from the vertex while the three furthest are of angle β. The factor 1/6 in equation (25) accounts for the presence of 6 axes of symmetry, and the load bearing element of each direction has stiffness characteristics given by equations (11) and (12).

The basic geometry of the pentagonal dodecahedron is taken from the International Tables of X-ray Crystallography (1959). From this source

$$\cos^4\theta = \frac{\tau^6}{45} \tag{26}$$

where

$$\tau = \frac{1}{2}(1+\sqrt{5}) \tag{27}$$

Angle β in equation (25) and Fig. 4, obtained by passing a plane through the pentagonal dodecahedron, can be written as

$$\beta = \theta + 2\alpha \tag{28}$$

but

$$\cos\alpha = \frac{\tau}{\sqrt{3}} \tag{29}$$

and

$$\cos\theta = \frac{\tau^{3/2}}{\sqrt{3}(5)^{1/4}}$$

Combining equations (28) and (29) gives

$$\cos\beta = \left(\frac{2\tau^3}{3} - 1\right)\frac{\tau^{3/2}}{\sqrt{3}(5)^{1/4}} - \frac{2}{3}\left(1 - \frac{\tau^3}{3\sqrt{5}}\right) \tag{30}$$

Using τ from equation (27) after some manipulation then gives the intermediate form

$$\cos\beta = \frac{\sqrt{5}-1}{2\sqrt{3}(5)^{1/4}\sqrt{\tau}} \tag{31}$$

where the identity $1/\tau = \tau - 1$ has been used. Using τ again and raising to the fourth power gives equation (31) as

$$\cos^4\beta = \frac{7-3\sqrt{5}}{45(3+\sqrt{5})} \tag{32}$$

Now, using equation (27) in (26) gives that term as

$$\cos^4\theta = \frac{1}{45}(9+4\sqrt{5}) \tag{33}$$

Combining equations (32) and (33) in (25) gives the final result for the modulus due to uniaxial strain in the direction of a vertex as

$$\eta\Big|_{\substack{\text{Vertex}\\ \text{Direction}}} = \frac{1}{5} \tag{34}$$

Next the modulus due to uniaxial strain in the direction of an axis of symmetry is determined. The appropriate expression for the non-dimensional modulus is

$$\eta = \frac{1}{6} + 5\left(\frac{1}{6}\right)\cos^4 2\gamma \tag{35}$$

where 2γ is the angle between any two axes of symmetry. Using the pentagonal dodecahedron, the direction cosine between the normal to the center of a face, and the direction from the center of the dodecahedron to the adjacent mid edge location, is given by

$$\cos\gamma = \frac{\sqrt{\tau}}{(5)^{1/4}} \tag{36}$$

Using equation (36) in (35) then gives

$$\eta\Big|_{\substack{\text{symm}\\ \text{axis}}} = \frac{1}{5} \tag{37}$$

From equations (34) and (37) the two extremum values for the modulus coincide. Furthermore, they are coincident with the known isotropic value from equation (6) thus the present case has been established as being isotropic. To state this explicitly, the condition of 6 axes of five-fold symmetry is sufficient to provide isotropy of the modulus tensor.

This proof of isotropy is related to material directions in low density materials microstructure. Similar considerations apply to fiber reinforced composite materials under the proviso that the material is fiber dominated, meaning that the matrix contribution to the stiffness is negligible compared with the fiber contribution. The fibers arranged according to the 6 directions specified above will then provide isotropy. Thus, 3 directions suffice for isotropy in two-dimensional continua while 6 directions are required for three-dimensional continua.

The present proof is restricted to the cases discussed involving load transfer according to the simplified stiffness matrix of equations (11) and (12). In that special case only a single element of the stiffness matrix was taken to be nonzero. This greatly simplified the number of terms to be considered in the tensor transformation relation. Nevertheless, these special cases probably cover the most important physical examples of materials whose microstructure exhibits special forms of directionality, i.e., low density, open cell materials, and fiber reinforced materials. It is likely that the present results can be generalized to the case where the individual material element is transversely isotropic rather than just being the special case considered here of transversely isotropic but with only one nonzero entry in the modulus matrix. Such a proof, however, would require retaining many more terms in the tensor transformation relations, and this may not be of any special physical relevance.

Interestingly this same five-fold symmetry has been detected by X-ray diffraction in full density materials. These are called quasi-crystalline metals, and they do not fit into any of the conventional crystal classes (see Rawls, 1986, for a recent overview of developments in this field). With regard to the low density materials of primary interest in the present work, reference can be made to Christensen (1986) for experimental data demonstrating the existence of these efficient materials.

Acknowledgment

I am appreciative to Professor B. Budiansky who has shown me a completely different, concise proof of isotropy in recent, private correspondence. He informs me that his proof will be published in forthcoming work.

I am also appreciative to Dr. R. W. Hopper for many helpful discussions on this problem, especially in the case of cubic symmetry. Also, thanks are due to Dr. G. S. Smith for help in locating the International Tables for X-ray Crystallography, and to Dr. J. F. Carley for reading the manuscript and offering clarifying comments.

This work performed under the auspices of the US Department of Energy by the Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

References

Budiansky, B., and Kimmel, E., 1986, "Elastic Moduli of Lungs," Harvard University, MECH-78.

Christensen, R. M., and Waals, F. M., 1976, "Effective Stiffness of Randomly Oriented Fiber Composites," *Journal of Composite Materials*, Vol. 12, pp. 518–532.

Christensen, R. M., 1986, "Mechanics of Low Density Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 34, pp. 563–578.

Gent, A. N., and Thomas, A. G., 1963, "Mechanics of Foamed Elastic Materials," *Rubber Chemistry and Technology*, Vol. 36, pp. 597–610.

Gibson, L. J., and Ashby, M. F., 1982, "The Mechanics of Three-Dimensional Cellular Materials," *Proceedings of the Royal Society of London*, Vol. A382, pp. 43–59.

International Tables for X-Ray Crystallography, 1959, Vol. II, Mathematical Tables, Kynoch Press, Birmingham, England.

Love, A. E. H., 1944, *The Mathematical Theory of Elasticity*, 4th Ed., Dover, New York.

Nye, J. F., 1957, *Physical Properties of Crystals*, Oxford University Press, Oxford.

Rawls, R. L., 1986, "Concept of Quasicrystalline Metal Alloys Becoming Clearer," *Chemical and Engineering News*, June 23, pp. 35–36.

Torsion of Cylinders With Shape Intrinsic Orthotropy

S. C. Cowin
Alden J. Laborde Professor of Engineering,
Department of Biomedical Engineering,
Tulane University,
New Orleans, LA 70118
Fellow, ASME.

Shape intrinsic orthotropy may be thought of as the type of elastic material symmetry possessed by the wood tissue of a tree. Each year's new growth rings form a laminate around a central core. The axes of material symmetry lie in the directions tangent and normal to the growth rings or laminates and along the axis of the cylinder. Let G_{tz} denote the linear elastic orthotropic shear modulus associated with the axial and tangential directions, the tangent plane of a laminate. It is shown here that, for a certain class of elastic cylinders with shape intrinsic orthotropy, the solution to the torsion problem is the same as the solution to the torsion problem for the isotropic cylinder of the same shape if the isotropic shear modulus G were replaced by the orthotropic shear modulus G_{tz}.

Introduction

Orthotropic symmetry is characterized by three mutually perpendicular planes of mirror symmetry. The normals to these three planes form a symmetry coordinate system for orthotropic symmetry and, relative to this coordinate system, there are only nine distinct orthotropic elastic constants for linear elasticity. One set of these nine constants consists of the technical elastic constants which include three Young's moduli E_1, E_2, and E_3, three shear moduli G_{12}, G_{13}, and G_{23}, and six Poisson's ratios, only three of which are independent. *Curvilinear orthotropy* is the term used to describe a material in which the orientation of the orthotropic symmetry coordinate system is different from point to point. Wood is generally considered to have curvilinear orthotropic elasticity, and the orthotropic symmetry coordinate system for wood is oriented so that one axis is coincident with the axis of the grain, one axis is tangent to the growth rings, and the third axis is perpendicular to the growth rings. The term *shape intrinsic orthotropy* is introduced here to describe the general situation when the symmetry coordinate system for orthotropic symmetry of a cylinder is coincident with the long axis of the cylinder, with the local tangent and normal to a closed family of curves, one of which forms the lateral boundary of the cylinder. The closed family of curves also delineate the laminates of the cylinder. The types of curvinlinear orthotropy known as circular (sometimes cylindrical) and elliptic orthotropy are special cases of shape intrinsic orthotropy and are associated with circular and elliptical cylinders, respectively. Shape intrinsic orthotropy is characteristic of cylinders that are formed in growth processes like plant and animal tissue or in industrial lamination processes. In the literature of elasticity curvilinear orthotropy is contrasted with *rectilinear orthotropy* in which the symmetry coordinate system for the material is Cartesian. For rectilinear orthotropy the symmetry coordinate system is selected so that one axis is coincident with the long axis of the cylinder and the other two axes lie in the plane of the cross section. The fact that the symmetry coordinate system is Cartesian means that the system at any point in the cylinder has axes that are parallel to the axes of the symmetry coordinate system at all other points.

The 327 page paper presented to the French Academy by St. Venant (1855) a century and a third ago remains the single most outstanding work on the problem of torsion of elastic cylinders. In that paper St. Venant reviews the basic equations of elasticity for both isotropic materials and rectilinearly orthotropic materials. He considers the problems of extension, contraction, flexure, and torsion of a cylinder of isotropic and rectilinearly orthotropic materials. In particular he solves the torsion problem for an isotropic cylinder of elliptic, of rectangular, and of equilateral triangular cross sections as well as of many cross sections described by higher order curves with which simple names are not associated. His figures illustrating his solutions have appeared in most subsequent texts on the theory of elasticity. For example, St. Venant's figure illustrating the warping of an elliptical cylinder due to torsion (1855, p. 339) is repeated by Love (1944, p. 320), Timoshenko and Goodier (1970, p. 298), and Sokolnikoff (1956, p. 123). Other warping illustrations derived from St. Venant (1855) appear in Love (1944, p. 321), Timoshenko and Goodier (1970, p. 301), and Sokolnikoff (1956, p. 125 and p. 133). St. Venant (1855) introduced a coordinate stretch transformation that reduced the problem of torsion of cylinders with rectilinear orthotropy to the problem of the torsion of an isotropic cylinder. Most subsequent work on elastic torsion, including the present contribution, could be considered as footnotes to the treatise of St. Venant (1855).

In this paper it is shown that St. Venant's method of solution of the isotropic elastic torsion problem for a particular cross section also, in many cases, solves the elastic torsion problem for a cylinder with shape intrinsic orthotropy of the same

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, 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, April 27, 1987.

Paper No. 88-WA/APM-3.

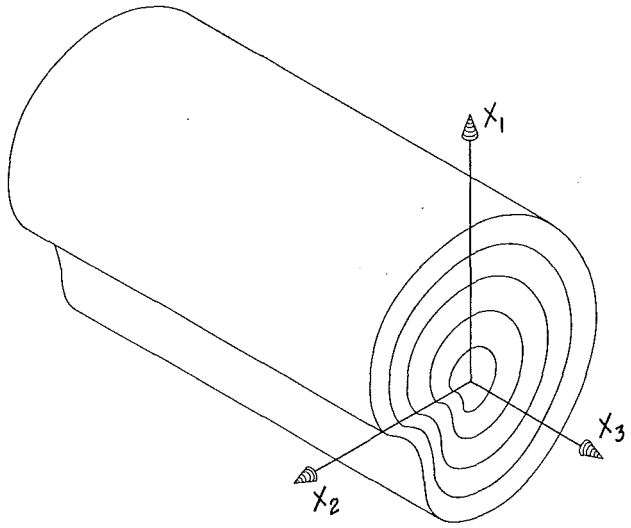


Fig. 1 An illustration of a cylinder of arbitrary cross section with shape intrinsic orthotropy and the rectangular coordinate system employed

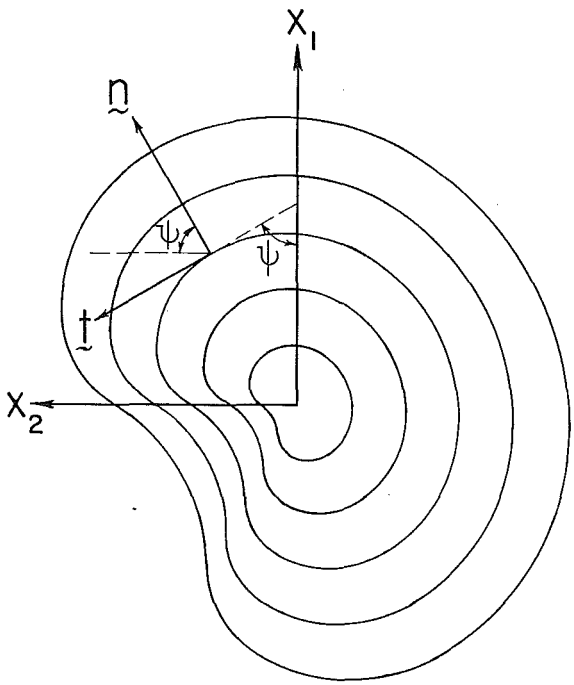


Fig. 2 The arbitrary cross section of the cylinder considered. The family of curves of which the lateral boundary is a member is illustrated. The local tangent and normal to this family form, along with the axis of the cylinder, the local symmetry coordinate system for shape intrinsic orthotropy.

cross section if the isotropic shear modulus G is replaced by the orthotropic shear modulus G_{tz}.

The torsion of a cylinder of arbitrary cross-sectional shape and with shape intrinsic orthotropy is illustrated with the Cartesian coordinate system x_1, x_2, x_3 in Fig. 1. The local coordinate system determined by the tangent and normal to one of the family of curves of which the lateral bounding curve is a member is shown in Fig. 2. The symmetry coordinate system for the shape intrinsic orthotropy consists of unit vectors in the direction of the local tangent and normal to the family of curves and a unit vector in the plane perpendicular to the plane of the family of curves. The two shear moduli of interest are denoted by G_{tz} and G_{nz} and are referred to this symmetry coordinate system.

In the following section St. Venant's solution of the torsion problem for an isotropic cylinder of arbitrary cross section is briefly summarized. In the subsequent section, a condition is obtained that determines the class of cylinders with shape intrinsic orthotropy for which the solution to the torsion problem is given by the solution to the torsion problem for an isotropic cylinder of the same cross-sectional shape, but with the isotropic shear modulus G replaced by the orthotropic shear modulus G_{tz}. In the next section, two specific examples of the class of cylinders are developed and, in the section following that, St. Venant's (1855) solution for the torsion of a cylinder with rectilinear orthotropy is reviewed and compared with the solutions obtained here. The final section contains a discussion of results.

The Torsion Problem for Isotropic Cylinders

St. Venant (1855) generalized the solution of Navier to the torsion problem for a right circular cylinder of isotropic material to the solution for a right cylinder of arbitrary cross section. St. Venant's assumption concerning the components u_1, u_2, u_3 of the displacement vector in an elastic cylinder of arbitrary cross-sectional shape can be expressed as

$$u_1 = -\alpha x_2 x_3,\ u_2 = \alpha x_1 x_3,\ u_3 = \alpha\phi(x_1, x_2), \tag{1}$$

relative to the Cartesian coordinate system illustrated in Fig. 1. The function $\phi(x_1, x_2)$ is called the warping function and α is a constant representing the twist per unit length of the cylinder. The only nonzero components of the strain tensor are determined from eqution (1) using the strain-displacement relations

$$E_{23} = \frac{\alpha}{2}\left(\frac{\partial\phi}{\partial x_2} + x_1\right),\ E_{13} = \frac{\alpha}{2}\left(\frac{\partial\phi}{\partial x_1} - x_2\right), \tag{2}$$

and the only nonzero components of the stress tensor are determined from equation (2) using the isotropic form of Hooke's law, thus

$$T_{23} = \alpha G\left(\frac{\partial\phi}{\partial x_2} + x_1\right),\ T_{13} = \alpha G\left(\frac{\partial\phi}{\partial x_1} - x_2\right). \tag{3}$$

The vanishing of all the components of the stress tensor except T_{23} and T_{13} satisfies all the stress equations of equilibrium except for the condition

$$\frac{\partial T_{23}}{\partial x_2} + \frac{\partial T_{13}}{\partial x_1} = 0. \tag{4}$$

This condition is satisfied if the warping function $\phi(x_1, x_2)$ is harmonic in the region of the cross section A,

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x_1^2} + \frac{\partial^2\phi}{\partial x_2^2} = 0 \text{ in } A, \tag{5}$$

as can be verified by substitution of equation (3) into (4). The lateral boundaries of the shaft are assumed to be free of surface tractions. This condition is satisfied if the shearing stresses T_{23} and T_{13} satisfy the condition

$$T_{23}\cos\psi + T_{13}\sin\psi = 0 \tag{6}$$

on the lateral boundary denoted by ∂A. It is convenient to define the quantity Λ by

$$\Lambda = T_{23}\cos\psi + T_{13}\sin\psi = \left(\frac{\partial\phi}{\partial x_1} - x_2\right)\sin\psi + \left(\frac{\partial\phi}{\partial x_2} + x_1\right)\cos\psi \tag{7}$$

where the second equality follows from the first and equation (3). The boundary condition (6) then takes the form

$$\Lambda = 0 \text{ on } \partial A. \tag{8}$$

The total torque T carried by the shaft is given by

$$T = \int_A (x_1 T_{23} - x_2 T_{13})\,dx_1\,dx_2, \tag{9}$$

and can be expressed by the simple formula

$$T = GJ\alpha \tag{10}$$

where

$$J = \int_A \left(x_1^2 + x_2^2 + x_1 \frac{\partial \phi}{\partial x_2} - x_2 \frac{\partial \phi}{\partial x_1} \right) dx_1 dx_2. \tag{11}$$

The method of St. Venant reduces the torsion problem for a right cylinder of arbitrary cross-sectional shape and of isotropic material to the determination of the warping function $\phi(x_1, x_2)$ that is harmonic in A and satisfies the boundary condition (8) on ∂A.

The Torsion Problem for a Class of Cylinders With Shape Intrinsic Orthotropy

The solution to the torsion problem in the case of shape intrinsic orthotropy will closely parallel the solution for the isotropic case; the only equations that will be different are those expressing Hooke's law. In the local tangent coordinate system in which the material is orthotropic, the two shear moduli of interest are G_{tz} and G_{nz}. The elastic coefficients C_{2323}, C_{1313}, and $C_{2313} = C_{1323}$ in the Cartesian coordinate system are obtained from G_{tz} and G_{nz} using the fourth rank Cartesian tensor transformation rule for rotation from the n, t, z system to the x_1, x_2, x_3 coordinate system, thus

$$\begin{aligned} C_{2323} &= \sin^2\psi G_{tz} + \cos^2\psi G_{nz}, \\ C_{1313} &= \cos^2\psi G_{tz} + \sin^2\psi G_{nz}, \\ C_{2313} &= \sin\psi \cos\psi (G_{nz} - G_{tz}). \end{aligned} \tag{12}$$

The only angle involved in this coordinate transformation is ψ, as can be seen from Fig. 2. The stresses T_{23} and T_{13} are related to the strains E_{23} and E_{13} in the Cartesian system by

$$\begin{aligned} T_{23} &= 2C_{2323}E_{23} + 2C_{2313}E_{13}, \\ T_{13} &= 2C_{2313}E_{23} + 2C_{1313}E_{13}. \end{aligned} \tag{13}$$

The assumption (1) of St. Venant is employed in the solution of the problem with shape intrinsic orthotropy. It follows that the strains E_{13} and E_{23} are then given by equation (2). Substitution of equations (2) and (12) into (13) and subsequent use of the definition (7) gives expressions for the stresses T_{23} and T_{13}:

$$\begin{aligned} T_{23} &= \alpha G_{tz}\left(\frac{\partial \phi}{\partial x_2} + x_1 \right) + \alpha \Lambda (G_{nz} - G_{tz}) \cos\psi, \\ T_{13} &= \alpha G_{tz}\left(\frac{\partial \phi}{\partial x_1} - x_2 \right) + \alpha \Lambda (G_{nz} - G_{tz}) \sin\psi. \end{aligned} \tag{14}$$

When these stresses are substituted into the boundary condition (6), condition (6) is again satisfied if relation (8) is satisfied. Observe that if the relation (8) is satisfied not only on ∂A, but also in A, then it can be used as an identity in the formulas (14) for the shearing stresses, and they reduce to the following simple expressions

$$T_{23} = \alpha G_{tz}\left(\frac{\partial \phi}{\partial x_2} + x_1 \right), \quad T_{13} = \alpha G_{tz}\left(\frac{\partial \phi}{\partial x_1} - x_2 \right). \tag{15}$$

These formulas for the shearing stresses are identical with those given by equation (3) for the isotropic case except that G_{tz} replaces G. Thus we can conclude that if $\phi(x_1, x_2)$ and the family of curves of which the lateral boundary is a member satisfy the relation (8) in A as well as on ∂A, then the solution for shape intrinsic orthotropy is the same as the solution for the isotropic case for a cylinder of the same shape if the isotropic shear modulus G is replaced by the orthotropic shear modulus G_{tz}. The class of cylinders with shape intrinsic orthotropy for which this is true are those for which the warping function $\phi(x_1, x_2)$ and the angle ψ satisfy the relation

$$\Lambda = \left(\frac{\partial \phi}{\partial x_1} - x_2 \right) \sin\psi + \left(\frac{\partial \phi}{\partial x_2} + x_1 \right) \cos\psi = 0 \tag{16}$$

everywhere in the cross section of the cylinder. Observing from Fig. 2 that

$$\tan\psi = -\frac{dx_2}{dx_1}, \tag{17}$$

the condition (15) may also be expressed in the form

$$\frac{\dfrac{\partial \phi}{\partial x_2} + x_1}{\dfrac{\partial \phi}{\partial x_1} - x_2} = \frac{dx_2}{dx_1}. \tag{18}$$

The left-hand side of this expression represents a property of the warping function $\phi(x_1, x_2)$ for A and the right-hand side represents a property of the family of bounding curves of which the lateral boundary, ∂A, is a member. Alternatively, from equations (2) and (3) it is easy to see how to express the result in terms of the shearing stresses or shearing strains.

The results presented above apply to the case of hollow cylinders for which the interior and exterior bounding curves of the lateral cross section are of the same family. It follows then that the formulas of Bredt (1896) for the torsion of thin-walled isotropic cylinders also apply to the torsion of thin-walled cylinders with shape intrinsic orthotropy if the isotropic shear modulus G is replaced by G_{tz}.

Examples of the Selected Class of Cylinders With Shape Intrinsic Orthotropy

In this section two examples are given of the class of cylinders with shape intrinsic orthotropy for which a solution to the torsion problem has been obtained. The examples are of a cylinder of elliptic cross section and a cylinder of general type considered by St. Venant (1855). For an elliptic cylinder made of an isotropic material the warping function $\phi(x_1, x_2)$ and family of ellipses of which the lateral boundary is a member are given by

$$\phi(x_1, x_2) = \frac{\xi^2 - 1}{\xi^2 + 1} x_1 x_2, \tag{19}$$

$$\left(\frac{x_1}{a} \right)^2 + \left(\frac{x_2}{\xi a} \right)^2 = 1, \quad 0 < \xi < 1, \tag{20}$$

respectively, where ξ is the ratio of minor axis length to the major axis length. The family of ellipses is obtained by variation of the major axis a while holding ξ fixed. Using equation (19) it is easy to see that

$$\frac{\dfrac{\partial \phi}{\partial x_2} + x_1}{\dfrac{\partial \phi}{\partial x_1} - x_2} = \frac{-\xi^2 x_1}{x_2}, \tag{21}$$

and from equation (20) it follows that

$$\frac{dx_2}{dx_1} = -\frac{\xi^2 x_1}{x_2}, \tag{22}$$

thus, from equation (18), the condition for this isotropic solution to be a solution for shape intrinsic orthotropy is satisfied. It follows that the nonzero strains and stresses in the elliptic cylinder with shape intrinsic orthotropy are given by

$$E_{23} = \frac{\xi^2 \alpha x_1}{1 + \xi^2}, \quad E_{13} = \frac{-\alpha x_2}{1 + \xi^2}, \tag{23}$$

and

$$T_{23} = 2G_{tz} \frac{\xi^2 \alpha x_1}{1 + \xi^2}, \quad T_{13} = -2G_{tz} \frac{\alpha x_2}{1 + \xi^2}, \tag{24}$$

respectively, and the total applied torque T by

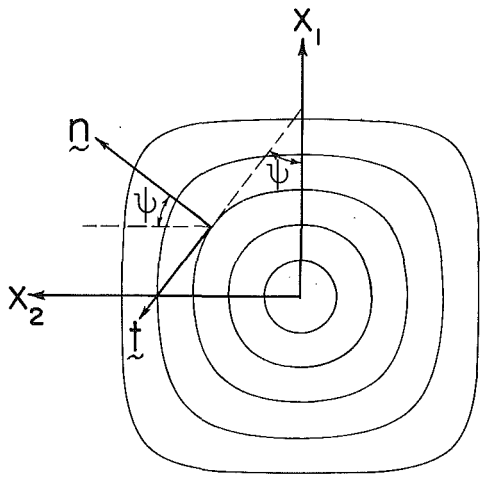


Fig. 3 The family of curves is given by equation (28), with $\xi = -0.1$. The outermost curve corresponds to $a = 1.5446$, the next ones to $a = 0.9045$, $a = 0.4438$, $a = 0.1348$, and $a = 0.0425$, respectively, as the closed curves tighten on the center of the shaft.

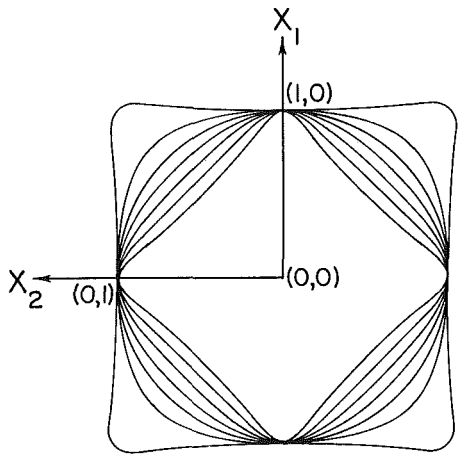


Fig. 4 The family of curves given by equation (28) with $a = 1$ and various values of ξ. From the outermost curve to the innermost, $\xi = -0.207$, $\xi = -0.1$, $\xi = 0.0$ (a circle), $\xi = 0.1$, $\xi = 0.2$, $\xi = 0.3$, $\xi = 0.4$, and $\xi = 0.5$. This figure is adapted from St. Venant (1855). In a similar plot given by Timoshenko and Goodier (1970, Fig. 152, p. 267), the curve corresponding to $\xi = -0.207$ above is mislabeled as $1/2\,(\sqrt{2}-1) = +0.207$.

$$T = \frac{\xi^3 \alpha}{1+\xi^2}\,\pi a^4 G_{tz}. \tag{25}$$

The shear stress T_{tz} referred to the symmetry coordinate system for shape intrinsic orthotropy is the only nonzero stress component in that system. It is related to T_{23} and T_{13} in the Cartesian system by

$$T_{23} = T_{tz}\sin\psi,\quad T_{13} = -T_{tz}\cos\psi, \tag{26}$$

and it follows from equations (24) and (26) that

$$T_{tz} = \frac{2\sqrt{x_2^2 + x_1^2\xi^4}}{1+\xi^2}\,G_{tz}\phi. \tag{27}$$

As a second example, a solution for an isotropic cylinder described by a fourth order curve given by St. Venant (1855) is considered. The warping function $\phi(x_1, x_2)$ and family of curves of which the lateral boundary is a member are given by

$$\phi = 2\xi x_1 x_2 (x_2^2 - x_1^2) \tag{28}$$

and

$$x_1^2 + x_2^2 - \xi(x_1^4 + x_2^4 - 6x_1^2x_2^2) = a - \xi, \tag{29}$$

respectively, where ξ is a parameter of curve shape, and different values of a correspond to different members of the family of curves. In order that the curves described by equation (29) be closed it is necessary to require that

$$\begin{aligned} &a > 0 \quad , \text{ if } \xi = 0, \\ &1 > 4(\xi - a)\xi > -1, \text{ if } \xi < 0, \text{ and} \\ &0 > 4(\xi - a)\xi > -1 \text{ if } \xi > 0. \end{aligned} \tag{30}$$

Fig. 3 represents a typical cross section of the shape given by equation (29) when $\xi = -0.1$. The curves for the same value of ξ and different values of a represent the everywhere-dense set of local curves which determine the direction of the local symmetry coordinate system. When ξ rather than a is varied in equation (29), different shaped curves are determined, as shown in Fig. 4. This figure shows that equation (29) represents cross-sectional shapes that vary from almost square to circular. Using equation (27) it is easy to show that

$$\frac{\dfrac{\partial\phi}{\partial x_2} + x_1}{\dfrac{\partial\phi}{\partial x_1} - x_2} = \frac{-(1 + 2\xi(3x_2^2 - x_1^2))x_1}{(1 + 2\xi(3x_1^2 - x_2^2))x_2}, \tag{31}$$

and from equation (29) it follows that

$$\frac{dx_2}{dx_1} = -\frac{(1 + 2\xi(3x_2^2 - x_1^2))x_1}{(1 + 2\xi(3x_1^2 - x_2^2))x_2}; \tag{32}$$

thus, from equation (18), the condition for this isotropic solution to a solution for shape intrinsic orthotropy is satisfied. It follows that the nonzero strains and stresses in this cylinder with shape intrinsic orthotropy are given by

$$\begin{aligned} E_{23} &= \alpha x_1(1 + 2\xi(3x_2^2 - x_1^2)), \\ E_{13} &= -\alpha x_2(1 + 2\xi(3x_1^2 - x_2^2)), \end{aligned} \tag{33}$$

and

$$\begin{aligned} T_{23} &= 2G_{tz}\alpha x_1(1 + 2\xi(3x_2^2 - x_1^2)), \\ T_{13} &= -2G_{tz}\alpha x_2(1 + 2\xi(3x_1^2 - x_2^2)), \end{aligned} \tag{34}$$

respectively.

St. Venant's Solution for an Elliptic Cylinder with Rectilinear Orthotropy

It is interesting to compare the results obtained here for an elliptic cylinder with shape intrinsic or elliptic orthotropy to the results obtained by St. Venant (1855) for an elliptic cylinder with rectilinear orthotropy. In the case of rectilinear orthotropy the shear moduli of significance are G_{13} and G_{23}. These are the orthotropic shear moduli in the rectilinear orthotropic symmetry coordinate system whose axes are coincident with the major and minor axes of the ellipse and the long axis of the shaft. This solution is described, for example, by Hearmon (1961) or Lekhnitskii (1963). In this case the warping function $\phi(x_1, x_2)$ is given by

$$\phi(x_1, x_2) = \frac{\xi^2 G_{13} - G_{23}}{\xi^2 G_{13} + G_{23}}\,x_1 x_2, \tag{35}$$

and the strains and stresses by

$$E_{23} = \frac{\xi^2 G_{13}\alpha x_1}{\xi^2 G_{13} + G_{23}},\quad E_{13} = \frac{-G_{23}\alpha x_2}{\xi^2 G_{13} + G_{23}}, \tag{36}$$

and

$$T_{23} = \frac{2\xi^2 G_{13}G_{23}\alpha x_1}{\xi^2 G_{13} + G_{23}},\quad T_{13} = \frac{-2G_{13}G_{23}\alpha x_2}{\xi^2 G_{13} + G_{23}}, \tag{37}$$

respectively. The torque T applied to the shaft is given by

$$T = \frac{G_{13}G_{23}\xi^3\pi a^4\alpha}{\xi^2 G_{13} + G_{23}}. \tag{38}$$

The results for the elliptic cylinders with rectilinear orthotropy

and shape intrinsic or elliptic orthotropy are easily compared. The warping functions for these two solutions are given by equations (35) and (19), the strains by equations (36) and (23), the stresses by equations (37) and (25), and the total torque by equations (38) and (25), respectively. The results recorded above reduce to the corresponding results for the elliptic shaft with shape intrinsic orthotropy if

$$G_{13} = G_{23} = G_{tz}. \tag{39}$$

Specifically the use of equation (39) reduces equation (35) to (19), equation (36) to (23), equation (37) to (24), and equation (38) to (25).

The Case When the Shear Modulus Varies

In this section the situation in which G_{tz} is a function of x_1 and x_2 is considered. The unit vectors $\mathbf{n}$ and $\mathbf{t}$ illustrated in Figs. 2 and 3 are represented in terms of ψ by

$$\mathbf{n} = \sin\psi\mathbf{e}_1 + \cos\psi\mathbf{e}_2, \mathbf{t} = -\cos\psi\mathbf{e}_1 + \sin\psi\mathbf{e}_2. \tag{40}$$

It is shown here that if G_{tz} does not vary in the tangential direction,

$$\mathbf{t}\cdot\nabla G_{tz} = 0, \tag{41}$$

then the solution to the torsion problem for shape intrinsic orthotropy given above is still applicable even if $\mathbf{n}\cdot\nabla G_{tz}$ is not zero. Thus G_{tz} can vary in the direction $\mathbf{n}$, or from laminate to laminate. To prove this result it is assumed that G_{tz} is a function of x_1 and x_2 and the stresses given by equation (15) are substituted into equation (4); then equation (4) is rewritten as

$$\alpha\cos\psi G_{tz}\nabla^2\phi + \alpha\Lambda\frac{\partial G_{tz}}{\partial x_2} + \alpha\left(\frac{\partial\phi}{\partial x_1} - x_2\right)\mathbf{t}\cdot\nabla G_{tz} = 0, \tag{42}$$

where the final representation was obtained by use of equations (40) and (7). The result (42) shows that if $\mathbf{t}\cdot\nabla G_{tz}$ and Λ are zero everywhere in the region, then equation (42) reduces to (5) and the solution to the torsion problem for shape intrinsic orthotropy is recovered even though $\mathbf{n}\cdot\nabla G_{tz}$ need no longer be zero.

In the remainder of this section the case when G_{tz} is a function of x_1 and x_2 satisfying the relation (41) is considered. Substitution of the stresses given by equation (15) into (9) yields the following formula for the total torque on the shaft:

$$T = \int_A \alpha G_{tz}(x_1, x_2)\left(x_1^2 + x_2^2 + x_1\frac{\partial\phi}{\partial x_2} - x_2\frac{\partial\phi}{\partial x_1}\right)dx_1\,dx_2 \tag{43}$$

In the case where the shaft is composed of N laminates with different G_{tz} shear moduli, the total torque is given by

$$T = \alpha\sum_{n=1}^{N} G_{tz}^n J^n \tag{44}$$

where G_{tz}^n is the shear modulus G_{tz} for the nth laminate and J^n is the J for the nth laminate, the formula for J^n being given by equation (11) with A set equal to A^n, the cross-sectional area of the nth laminate.

Discussion

The advantage of the result presented here is that it extends a class of known isotropic solutions for the torsion of a shaft to the case of shape intrinsic orthotropy. It can be shown that the pure bending and axial extension or compression of a (shape intrinsic orthotropic) shaft with $\nu_{nz} = \nu_{tz}$ are the same as for an isotropic material (shaft) if one replaces the isotropic Young's modulus E by the orthotropic axial modulus E_z. Thus it has been shown here that the problem of the combined torsion, pure bending, and axial extension or compression of a shaft with shape intrinsic orthotropy, for which $\nu_{nz} = \nu_{tz}$, has the same solution as the same problem for an isotropic material, provided the isotropic elastic moduli E and G are replaced by E_z and G_{tz}, respectively, and the condition (18) is satisfied. Since the determination of the transverse shear in engineering beam theory depends on the axial normal stress, it follows that engineering beam theory can be extended from beams of isotropic material to beams with shape intrinsic orthotropy, for which $\nu_{nz} = \nu_{tz}$, by the same prescription. In fact, much of the content of texts on the mechanics of materials or strength of materials can be extended from isotropic materials to materials with shape intrinsic orthotropy, for which $\nu_{nz} = \nu_{tz}$.

There is an interesting point concerning the expeimental evaluation of the shear moduli G_{tz} and G_{nz} for shape intrinsic orthotropy. For any of the class of solutions obtained here for the torsion of a shaft with shape intrinsic orthotropy, the shear modulus G_{tz} can be experimentally determined by a torsion test of the entire shaft. Since G_{tz} is the only elastic constant involved in the formula relating to angular deflection, only one test is necessary. For example, if the shaft is circular it follows from equation (25), by setting ξ equal to one, that

$$T = \frac{\pi\alpha a^4}{2} G_{tz}. \tag{45}$$

Thus, by measuring T, a and α in an experiment, one can determine G_{tz}. However, if one takes a small circular torsion specimen from a cylinder with shape intrinsic orthotropy, a specimen whose axis is parallel to the axis of the shaft but which is not concentric with the shaft, the small specimen will appear to have rectilinear orthotropy and the relation between applied torque and angular deflection will involve both G_{tz} and G_{nz}. The formula relating the applied torque T to the angle of twist per unit length for a circular shaft with rectilinear orthotropy is given by equation (38) with ξ equal to one, thus

$$T = \frac{\pi\alpha a^4}{2}\left(\frac{2G_{nz}G_{tz}}{G_{nz} + G_{tz}}\right). \tag{46}$$

The effective shear modulus of the small specimen is thus seen to be a combination of G_{nz} and G_{tz}.

Acknowledgment

This investigation was supported by USPHS, Research Grant DE06859 from the National Institute of Dental Research, National Institutes of Health, Bethesda, MD 20205. The author thanks Richard T. Hart and Morteza M. Mehrabadi for helpful comments on an earlier draft. The author also thanks Richard T. Hart for the calculation of the curves illustrated in Figs. 3 and 4 and Louis P. Orth, Jr., for the preparation of the figures.

References

Bredt, R., 1896, "Kritische Bemerkungen zur Drehungs-Elastizitaet," *Zeitschritt Verein Deutscher Ingenieure*, Vol. 40, pp. 785–813.

Hearmon, R. F. S., 1961, *Applied Anisotropic Elasticity*, Oxford University Press, Oxford, England.

Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Elastic Body*, Holden Day, San Francisco, CA.

Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, 4Ed., Dover, New York.

Saint-Venant, A. -J. -C. B., 1855, "Memoire sur la Torsion des Prismes," *Memoires Presentes par Divers Savants a l'Academie des Sciences, Sciences Math. et Phys.*, Vol. 14, pp. 233–560.

Sokolnikoff, I. S., 1956, *Mathematical Theory of Elasticity*, McGraw Hill, New York.

Timoshenko, S., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd Ed., McGraw Hill, New York.

Poroelastic Solution of a Plane Strain Point Displacement Discontinuity

E. Detournay
Dowell Schlumberger Technology Center,
Tulsa, OK

A. H-D. Cheng
Department of Civil Engineering,
University of Delaware,
Newark, DE

The plane strain fundamental solution of an instantaneous and a continuous point displacement discontinuity is presented in this paper. These solutions, together with the one of a fluid source, are obtained on the basis of a decomposition technique proposed by Biot, which separates the displacement field into a time independent part satisfying an elasticity equation, and an irrotational part governed by a diffusion equation. We begin the derivation by presenting a continuous edge dislocation. The continuous point displacement discontinuity is obtained by differentiating, along the direction of the cut, the corresponding edge dislocation solution. The instantaneous influence functions are determined by further differentiating with respect to time. The displacement discontinuity and source singularities can be distributed on a crack surface to create displacement and flux jumps required for the numerical modeling of a fracture in a poroelastic medium.

Introduction

Crouch and Starfield (1983) pioneered the displacement discontinuity method (DDM) as a means of solving boundary value problems in elasticity. It has become a popular numerical method in the field of geomechanics, because of its ability of handling rock discontinuities and fractures (Crouch and Starfield, 1983; Wiles and Curran, 1982; Vandamme, 1986). The concept of the DDM can be traced back to the mathematical theory of dislocations (Bilby, 1968; Dundurs, 1969; Mura, 1982); in its current implementation, however, we view it as a variation of the indirect boundary element method, in which a fictitious density of a displacement discontintuity singularity, instead of Kelvin's point force, is distributed along the boundary of the elastic domain. A boundary element procedure, which includes boundary discretization, polynomial interpolation, etc., is then applied for the numerical solution.

The DDM is particularly appealing for problems involving fractures and discontinuities because the fundamental solution contains a displacement jump, thus requiring only one layer of singularity to be distributed along the crack. In addition, the so-called fictitious density of singularity on the crack surface has the physical significance of being the actual displacement discontinuity associated with the fracture.

The DDM is presently limited to solving elastic problems. However, many geomechanics problems, such as soil consolidation (Cheng and Liggett, 1984), hydraulic fracturing (Cleary, 1978), and stabilization of faults (Rudnicki, 1985) are dominated by coupled diffusion-deformation effects, and should therefore be analyzed within the framework of Biot's theory of poroelasticity (Biot, 1941). In order to extend the original DDM to poroelasticity, the fundamental solutions of an impulse point displacement discontinuity and an impulse point source are required.

The objective of this paper is to derive these solutions. In the process, a general methodology for deriving singular solutions in poroelasticity is outlined.

Poroelasticity

The theory of linear, isotropic poroelasticity was introduced by Biot (1941) for modeling the response of fluid-saturated porous solids. As in the original formulation of Biot (see also Rice and Cleary, 1976), we choose the basic dynamic variables to be the total stress σ_{ij} and the pore pressure p. The corresponding conjugate kinematic quantities are the solid strain e_{ij}, derivable from an average solid displacement vector u_i, and the variation of fluid volume per unit reference volume, ζ.

A consistent set of five material parameters for the linear isotropic theory is (Rice and Cleary, 1976): the shear modulus G, the drained and undrained Poisson's ratio ν and ν_u, Skempton's pore pressure coefficient B (ratio of the induced pore pressure over the variation of confining pressure under undrained conditions) and the permeability coefficient κ (which can be expressed as k/μ, where k is the intrinsic permeability and μ the dynamic fluid viscosity).

The governing equations of linear poroelasticity consist of the following (Biot, 1941; Rice and Cleary, 1976):

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 3, 1986; final revision April 24, 1987.
Paper No. 87-WA/APM-22.

- constitutive relations

$$\sigma_{ij} = 2Ge_{ij} + \frac{2G\nu}{1-2\nu}\delta_{ij}e - \alpha\delta_{ij}p \tag{1}$$

$$p = -\frac{2GB(1+\nu_u)}{3(1-2\nu_u)}e + \frac{2GB^2(1-2\nu)(1+\nu_u)^2}{9(\nu_u-\nu)(1-2\nu_u)}\zeta \tag{2}$$

- equilibrium equations

$$\sigma_{ij,j} = -F_i \tag{3}$$

- Darcy's law

$$q_i = -\kappa(p_{,i} - f_i) \tag{4}$$

- continuity equation

$$\frac{\partial\zeta}{\partial t} + q_{i,i} = \gamma \tag{5}$$

where $e = e_{ii}$ is the volumetric strain, q_i the specific discharge vector, F_i the bulk body force (fluid and solid), f_i the fluid body force, γ the volume rate of injection from fluid source, and α is the Biot coefficient of effective stress (Biot and Willis, 1957), defined as

$$\alpha = \frac{3(\nu_u-\nu)}{B(1-2\nu)(1+\nu_u)} \tag{6}$$

The foregoing can be combined to yield a set of field equations in terms of u_i and ζ. Joining equations (1) through (3) gives an elasticity equation with a fluid coupling term

$$G\nabla^2 u_i + \frac{G}{1-2\nu_u}e_{,i} - \frac{2GB(1+\nu_u)}{3(1-2\nu_u)}\zeta_{,i} = -F_i \tag{7}$$

Combining equations (2), (4), and (5), and also using equation (7), produces the following diffusion equation:

$$\frac{\partial\zeta}{\partial t} - c\nabla^2\zeta = \frac{\kappa B(1+\nu_u)}{3(1-\nu_u)}F_{i,i} - \kappa f_{i,i} + \gamma \tag{8}$$

where

$$c = \frac{2\kappa B^2 G(1-\nu)(1+\nu_u)^2}{9(1-\nu_u)(\nu_u-\nu)} \tag{9}$$

is a generalized consolidation coefficient (Rice and Cleary, 1976).

It should be noted that the body force and source terms in the above field equations can be used to introduce field singularities, such as point force, source, dipole, etc., into the solution.

Decomposition of the Displacement

Biot (1956) introduced the following decomposition which allows further uncoupling of the field equations (7) and (8):

$$u_i = u_i^o + \Delta u_i = u_i^o + \frac{B(1+\nu_u)}{3(1-\nu_u)}\Phi_{,i} \tag{10}$$

If the first part of displacement field is required to satisfy Navier's equation of elasticity with undrained coefficients, i.e.,

$$G\nabla^2 u_i^o + \frac{G}{1-2\nu_u}e_{,i}^o = -F_i \tag{11}$$

then it can be proven, from equation (7), that the potential Φ is defined by the identity

$$\zeta = \nabla^2\Phi \tag{12}$$

Substituting the preceding expression into equation (8), and relaxing a Laplacian leads to a diffusion equation for Φ

$$\frac{\partial\Phi}{\partial t} - c\nabla^2\Phi = g_1 + g_2 + g_3 + \psi_1 \tag{13}$$

in which

$$\begin{aligned}
\nabla^2 g_1 &= \frac{\kappa B(1+\nu_u)}{3(1-\nu_u)}F_{i,i} \\
\nabla^2 g_2 &= -\kappa f_{i,i} \\
\nabla^2 g_3 &= \gamma \\
\nabla^2\psi_1 &= 0
\end{aligned} \tag{14}$$

The completeness of the above solution has been proven by Biot (1956).

It is sometimes more convenient to use an alternate right-hand side for equation (13), in the form:

$$\frac{\partial\Phi}{\partial t} - c\nabla^2\Phi = -\frac{3c(1-\nu_u)(\nu_u-\nu)}{B(1-\nu)(1-2\nu_u)(1+\nu_u)}e^o + g_2 + g_3 + \psi_1 \tag{15}$$

In the above, the body force F_i has been eliminated using the divergence of equation (11).

Another useful relation can be obtained from equations (4), (5), and (12)

$$\frac{\partial\Phi}{\partial t} - \kappa p = g_2 + g_3 + \psi_2 \tag{16}$$

where $\nabla^2\psi_2 = 0$. The previous equation leads to an alternate definition of Φ:

$$\Phi = \kappa\int_0^t p\,dt + \int_0^t (g_2 + g_3 + \psi_2)dt \tag{17}$$

The above results suggest that the displacement field can be decomposed into an "undrained" part, u_i^o, satisfying an elasticity equation with undrained coefficient; and an irrotational part derivable from a potential, Φ, that is governed by a diffusion equation. Although the governing equations, (11) and (13), appear to be fully uncoupled, except for a body force term, the coupling generally persists through the boundary conditions in terms of the new variables. Hence, the "undrained" part, u_i^o, is generally time-dependent and cannot be solved independently from the variable Φ.

However, for problems in infinite domains, such as solutions of free-space Green's function, the above restriction does not apply. The undrained part, no longer subjected to a transient boundary condition, becomes time-independent; it then follows that the time-dependency is completely absorbed in the irrotational part. The harmonic functions ψ_1 and ψ_2 can also be dropped. The task of finding the fundamental solutions for poroelasticitiy is thus reduced to seeking the particular solution of two uncoupled, singular equations. The complete poroelastic solution is then the summation of the undrained and the transient parts.

Displacement Discontinuity

To derive the solution of a plane strain point displacement discontinuity, let us start from the solution of a continuous (in time) edge dislocation, that is located on the positive x_1 axis (see Fig. 1). Two in-plane dislocation modes are considered, namely the slip mode and the normal mode. The displacement jump across the cut is characterized by the following relationship:

$$\lim_{x_2\to 0-} u_i^{ec} - \lim_{x_2\to 0+} u_i^{ec} = \mathrm{H}(t)\mathrm{H}(x_1)d_i \tag{18}$$

where d_i is the displacement discontinuity, and H the Heaviside unit step function. The superscript *ec* stands for edge dislocation and continuous.

The instantaneous point displacement discontinuity required for the DDM can be obtained by differentiating the edge dislocation solution with respect to time and along the direction of the cut. The following displacement jump is then observed

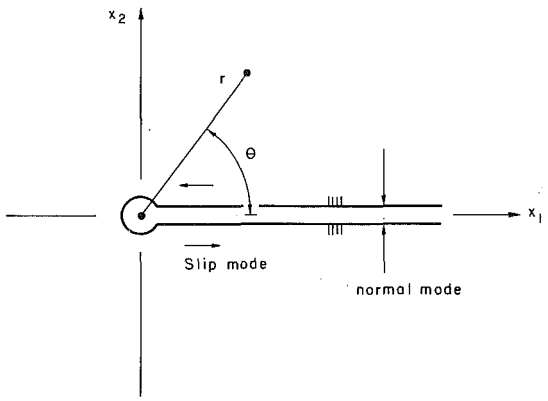


Fig. 1 Edge dislocation

$$\lim_{x_2\to 0-} u_i^{di} - \lim_{x_2\to 0+} u_i^{di} = \delta(t)\delta(x_1)d_i \tag{19}$$

in which δ is the one-dimensional Dirac delta function and where the superscript *di* denotes displacement discontinuity and instantaneous. The above instantaneous displacement discontinuity can be integrated along a segment of fracture and over an interval of time using certain shape functions, to create the desired displacement discontinuity element needed in the numerical method.

The two-dimensional, poroelastic solution of a continuous edge dislocation can be obtained using the following procedure. First, it is recognized that, at the instant of loading, a poroelastic material behaves as an elastic one with undrained material coefficients. Therefore, in the decomposition, the undrained displacement part, u_i^o, is required to fully absorb the initial value and to be time-independent. This ensures that equation (11) is satisfied. Secondly, the irrotational part, Φ, must satisfy equation (13) with zero initial condition. As a side condition one must also check that the time-dependent displacement field, obtained by differentiating Φ, is continuous and does not contain any additional displacement discontinuity. The displacement jump condition, equation (18), is then satisfied at all times.

Consequently, the initial displacement field is given by the classical elastic solution (Dundurs, 1969), with undrained coefficients:

$$(u_{11}^{ec})^o = \frac{1}{4\pi(1-\nu_u)}\,\mathrm{H}(t)[2(1-\nu_u)\theta + r_{,1}r_{,2}]$$

$$(u_{21}^{ec})^o = -\frac{1}{4\pi(1-\nu_u)}\,\mathrm{H}(t)[(1-2\nu_u)\ln r + r_{,1}r_{,1}]$$

$$(u_{12}^{ec})^o = \frac{1}{4\pi(1-\nu_u)}\,\mathrm{H}(t)[(1-2\nu_u)\ln r + r_{,2}r_{,2}] \tag{20}$$

$$(u_{22}^{ec})^o = \frac{1}{4\pi(1-\nu_u)}\,\mathrm{H}(t)[2(1-\nu_u)\theta - r_{,1}r_{,2}]$$

where r is the radial distance from the origin; $r_{,i} = x_i/r$; and, θ is the polar angle, which has a branch cut along the positive x_1 axis ($0 \le \theta \le 2\pi$), see Fig. 1. Note that the first subscript of displacement denotes the displacement component, while the second subscript is reserved for the dislocation mode (1 for slip, 2 for normal mode).

The presence of an edge dislocation is equivalent to the introduction of a singular body force F_{ij}, to the right-hand side of equation (11) (the second index again refers to the edge dislocation mode). The body force in turn becomes a forcing function in the diffusion equation (13) or (15). Substituting equation (20) into (15) and arbitrarily setting $g_2 = g_3 = \psi_1 = 0$ yields:

$$\frac{\partial \Phi_i^{ec}}{\partial t} - c\nabla^2\Phi_i^{ec} = \frac{3c(\nu_u-\nu)}{2\pi B(1-\nu)(1+\nu_u)}\,\mathrm{H}(t)\epsilon_{ij}(\ln r)_{,j} \tag{21}$$

where the subscript for Φ denotes the dislocation mode, and ϵ_{ij} is the two-dimensional permutation symbol, (i.e., $\epsilon_{11} = \epsilon_{22} = 0$ and $\epsilon_{12} = -\epsilon_{21} = 1$).

The physical significance of the preceding equation becomes clear after applying the Laplacian operator:

$$\frac{\partial \zeta_i^{ec}}{\partial t} - c\nabla^2\zeta_i^{ec} = \frac{3c(\nu_u-\nu)}{B(1-\nu)(1+\nu_u)}\,\mathrm{H}(t)\epsilon_{ij}\delta(\mathbf{x})_{,j} \tag{22}$$

in which $\delta(\mathbf{x}) = \delta(x_1)\delta(x_2)$ is the two-dimensional Dirac delta function. The time-dependent part of the poroelastic edge dislocation is thus created by the spatial derivative of a fluid volume source, which is also known as a dipole (opposing source and sink). In the above, the dipole is in the positive x_2 and negative x_1 direction, for slip and normal dislocation modes, respectively. As pointed out by Rudnicki (1987) the orientation of the dipole enforces a no-flux boundary condition across the dislocation line for the normal mode, and a constant (zero) pressure condition for the slip mode.

To solve equation (21), we first begin with the solution of a continuous source in a poroelastic medium (see Appendix) and then differentiate the expression. After some algebra, we obtain

$$\Phi_i^{ec} = \frac{3(\nu_u-\nu)}{8\pi B(1+\nu_u)(1-\nu)}\,\epsilon_{ij}rr_{,j}[E_1(\xi^2) + \xi^{-2}(1-e^{-\xi^2})] \tag{23}$$

where $\xi = \sqrt{r^2/4ct}$. From equation (10) the time-dependent displacement field for the edge dislocation is then given by:

$$\Delta u_{ij}^{ec} = \frac{\nu_u-\nu}{8\pi(1-\nu)(1-\nu_u)}\,\epsilon_{jk}[(\delta_{ik} - 2r_{,i}r_{,k})\xi^{-2}(1-e^{-\xi^2}) + \delta_{ik}E_1(\xi^2)] \tag{24}$$

Once again, the second subscript of the displacement denotes the edge dislocation mode. The foregoing field is continuous across the positive x_1 axis, as required in the original postulate.

The influence function for the stress field, σ_{ijk}^{ec}, is deduced from equation (24) using the strain displacement relations and the constitutive relations, (1) and (2). As before, the solution can be separated into an undrained part, $(\sigma_{ijk}^{ec})^o$, and an evolving part, $\Delta\sigma_{ijk}^{ec}$

$$(\sigma_{ijk}^{ec})^o = \frac{G}{2\pi(1-\nu_u)}\,\mathrm{H}(t)\epsilon_{k\ell}\frac{1}{r}(\delta_{i\ell}r_{,j} + \delta_{j\ell}r_{,i} - \delta_{ij}r_{,\ell} - 2r_{,i}r_{,j}r_{,\ell}) \tag{25}$$

$$\Delta\sigma_{ijk}^{ec} = -\frac{G(\nu_u-\nu)}{2\pi(1-\nu)(1-\nu_u)}\,\epsilon_{k\ell}\frac{1}{r}[2(r_{,i}r_{,j}r_{,\ell} - \delta_{ij}r_{,\ell})e^{-\xi^2} + (\delta_{i\ell}r_{,j} + \delta_{j\ell}r_{,i} + \delta_{ij}r_{,\ell} - 4r_{,i}r_{,j}r_{,\ell})\xi^{-2}(1-e^{-\xi^2})] \tag{26}$$

where the last subscript of the influence function refers to the dislocation mode. It is of interest to remark that by setting $k = 1$ in equations (25) and (26), a solution identical to the slip dislocation solution derived by Rice and Cleary (1976), using Muskhelishvili's complex variable technique, is obtained.

The aim here is to derive the fundamental solution of an instantaneous point displacement discontinuity. Before presenting such solutions, however, it is instructive to examine the limiting behavior of a continuous point displacement discontinuity. Differentiating equations (20) and (24) with respect to x_1, the following set of equations is obtained

$$(u_{ij}^{dc})^o = -\frac{1}{4\pi(1-\nu_u)} H(t)\frac{1}{r}[(1-2\nu_u)(\delta_{ij}r_{,2} + \delta_{i2}r_{,j} - \delta_{j2}r_{,i}) + 2r_{,i}r_{,j}r_{,2}] \quad (27)$$

$$\Delta u_{ij}^{dc} = -\frac{\nu_u - \nu}{4\pi(1-\nu_u)(1-\nu)}\frac{1}{r}[2(r_{,i}r_{,j}r_{,2} - \delta_{j2}r_{,i})e^{-\xi^2} + (\delta_{i2}r_{,j} + \delta_{j2}r_{,i} + \delta_{ij}r_{,2} - 4r_{,i}r_{,j}r_{,2})\xi^{-2}(1-e^{-\xi^2})] \quad (28)$$

The asymptotic behavior of the combined displacement field is such that

$$\lim_{t\to 0 \text{ or } \infty} u_{ij}^{dc} = -\frac{1}{4\pi(1-\bar{\nu})}\frac{1}{r}[(1-2\bar{\nu})(\delta_{ij}r_{,2} + \delta_{i2}r_{,j} - \delta_{j2}r_{,i}) + 2r_{,i}r_{,j}r_{,2}] \quad (29)$$

where $\bar{\nu} = \nu_u$ as $t \to 0$, and, $\bar{\nu} = \nu$ as $t \to \infty$. Thus, as expected, the poroelastic medium responds to the loading generated by a displacement discontinuity as an undrained material at $t = 0^+$ and as a drained material as $t \to \infty$. Note that the above expression is identical to the solution of an elastic point displacement discontinuity obtained by Wiles and Curran (1982) with the appropriate Poisson's ratio substituted.

Finally, we differentiate with respect to time to obtain the instantaneous displacement discontinuity solutions:

$$(u_{ij}^{di})^o = -\frac{1}{4\pi(1-\nu_u)}\delta(t)\frac{1}{r}[(1-2\nu_u)(\delta_{ij}r_{,2} + \delta_{i2}r_{,j} - \delta_{j2}r_{,i}) + 2r_{,i}r_{,j}r_{,2}] \quad (30)$$

$$\Delta u_{ij}^{di} = \frac{c(\nu_u - \nu)}{\pi(1-\nu_u)(1-\nu)}\frac{1}{r^3}[2(\delta_{j2}r_{,i} - r_{,i}r_{,j}r_{,2})\xi^4 e^{-\xi^2} - (\delta_{i2}r_{,j} + \delta_{j2}r_{,i} + \delta_{ij}r_{,2} - 4r_{,i}r_{,j}r_{,2})(1 - e^{-\xi^2} - \xi^2 e^{-\xi^2})] \quad (31)$$

$$(\sigma_{ijk}^{di})^o = \frac{G}{2\pi(1-\nu_u)}\delta(t)\frac{1}{r^2}[8r_{,i}r_{,j}r_{,k}r_{,2} - 2(\delta_{k2}r_{,i}r_{,j} + \delta_{ij}r_{,k}r_{,2}) - (\delta_{ik}\delta_{j2} + \delta_{jk}\delta_{i2} - \delta_{ij}\delta_{k2})] \quad (32)$$

$$\Delta\sigma_{ijk}^{di} = -\frac{2Gc(\nu_u - \nu)}{\pi(1-\nu)(1-\nu_u)}\frac{1}{r^4}\{[24r_{,i}r_{,j}r_{,k}r_{,2} - 12(\delta_{ij}r_{,k}r_{,2} + \delta_{k2}r_{,i}r_{,j}) - 3(\delta_{ik}\delta_{j2} + \delta_{jk}\delta_{i2} - 3\delta_{ij}\delta_{k2})][1 - (1+\xi^2)e^{-\xi^2}] - [12r_{,i}r_{,j}r_{,k}r_{,2} - 6(\delta_{k2}r_{,i}r_{,j} + \delta_{ij}r_{,k}r_{,2}) - 2\delta_{ik}\delta_{j2} - 2\delta_{jk}\delta_{i2} + 4\delta_{ij}\delta_{k2}]\xi^4 e^{-\xi^2} - [4r_{,i}r_{,j}r_{,k}r_{,2} - 4(\delta_{ij}r_{,k}r_{,2} + \delta_{k2}r_{,i}r_{,j}) + 4\delta_{ij}\delta_{k2})]\xi^6 e^{-\xi^2}\} \quad (33)$$

$$(p_i^{di})^o = \frac{BG(1+\nu_u)}{3\pi(1-\nu_u)}\delta(t)\frac{1}{r^2}(\delta_{i2} - 2r_{,i}r_{,2}) \quad (34)$$

$$\Delta p_i^{di} = \frac{4BGc(1+\nu_u)}{3\pi(1-\nu_u)}\frac{1}{r^4}[\delta_{i2}\xi^4 e^{-\xi^2} + 2(r_{,i}r_{,2} - \delta_{i2})\xi^6 e^{-\xi^2}] \quad (35)$$

$$(q_{ij}^{di})^o = \frac{3c(\nu_u - \nu)}{\pi B(1-\nu)(1+\nu_u)}\delta(t)\frac{1}{r^3}(\delta_{i2}r_{,j} + \delta_{j2}r_{,i} + \delta_{ij}r_{,2} - 4r_{,i}r_{,j}r_{,2}) \quad (36)$$

$$\Delta q_{ij}^{di} = -\frac{6c^2(\nu_u - \nu)}{\pi B(1-\nu)(1+\nu_u)}\frac{1}{r^5}[2(\delta_{i2}r_{,j} + \delta_{ij}r_{,2} - 3\delta_{j2}r_{,i})\xi^6 e^{-\xi^2} + 4(\delta_{j2}r_{,i} - r_{,i}r_{,j}r_{,2})\xi^8 e^{-\xi^2}] \quad (37)$$

In problems involving fluid flow in the fracture, it is also expected that there will be a flux jump across the crack surface. To account for this flux discontinuity, sources need to be distributed along the surface. The instantaneous source solution for poroelastic medium is obtained by differentiating the continuous source solution (given in the Appendix) with respect to time, which yields the following:

$$u_i^{si} = \frac{B(1+\nu_u)}{6\pi(1-\nu_u)}\frac{1}{r}r_{,i}(1-e^{-\xi^2}) \quad (38)$$

$$\sigma_{ij}^{si} = \frac{BG(1+\nu_u)}{3\pi(1-\nu_u)}\frac{1}{r^2}[(\delta_{ij} - 2r_{,i}r_{,j})(1-e^{-\xi^2}) - 2(\delta_{ij} - r_{,i}r_{,j})\xi^2 e^{-\xi^2}] \quad (39)$$

$$p^{si} = \frac{2B^2G(1-\nu)(1+\nu_u)^2}{9\pi(1-\nu_u)(\nu_u - \nu)}\frac{1}{r^2}\xi^2 e^{-\xi^2} \quad (40)$$

$$q_i^{si} = \frac{2c}{\pi}\frac{1}{r^3}r_{,i}\xi^4 e^{-\xi^2} \quad (41)$$

Discussion

The influence functions of an instantaneous displacement discontinuity and source can be distributed on the locus of a fracture to generate a desirable solution field. In particular, the following integral equations can be exploited for the numerical solution of a boundary value problem

$$\sigma_{ij}(\mathbf{x};t) = \int_0^t\int_\Gamma d_k(\chi;\tau)\sigma_{ijk}^*(\mathbf{x}-\chi;t-\tau)d\Gamma(\chi)d\tau + \int_0^t\int_\Gamma 9(\chi;\tau)\sigma_{ij}^*(\mathbf{x}-\chi;t-\tau)d\Gamma(\chi)d\tau \quad (42)$$

$$p(x;t) = \int_0^t\int_\Gamma d_i(\chi;\tau)p_i^*(\mathbf{x}-\chi;t-\tau)d\Gamma(\chi)d\tau + \int_0^t\int_\Gamma 9(\chi;\tau)p^*(\mathbf{x}-\chi;t-\tau)d\Gamma(\chi)d\tau \quad (43)$$

in which 9 is the normal flux discontinuity, $\mathbf{x}$ is the field point where the stress or the pressure is evaluated, and χ is a point on the fracture; both $\mathbf{x}$ and χ are referred to a global coordinate system (see Fig. 2). The influence functions, denoted by an asterisk (*) in equations (42) and (43), are defined as follows:

$$\sigma_{ijk}^* = \ell_{it}\ell_{jm}\sigma_{\ell mk}^{di}(\mathbf{x}';t-\tau)$$
$$p_i^* = p_i^{di}(\mathbf{x}';t-\tau)$$
$$\sigma_{ij}^* = \ell_{ik}\ell_{j\ell}\sigma_{k\ell}^{si}(\mathbf{x}';t-\tau) \quad (44)$$
$$p^* = p^{si}(\mathbf{x}';t-\tau)$$

where $\ell_{ij} = \ell_{ij}(\chi)$ are the directional cosines between the local (crack) and the global coordinate system, and $\mathbf{x}'$ the field point in the local coordinate system (see Fig. 2). Transformation between the two coordinate systems is done according to

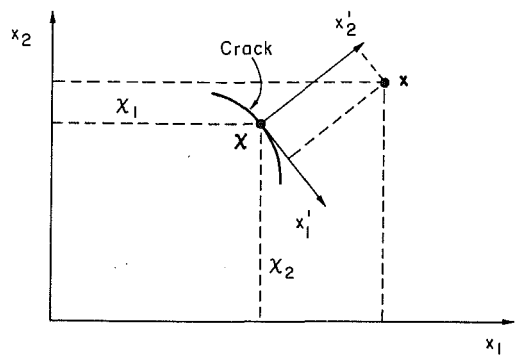


Fig. 2 Global and local crack coordinate systems

$$x_i' = \ell_{ji}(x_j - \chi_j) \tag{45}$$

Equations (42) and (43) can be discretized, numerically integrated in both time and spatial coordinates, and collocated for the known stress and pressure boundary conditions along the crack surface. These operations result in the formulation of a linear system of algebraic equations which need to be solved for the displacement and flux jumps at each time step. The influence functions defined in equations (30) through (37) are then used to evaluate quantities such as displacement, stress, pressure, and flux in the poroelastic medium. The details of the numerical implementation and also the applications to hydraulic fracturing problems will be separately reported (Detournay et al., 1987; Vandamme et al., 1987).

Conclusion

The fundamental plane strain solutions of a point displacement discontinuity and a point source in a poroelastic medium have been derived. The methodology is based on the decomposition technique proposed by Biot (1956). This procedure offers a systematic approach for finding Green's functions for poroelasticity. In addition to the displacement discontinuity and fluid source solutions provided herein, other types of Green's functions, such as point force and fluid dilatation, can also be derived (Cheng and Predeleanu, 1987).

The influence functions presented here can be used in a boundary element procedure to solve general boundary value problems governed by poroelasticity. This numerical technique would be particularly appealing for solving problems involving fractures and discontinuities.

Acknowledgment

The authors would like to acknowledge the management of Dowell Schlumberger for allowing publication of this material. They also want to express their special thanks to Dr. J-C. Roegiers for his continuing support of this project. The partial support of a grant from the University of Delaware Research Foundation to the second author is also appreciated.

References

Abramowitz, M., and Stegun, I. A., 1972, *Handbook of Mathematical Functions*, Dover.

Bilby, B., and Eshelby, J. D., 1968, "Dislocations and the Theory of Fracture," *Fracture*, Liebowitz, H., ed., Academic Press, pp. 99–182.

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

Biot, M. A., 1956, "General Solutions of the Equations of Elasticity and Consolidation for a Porous Material," ASME JOURNAL OF APPLIED MECHANICS, Vol. 78, pp. 91–96.

Biot, M. A., and Willis, D. G., 1957, "The Elastic Coefficients of the Theory of Consolidation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 24, pp. 594–601.

Carslaw, H. S., and Jaeger, J. C., 1959, *Conduction of Heat in Solids*, 2nd Ed., Oxford Univ. Press (Clarendon), London.

Cheng, A. H-D., and Liggett, J. A., 1984, "Boundary Integral Equation Method for Linear Porous-Elasticity with Applications to Soil Consolidation," *Int. J. Numer. Meth. Engng.*, Vol. 20, No. 2, pp. 255–278.

Cheng, A. H-D., and Predeleanu, M., 1987, "Transient Boundary Element Formulation for Poroelasticity," *Appl. Math. Modelling*, to appear.

Cleary, M. P., 1978, "Moving Singularities in Elasto-Diffusive Solids with Applications to Fracture Propagation," *Int. J. Solids Struct.*, Vol. 14, pp. 81–97.

Crouch, S. L., and Starfield, A. M., 1983, *Boundary Element Methods in Solid Mechanics*, Allen and Unwin.

Detournay, E., Roegiers, J-C., and Cheng, A. H-D., 1987, "Some New Examples of Poroelastic Effects in Rock Mechanics," *Rock Mechanics, Proc. 28th U.S. Symp.*, University of Arizona, I. W. Farmer, J. J. K. Daemen, C. S. Desai, C. E. Glass, and S. P. Neuman, eds., Balkema Publishers, pp. 575–584.

Dundurs, J., 1969, "Elastic Interaction of Dislocation with Inhomogeneities," *Mathematical Theory of Dislocations*, Mura, T., ed., Joint Appl. Mech. and Fluid Engng. Conf., pp. 70–115.

Mura, T., 1982, *Micromechanics of Defects in Solids*, Martinus Nijhoff Publ., Hague.

Rice, J. R., and Cleary, M. P., 1976, "Some Basic Stress-Diffusion Solutions for Fluid Saturated Elastic Porous Media with Compressible Constituents," *Rev. Geophys. Space Phys.*, Vol. 14, pp. 227–241.

Rudnicki, J. W., 1985, "Effect of Pore Fluid Diffusion on Deformation and Failure of Rock," *Mechanics of Geomaterials*, Bazant, Z., ed., Wiley, pp. 315–347.

Rudnicki, J. W., 1987, "Plane Strain Dislocations in Linear Elastic Diffusive Solids," ASME Journal of Applied Mechanics, in print.

Vandamme, L., 1986, "A Three-Dimensional Displacement Discontinuity Model for the Analysis of Hydraulic Fractures," Ph.D. Dissertation, University of Toronto, Canada.

Vandamme, L., Detournay, E., and Cheng, A. H-D., 1987, "A Poroelastic Displacement Discontinuity Method for Hydrofracture Simulation," submitted to *Int. J. Numer. Anal. Meth. Geomech.*

Wiles, T. D., and Curran, J. H., 1982, "A General 3-D Displacement Discontinuity Method," *Proc. Int. Conf. Num. Anal. Metho. Geomech.*, Edmonton, pp. 103–111.

APPENDIX

The governing equation for a two-dimensional continuous fluid source of constant injection rate Q is given by:

$$c\nabla^2\zeta^{sc} - \frac{\partial\zeta^{sc}}{\partial t} = -QH(t)\delta(\mathbf{x}) \tag{A-1}$$

Solution of the above is well-known (Carslaw and Jaeger, 1959), i.e.,

$$\zeta^{sc} = \frac{Q}{4\pi c} E_1(\xi^2) \tag{A-2}$$

where $\xi = \sqrt{r^2/4ct}$; and E_1 is the exponential integral (Abramowitz and Stegun, 1972)

$$E_1(\eta) = \int_\eta^\infty \frac{e^{-z}}{z}\, dz \tag{A-3}$$

Because of axial symmetry, the scalar potential Φ can be obtained by integrating equation (12) in the radial direction, and

$$\Phi^{sc} = \frac{Q}{16\pi c} r^2[(1+\xi^{-2})E_1(\xi^2) + \xi^{-2}(2\ln r - e^{-\xi^2})] \tag{A-4}$$

According to equation (10), the poroelastic displacement field due to a continuous source is given by

$$u_i^{sc} = \frac{QB(1+\nu_u)}{24\pi c(1-\nu_u)} r r_{,i}[E_1(\xi^2) + \xi^{-2}(1-e^{-\xi^2})] \tag{A-5}$$

The pressure and flux expressions are the following

$$p^{sc} = \frac{Q}{4\pi\kappa} E_1(\xi^2) \tag{A-6}$$

$$q_i^{sc} = \frac{Q}{2\pi} \frac{r_{,i}}{r} e^{-\xi^2} \tag{A-7}$$

K. Walton
School of Mathematics,
University of Bath,
Claverton Down,
Bath BA2 7AY, England

P. J. Digby
Division of Rock Mechanics,
University of Luleå,
S-951 87 Luleå, Sweden

Wave Propagation Through Fluid Saturated Porous Rocks

A sedimentary rock is modeled by a random packing of identical spherical particles. The connected pore space is filled with an inviscid, compressible fluid. A low-frequency expansion technique is used to calculate the effective wave speeds explicitly in terms of the microstructural properties of the rock considered. The effect of both the pore fluid and the initial confining pressure to which the rock is subjected can be included in the calculations.

1 Introduction

In the present paper we derive the basic equations for wave propagation in a fluid saturated rock having connected pore space. Here, as in Biot's work (1956), the rock will always be elastically isotropic and homogeneous over lengths large compared with a typical grain diameter. We also consider the case in which the pore space is filled with an inviscid, compressible fluid. We derive the basic equations of motion for the medium when a steady state disturbance, whose amplitude is small compared with the contact radius of the contacting grains, propagates through the medium. The equations first derived by Biot (1956) to study this problem have long been regarded as standard. They may be written in the form

$$\left.\begin{aligned} N\nabla^2<\mathbf{u}> + \nabla[(A+N)\nabla\cdot<\mathbf{u}> + Q\nabla\cdot<\mathbf{v}>] \\ = -\omega^2[\rho_{11}<\mathbf{u}> + \rho_{12}<\mathbf{v}>] \\ \text{and} \\ \nabla[Q\nabla\cdot<\mathbf{u}> + R\nabla\cdot<\mathbf{v}>] \\ = -\omega^2[\rho_{12}<\mathbf{u}> + \rho_{22}<\mathbf{v}>] \end{aligned}\right\} \quad (1)$$

In equation (1) above, and for the special case of a dry porous rock, the constitutive coefficients N and A may be interpreted as the effective Lamé coefficients for the dry rock framework. Explicit expressions for the effective Lamé coefficients of an initially hydrostatically stressed dry granular rock, modeled by a random packing of contacting, identical spherical particles, have been obtained by Digby (1981) and Walton (in press). The coefficients ρ_{11}, ρ_{22}, and ρ_{12} in equation (1) above describe the coupling between the motion of the fluid and that of the solid. The components of the average solid and fluid displacement vectors, $<\mathbf{u}>$ and $<\mathbf{v}>$, corresponds to solid and fluid displacements, averaged, in some sense, over regions whose dimensions are large compared with a typical grain diameter but small compared with lengths over which the macroscopic processes are occurring. Similar remarks apply to the other relevant field quantities appearing in a number of other equations derived in Biot's paper.

The application of Biot's theory to the study of diffusion and wave propagation processes in fluid saturated, poroelastic rocks is, of course, a subject of considerable importance. A number of equally plausible interpretations, on both the method of averaging the various field quantities and the physical meanings of the coefficients which appear in Biot's paper have, therefore, been advanced by both Biot himself and a large number of subsequent writers. Despite this considerable effort, however, we believe that with the exception of a few special cases, considerable ambiguity still exists here. We believe that some progress towards a solution of the problem described above was first made by Burridge and Keller (1981) and Walton (1977).

Burridge and Keller (1981), in their elegant application of the two-space method of homogenization, related the coefficients in equations (1) above to the microstructural properties of the rock considered.

Walton (1977), on the other hand, performed the same exercise rigorously by a rather different method. A low-frequency expansion technique was used. In its original form, Walton's work (1977) deals only with the propagation of longitudinal waves through an initially stressed fluid saturated rock whose framework is modeled by a simple cubic packing of identical, contacting spherical particles, and is, of course, far less general than Burridge and Keller's work. However, the results obtained are more directly applicable to the study of wave propagation through initially stressed, fluid saturated rocks in which the dominant mode of deformation of the rock framework is by grain contact deformation. In addition, Walton's work (1977) allows a more explicit interpretation of Biot's dynamic coupling coefficients ρ_{11}, ρ_{22}, and ρ_{12}, which appear in equations (1) above.

In the present paper we will therefore generalize Walton's original work (1977). The propagation of *both* longitudinal and transverse steady state disturbances through fluid saturated rocks will be studied.

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 6, 1987.

2 The Low Frequency Expansion Technique and Averaging Procedure

(a) The Governing Microstructural Equations. The motion of the fluid and solid components will be assumed small enough to be governed by the linearized equations of fluid mechanics and elasticity. Then with the $e^{-i\omega t}$ time dependence suppressed (steady state disturbance of frequency ω), the equations of motion governing vibrations about an equilibrium state are

$$\left.\begin{aligned} \sigma_{ij,j} &= -\rho_s \omega^2 u_i \quad \text{in the solid grains} \\ \sigma_{ij} &= \lambda \delta_{ij} u_{k,k} + \mu(u_{i,j} + u_{j,i}) \end{aligned}\right\} \quad (2)$$

and

$$\left.\begin{aligned} p_{,i} &= \rho_f \omega^2 v_i \quad \text{in the inviscid pore fluid} \\ p &= -\kappa_f v_{k,k} \end{aligned}\right\} \quad (3)$$

In equations (2) and (3) above and throughout the remainder of this paper, standard suffix notation will be used, with δ_{ij} denoting the usual Kronecker delta symbol. ρ_s, λ, and μ denote the density and Lamé coefficients for the solid grains (assumed elastically isotropic and homogeneous). ρ_f and κ_f denote the density and bulk modulus of the pore fluid component (assumed inviscid and compressible). σ_{ij} and u_i denote the increments in the Cartesian components of the Cauchy stress tensor and displacement vector, respectively, for the solid component. Similar remarks apply to the fluid displacement vector $\mathbf{v}$ for the fluid component. Across the fluid-solid boundaries we also require continuity of the traction and normal displacements, that is,

$$\left.\begin{aligned} n_k u_k &= n_k v_k \\ \text{and} \quad \sigma_{ij} n_j &= -p n_i \end{aligned}\right\} \quad (4)$$

where, in equation (4), n_i denotes a component of the unit normal to the fluid-solid boundary pointing out of the solid, that is, into the pore fluid.

(b) Low-Frequency Power Series Expansion Technique. We follow the approach of Walton (1977). Thus we write each field quantity appearing in equations (2)–(4), for example σ_{ij}, in the form

$$\sigma_{ij} = \sigma_{ij}^{(0)} + \omega \sigma_{ij}^{(1)} + \omega^2 \sigma_{ij}^{(2)} + \quad (5)$$

This power series expansion is supposed to be valid only for those values of the frequency ω for which the effective wavelength of a wave travelling in an infinite medium made of the fluid saturated granular material would be much larger than the grain radius, R. In equation (5), the superscripts denote the order of approximation used. Strictly speaking, one should make equations (2)–(4) dimensionless, introduce a dimensionless frequency, and expand all quantities in powers of this. For convenience, however, all of our equations will be expressed in terms of the physical frequency ω.

(c) The Averaging Procedure. Our ultimate aim is to obtain equations expressed in terms of "averaged" field and physical quantities. We will again follow Walton's approach (1977), but with one major difference. Our representative volume element will now contain many spheres (grains) in the medium. The average of any field quantity, for example $\sigma_{ij}(\mathbf{y})$ is denoted by $\sigma_{ij}(\mathbf{x}) = <\sigma_{ij}>$ and is defined by the equation

$$<\sigma_{ij}> = \frac{1}{V}\int_{V_s} \sigma_{ij}(\mathbf{y}) dV_y \quad (6)$$

The integral in equation (6) is now taken over a large volume (the representative volume), centered on the point $\mathbf{x}$. V denotes its total volume. V_f and V_s denote those parts of V occupied by the fluid and solid components, respectively. Throughout this paper, the porosity η of the rock considered is defined to be the ratio of the volume of fluid V_f contained in the representative element to its total volume V. Remarks analogous to those made above apply to quantities defined only in the fluid part of the representative element. We shall also make frequent use of equations of the type

$$<\sigma_{ij}>_{,j} = \frac{\partial <\sigma_{ij}>}{\partial x_j} = \frac{1}{V}\int_{S_s(\mathbf{x})} \sigma_{ij} n_j dS_y \quad (7)$$

which follow directly from equation (6). In equation (7), $s_s(\mathbf{x})$ denotes the external bounding surface of the representative volume intersected by the solid grains, having a unit normal n_j and an element of area dS_y. In equation (7), and throughout this paper, differentiation of averaged field quantities is taken with respect to the coordinates of the centroid of the representative volume element.

3 Some Basic Equations Derived from the Low Frequency Expansion Technique

(a) Some General Considerations. We first present some basic equations which may be derived independently of any expansion technique. We first note that one immediate consequence of the microstructural equations (3) is that the fluid motion is irrotational, that is

$$\text{curl}\mathbf{v} = \mathbf{0} \quad (8)$$

Hence, (at every order of approximation in equation (5)) there exists a displacement potential ϕ defined by the equation

$$\mathbf{v} = \nabla \phi \quad (9)$$

To obtain further equations, we note that by direct application of the averaging procedure described in Section 2(*c*) of this paper and the first of the boundary conditions (4) to the second of equations (2) and (3) gives

$$\frac{<\sigma_{kk}>}{3\kappa_s} - \frac{<p>}{\kappa_f} = \frac{1}{V}\int_{V_s} u_{k,k} dV + \frac{1}{V}\int_{V_f} v_{k,k} dV = <u_j>_{,j} + <v_j>_{,j} \quad (10)$$

Similarly, from the first of equations (2) and (3) and the second of the boundary conditions (4), one obtains

$$<\sigma_{ij}>_{,j} - <p>_{,i} = -\rho_s \omega^2 <u_i> - \rho_f \omega^2 <v_i> \quad (11)$$

Equations (10) and (11) above correspond to equations (55) and (61), respectively, in Walton (1977). They are also valid at every order of approximation in the power series representation (5).

(b) Further Equations Obtained from the Low Frequency Expansion Technique. These equations are all obtained by first substituting the power series expansions (5) into the basic microstructural equations (2)–(4) and then equating coefficients of like powers of the frequency ω. We then derive the following equations:

Zeroth Order Motion. In the solid we have

$$\left.\begin{aligned} \sigma_{ij,j}^{(0)} &= 0 \\ \sigma_{ij}^{(0)} &= \mu(u_{i,j}^{(0)} + u_{j,i}^{(0)}) + \lambda \delta_{ij} u_{k,k}^{(0)} \end{aligned}\right\} \quad (12)$$

and in the fluid we have

$$\nabla p^{(0)} = \mathbf{0} \quad (13)$$

Hence $p^{(0)} = \text{constant}$ (which we take as zero), therefore

$$p^{(0)} = 0 \ \text{(throughout the fluid)} \quad (14)$$

Also from boundary conditions (4)

$$\sigma_{ij}^{(0)} = 0 \ \text{(on the solid-fluid interface)} \quad (15)$$

We also suppose that the zeroth order motion corresponds to zero solid stress, that is,

$$\sigma_{ij}^{(0)} = 0 \text{ (throughout the solid)} \tag{16}$$

From equations (3), (5), and (14) it follows that

$$\nabla \cdot \mathbf{v}^{(0)} = 0 \text{ (throughout the fluid)} \tag{17}$$

From equations (8), (9), (17),

$$\left.\begin{aligned} \text{curl}\mathbf{v}^{(0)} &= \mathbf{0} \\ \mathbf{v}^{(0)} &= \nabla\phi^{(0)} \\ \nabla^2\phi^{(0)} &= 0 \end{aligned}\right\} \tag{18}$$

and on the inner boundaries (the solid-fluid interface),

$$u_k^{(0)} n_k = v_k^{(0)} n_k \tag{19}$$

The zeroth order motion considered here, therefore, consists of a solid motion which, from equations (12) and (16), is that of a rigid body translation (since there is zero stress and the displacement is bounded at infinity), and a fluid motion which has zero accompanying pressure. Thus, with a zero superscript denoting zeroth order, we have also

$$\mathbf{u}^{(0)} = \text{constant (throughout the solid)} \tag{20}$$

First Order Motion. In the solid we have

$$\left.\begin{aligned} \sigma_{ij,j}^{(1)} &= 0 \\ \sigma_{ij}^{(1)} &= \mu(u_{i,j}^{(1)} + u_{j,i}^{(1)}) + \lambda\delta_{ij}u_{k,k}^{(1)} \end{aligned}\right\} \tag{21}$$

and in the fluid,

$$\left.\begin{aligned} &\nabla p^{(1)} = \mathbf{0} \\ &\text{that is, } p^{(1)} = \text{constant throughout the fluid} \\ &\text{Also,} \quad \nabla \cdot \mathbf{v}^{(1)} = -p^{(1)}/\kappa_f \end{aligned}\right\} \tag{22}$$

Again, from equations (8) and (9) we have

$$\left.\begin{aligned} \text{curl}\mathbf{v}^{(1)} &= \mathbf{0} \\ \mathbf{v}^{(1)} &= \nabla\phi^{(1)} \\ \text{Then,} \quad \nabla^2\phi^{(1)} &= -p^{(1)}/\kappa_f \end{aligned}\right\} \tag{23}$$

Also, on the inner boundaries, we have

$$\left.\begin{aligned} n_k u_k^{(1)} &= n_k v_k^{(1)} \\ \text{and} \quad \sigma_{ij}^{(1)} n_j &= -n_i p^{(1)} \end{aligned}\right\} \tag{24}$$

where from equation (22), $p^{(1)}$ in equation (24) is constant.

Second Order Motion. Here, it will be necessary to consider only the fluid motion in the subsequent sections of this paper. Thus, to this order, we need only consider (from equation (18)), the equation

$$\nabla p^{(2)} = \rho_f \mathbf{v}^{(0)} = \rho_f \nabla\phi^{(0)} \tag{25}$$

We then obtain

$$p^{(2)} = \rho_f \phi^{(0)} + \text{constant} \tag{26}$$

4 The Complete System of Equations for Wave Propagation Through a Fluid Saturated Granular Rock

(*a*) Fluid Motion. It will be seen that the fluid motion will, to the lowest order, be completely determined by the function $\phi^{(0)}$ appearing in equations (18) above for the zeroth order fluid motion. To satisfy equations (18) above for $\phi^{(0)}$, subject to the boundary conditions (19) on the inner boundaries, we first consider a potential function $\phi^{(0)}$ of the form

$$\phi^{(0)} = (\mathbf{\Phi} - \mathbf{u}^{(0)}) \cdot \mathbf{\Psi} + \mathbf{u}^{(0)} \cdot \mathbf{r} \tag{27}$$

In equation (27) $\mathbf{u}^{(0)}$ is a constant vector determined from the zeroth order solid motion (equation (20)), and $\mathbf{\Phi}$ is also a constant vector. The vector potential function $\mathbf{\Psi}(\mathbf{r})$, depending only on the geometry of the medium considered, satisfies the following conditions:

$$\left.\begin{aligned} &\nabla^2\mathbf{\Psi} = \mathbf{0} \text{ in the fluid} \\ \text{and} \quad &(\hat{n} \cdot \nabla)\mathbf{\Psi} = \mathbf{0} \text{ on the inner boundaries} \end{aligned}\right\} \tag{28}$$

These conditions are of course insufficient to determine either $\mathbf{\Psi}(\mathbf{r})$ or $\phi^{(0)}$ uniquely, since the boundary conditions on the external surface of the representative volume element are unknown. However, the form of the potential function for $\phi^{(0)}$ given in equations (27), (28) may be regarded as a particular integral to the original problem described by equations (18) and (19). By applying the averaging procedure described in Section 2(*c*) of this paper, it then follows from equations (18) and (26)–(28) that

$$< v_i^{(0)} > = \frac{\eta < u_i^{(0)} >}{(1-\eta)} + (\Phi_j - u_j^{(0)}) \frac{1}{V} \int_{V_f} \frac{\partial \Psi_j}{\partial x_i} dV \tag{29}$$

and

$$< p^{(2)} >_{,i} = \rho_f < \phi^{(0)} >_{,i}$$

$$= \frac{\eta \rho_f < u_i^{(0)} >}{(1-\eta)} + \rho_f (\Phi_j - u_j^{(0)}) \frac{1}{V} \int_{S_f} \Psi_j n_i dS_y \tag{30}$$

We now use equation (30) to provide the required additional constraint on the values taken by the function $\mathbf{\Psi}(\mathbf{x})$ on the external boundary S_f of the representative volume element. Thus, we postulate that the averaged pressure, $< p^{(2)} >$ should be linear in $\mathbf{x}$, that is, $< p^{(2)} >$ is of the form $\mathbf{A} \cdot \mathbf{x}$ where $\mathbf{A}$ is a constant vector. This, from equations (27) and (30), is equivalent to supposing that we must find a particular integral of equations (28), such that for a statistically isotropic medium,

$$\frac{1}{V} \int_{S_f} \Psi_j n_i dS_y = \eta \delta_{ij} \tag{31}$$

A reasonable choice is to take $\Psi_j = x_j$ on S_f. From equations (29) and (30) we then have

$$\left\{ < v_i^{(0)} > - \frac{\eta < u_i^{(0)} >}{(1-\eta)} \right\} \rho_f = K_{ij} \left\{ < p^{(2)} >_{,j} - \frac{\eta \rho_f < u_j^{(0)} >}{(1-\eta)} \right\} \tag{32}$$

In which

$$\eta K_{ij} = \frac{1}{V} \int_{V_f} \frac{\partial \Psi_j}{\partial x_i} dV \tag{33}$$

In particular, if the medium is stastistically isotropic, $K_{ij} = \delta_{ij} K$ where K is a scalar depending only on the geometrical properties of the packing used to model the rock we are considering. Hence K can, in principle, be determined numerically. Also, in this case, equation (32) reduces to

$$< p^{(2)} >_{,i} = \frac{\eta \rho_f < u_i^{(0)} >}{(1-\eta)} + \rho_f \chi \left\{ < v_i^{(0)} > - \frac{\eta < u_i^{(0)} >}{(1-\eta)} \right\} \tag{34}$$

Where, in equation (34), $\chi = K^{-1}$. From equations (5), (14), (22), and (34) we then obtain, to the lowest order,

$$< p >_{,i} = \omega^2 < p^{(2)} >_{,i}$$

$$= \rho_f \omega^2 \left\{ \chi < v_i > + \frac{\eta(1-\chi) < u_i >}{(1-\eta)} \right\} \tag{35}$$

Equation (35) is equivalent to equation (58) in Walton's work (1977) for the case of a simple cubic packing.

It is very unlikely that the quantity $\chi (= K^{-1})$ in equations (33)–(35) above could ever be determined analytically, in a general case. However, an upper bound on the scalar K can be obtained from the following considerations.

The scalar K, for a statistically isotropic medium, is defined in terms of the vector potential function $\mathbf{\Psi}(\mathbf{x})$ through equation (33). We note that in addition to equations (28), each component $\Psi_j(\mathbf{x})$ now also satisfies the condition that

$\Psi_j(\mathbf{x})=x_j$ on the external boundary of the representative volume element. Suppose now that $\Psi_j^*(\mathbf{x})$ is *any* other function, not necessarily satisfying equations (28) but is such that $\Psi_j^*(\mathbf{x})=x_j$ on the same external boundary. We consider the volume integral

$$\int_{V_f} |\nabla(\Psi_j^* - \Psi_j)|^2 dV \geq 0.$$

We can then easily show that

$$\int_{V_f} |\nabla \Psi_j^*|^2 dV \geq \int_{V_f} |\nabla \Psi_j|^2 dV.$$

We then have,

$$\int_{V_f} |\nabla \Psi_j^*|^2 dV \geq \int_{V_f} |\nabla \Psi_j|^2 dV$$

$$= \int_S \Psi_j \hat{n} \cdot \nabla \Psi_j dS$$

$$= \int_S x_j \hat{n} \cdot \nabla \Psi_j dS$$

$$= \int_{V_f} \frac{\partial \Psi_j}{\partial x_j} dV$$

$$= V_f K_{jj} \quad \text{(no sum over } j\text{!)}$$

$$= V_f K$$

Hence,

$$K \leq \frac{1}{V_f} \int_{V_f} |\nabla \Psi_j^*|^2 dV$$

Since Ψ_j^* is *any* function satisfying $\Psi_j^*=x_j$ on the external boundary, we now take $\Psi_j^*=x_j$ *throughout* the fluid volume V_f and we have

$$\left.\begin{array}{l} K \leq 1 \\ \text{that is} \quad \chi \geq 1 \end{array}\right\} \tag{36}$$

Our required lower bound for the numerical factor χ.

(*b*) Solid Motion. We now derive the constitutive equation for the fluid saturated rock framework. We consider the problem for the solid component defined by equations (21) and the second of equations (24) for the first order motion. To complete the specification of this problem, conditions on the external boundary of the representative volume intersected by the solid grains are now required. Here, following our earlier work (Digby, 1981; Walton, in press), we suppose that the displacement $\mathbf{u}^{(1)}$ will be linear on this external boundary, that is, we write

$$u_i^{(1)} = e_{ij}^{(1)} x_j \tag{37}$$

where in equation (37) we write

$$e_{ij}^{(1)} = \frac{1}{2V_s} \int_{S_s} (u_j^{(1)} n_i + u_i^{(1)} n_j) dS_y \tag{38}$$

We now consider the displacement field $u_i^{(1)} + (p^{(1)}x_i/3\kappa_s)$, denoted by $w_i^{(1)}$, and the associated stress field $\sigma_{ij}^{(1)} + \delta_{ij}p^{(1)}$, denoted by $\tau_{ij}^{(1)}$. This displacement field and stress field satisfies the following boundary value problem:

$$\left.\begin{array}{l} \tau_{ij,j}^{(1)} = 0 \\ \qquad\qquad \text{in the solid grains} \\ \tau_{ij}^{(1)} = \lambda \delta_{ij} w_{k,k}^{(1)} + \mu (w_{i,j}^{(1)} + w_{j,i}^{(1)}) \end{array}\right\} \tag{39}$$

$$\left.\begin{array}{l} \tau_{ij}^{(1)} n_j = 0 \text{ on the inner boundaries} \\ w_i^{(1)} = \left(e_{ij}^{(1)} + \dfrac{p^{(1)}\delta_{ij}}{3\kappa_s}\right) x_j \quad \text{on the external boundary} \end{array}\right\} \tag{40}$$

We note that the problem defined by equations (39) and (40) corresponds to the problem which would arise when we calculate the effective elastic Lamé coefficients for a *dry* packing. Thus $\tau_{ij}^{(1)}$ and $w_i^{(1)}$ will be related through the effective Lamé coefficients for the dry packing. Denoting these by λ^* and μ^* we then have

$$\frac{1}{V}\int_{V_s} \tau_{ij}^{(1)} dV = 2\mu^* \left(e_{ij}^{(1)} + \frac{p^{(1)}\delta_{ij}}{3\kappa_s}\right) + \lambda^* \delta_{ij} \left(e_{kk}^{(1)} + \frac{p^{(1)}}{\kappa_s}\right) \tag{41}$$

That is, from equation (38)

$$\langle \sigma_{ij}^{(1)} \rangle + \left(1 - \eta - \frac{\kappa_s^*}{\kappa_s}\right) \frac{\langle p^{(1)} \rangle}{\eta} \delta_{ij}$$

$$= \frac{\mu^*}{(1-\eta)} (\langle u_i^{(i)} \rangle_{,j} + \langle u_j^{(1)} \rangle_{,i}) + \frac{\lambda^*}{(1-\eta)} \delta_{ij} \langle u_k^{(1)} \rangle_{,k} \tag{42}$$

Hence, from equations (5), (16), (20), and (42), our required constitutive equations are, to the lowest order in ω,

$$\langle \sigma_{ij} \rangle + \left(1 - \eta - \frac{\kappa_s^*}{\kappa_s}\right) \frac{\langle p \rangle}{\eta} \delta_{ij}$$

$$= \frac{\mu^*}{(1-\eta)} (\langle u_i \rangle_{,j} + \langle u_j \rangle_{,i}) + \frac{\lambda^*}{(1-\eta)} \delta_{ij} \langle u_k \rangle_{,k} \tag{43}$$

(*c*) Summary of the Full System of Equations. Collecting together equations (10), (11), (35), and (43), we now have the required full system of equations which will be used in Section 5 of this paper to relate the various coefficients appearing in Biot's equations (1) to the micro-structural properties of the rock we are considering. This full system of equations is as follows:

$$\frac{\langle \sigma_{kk} \rangle}{3\kappa_s} - \frac{\langle p \rangle}{\kappa_f} = \langle u_j \rangle_{,j} + \langle v_j \rangle_{,j} \tag{44}$$

$$\langle \sigma_{ij} \rangle_{,j} - \langle p \rangle_{,i} = -\rho_s \omega^2 \langle u_i \rangle - \rho_f \omega^2 \langle v_i \rangle \tag{45}$$

$$\langle p \rangle_{,i} = \rho_f \omega^2 \left\{ \chi \langle v_i \rangle + \frac{\eta(1-\chi)\langle u_i \rangle}{(1-\eta)} \right\} \tag{46}$$

$$\langle \sigma_{ij} \rangle + \left(1 - \eta - \frac{\kappa_s^*}{\kappa_s}\right) \delta_{ij} \frac{\langle p \rangle}{\eta}$$

$$= \frac{\mu^*}{(1-\eta)} (\langle u_i \rangle_{,j} + \langle u_j \rangle_{,i}) + \frac{\delta_{ij}\lambda^*}{(1-\eta)} \langle u_k \rangle_{,k} \tag{47}$$

The only "unknown" quantities appearing in the above system of equations are χ (or equivalently, K) and the effective Lamé frame elastic coefficients λ^* and μ^*. Now the effective Lamé coefficients λ^* and μ^* are known for dry packings. However, it must be remembered that the above equations apply to the case in which fluid is present, but it is still a relatively straight-forward matter to apply our earlier results (Digby, 1981; Walton, in press) here since the only effect of a uniform fluid pressure is to create a hydrostatic compression everywhere.

The quantity K_{ij} appearing in equation (33) could be determined numerically for a granular rock modeled by a spatially periodic packing of spherical particles (see, for example, the case of the simple cubic packing treated by Walton, 1977). In principle, it should also be possible to perform the same exercise for K or equivalently χ for random packings. We have also obtained a lower bound for the numerical factor χ, namely $\chi \geq 1$. It will be seen in the following section that this result

has important physical implications on the final results obtained from our full system of equations (44)–(47).

5 Results Obtained

(a) Comparison with Biot's Equations. From our basic equations (44)–(47) summarized in Section 4, we have, in particular, from equations (44), (47)

$$<p> = -\left\{\frac{(\kappa_s(1-\eta)-\kappa_s^*)}{(1-\eta)}\nabla\cdot<\mathbf{u}> + \kappa_s\nabla\cdot<\mathbf{v}>\right\}\frac{\eta\kappa_R\kappa_s}{(\kappa_s^2-\kappa_R\kappa_s^*)} \tag{48}$$

Where, in equation (48), κ_R denotes the usual Reuss averaged effective bulk modulus $1/\kappa_R = (\eta/\kappa_f) + ((1-\eta)/\kappa_s)$. From equations (47), (48) one then obtains

$$<\sigma_{ij}> = \frac{\mu^*}{(1-\eta)}(<u_i>_{,j} + <u_j>_{,i}) + \frac{\lambda^*\delta_{ij}}{(1-\eta)}<u_k>_{,k} + \frac{\delta_{ij}(\kappa_s(1-\eta)-\kappa_s^*)\kappa_R\kappa_s}{(\kappa_s^2-\kappa_R\kappa_s^*)} \times\left\{\frac{(\kappa_s(1-\eta)-\kappa_s^*)}{\kappa_s(1-\eta)}<u_k>_{,k} + <v_k>_{,k}\right\} \tag{49}$$

From equations (46) and (48) one then obtains

$$\left\{\frac{(\kappa_s(1-\eta)-\kappa_s^*)}{(1-\eta)}\nabla\nabla\cdot<\mathbf{u}> + \kappa_s\nabla\nabla\cdot<\mathbf{v}>\right\}\frac{\eta\kappa_R\kappa_s}{(\kappa_s^2-\kappa_R\kappa_s^*)} = -\rho_f\omega^2\left\{\frac{\eta(1-\chi)<\mathbf{u}>}{(1-\eta)} + \chi<\mathbf{v}>\right\} \tag{50}$$

Also, from equations (45), (46), and (49), we obtain

$$\frac{\mu^*\nabla^2<\mathbf{u}>}{(1-\eta)} + \frac{(\lambda^*+\mu^*)}{(1-\eta)}\nabla\nabla\cdot<\mathbf{u}> + \left\{\frac{(\kappa_s(1-\eta)-\kappa_s^*)}{\kappa_s(1-\eta)}\nabla\nabla\cdot<\mathbf{u}> + \nabla\nabla\cdot<\mathbf{v}>\right\}\frac{\kappa_R\kappa_s(\kappa_s(1-\eta)-\kappa_s^*)}{(\kappa_s^2-\kappa_R\kappa_s^*)} = -\omega^2\left\{\left(\rho_s + \eta\frac{(\chi-1)}{(1-\eta)}\rho_f\right)<\mathbf{u}> + \rho_f(1-\chi)<\mathbf{v}>\right\} \tag{51}$$

Regarded as two equations in the two unknown displacements $\mathbf{u}$ and $(1-\eta)<\mathbf{v}>/\eta$, equations (50) and (51) will in fact be identical to Biot's equations (1), discussed in the introduction to this paper, provided we make the following identifications:

$$\left.\begin{aligned} N &= \mu^* \\ A &= \lambda^* + \frac{\kappa_R(\kappa_s(1-\eta)-\kappa_s^*)^2}{(\kappa_s^2-\kappa_R\kappa_s^*)} \\ Q &= \frac{\eta\kappa_R\kappa_s(\kappa_s(1-\eta)-\kappa_s^*)}{(\kappa_s^2-\kappa_R\kappa_s^*)} \\ R &= \frac{\eta^2\kappa_R\kappa_s^2}{(\kappa_s^2-\kappa_R\kappa_s^*)} \end{aligned}\right\} \tag{52}$$

From which

$$A = \lambda^* + (Q^2/R)$$

and

$$\left.\begin{aligned} \rho_{11} &= (1-\eta)\rho_s + \eta(\chi-1)\rho_f \\ \rho_{12} &= \eta(1-\chi)\rho_f \\ \rho_{22} &= \eta\chi\rho_f \end{aligned}\right\} \tag{53}$$

Following Biot's notation (1956) we now also write, from equations (52)

$$\left.\begin{aligned} P &= A + 2N = \lambda^* + 2\mu^* + (Q^2/R) \\ H &= P + R + 2Q = \lambda^* + 2\mu^* + ((Q+R)^2/R). \\ \sigma_{11} &= P/H,\ \sigma_{12} = Q/H,\ \sigma_{22} = R/H \end{aligned}\right\} \tag{54}$$

from which

$$\sigma_{11} + \sigma_{22} + 2\sigma_{12} = 1$$

and from equations (53),

$$\left.\begin{aligned} \rho_{12} &= \eta(1-\chi)\rho_f = -\rho_a \\ \rho_1 &= (1-\eta)\rho_s \\ \rho_2 &= \eta\rho_f \\ \tilde{\rho} &= \rho_1 + \rho_2 \\ \gamma_{11} &= \rho_{11}/\tilde{\rho},\ \gamma_{12} = \rho_{12}/\tilde{\rho},\ \gamma_{22} = \rho_{22}/\tilde{\rho} \end{aligned}\right\} \tag{55}$$

from which

$$\gamma_{11} + \gamma_{22} + 2\gamma_{12} = 1$$

We note that our explicit expressions for Biot's constitutive coefficients in terms of the microstructural properties of the rock we are considering (equations (52)) are all positive. Our explicit expressions for Biot's dynamic coupling coefficients (equation (53)) also all have the correct sign since we have shown earlier that the numerical factor χ always satisfies the inequality $\chi \geq 1$. It is also easy to show that our expressions for Biot's coefficients, $\sigma_{11}\ \sigma_{12}\ \sigma_{22}$ and $\gamma_{11}\ \gamma_{12}\ \gamma_{22}$ (equations (54) and (55)) also satisfy the inequalities $\sigma_{11}\ \sigma_{22} - \sigma_{12}^2 \geq 0$ and $\gamma_{11}\gamma_{22} - \gamma_{12}^2 \geq 0$.

(b) The Longitudinal Wave Speeds. We have shown above that our averaged equations of motion (equations (50), (51)) are in fact identical to Biot's equations (1) discussed in the introduction to this paper, provided the identifications defined by equations (52)–(55) are made. All of Biot's analysis (1956) for the determination of the properties of the elastic waves and their propagation speeds may therefore be applied unchanged. Only a brief discussion will, therefore, be given here.

For the longitudinal wave speeds, for example, we write $(<\mathbf{u}>, <\mathbf{v}>) = (\mathbf{U},\mathbf{V})e^{i\mathbf{k}\cdot\mathbf{x}}$ where $(\mathbf{U},\mathbf{V}) = \hat{k}(U,V)$. That is, the displacement amplitudes are parallel to the wave vector $\mathbf{k}$. One can show, for example, exactly as in Biot's work (1956), that in general there are exactly two real P-wave speeds $\alpha = w/k$ $(k = |\mathbf{k}|)$, whose values are determined from the quartic equation

$$(\sigma_{11}\sigma_{22} - \sigma_{12}^2) - (\sigma_{22}\gamma_{11} + \sigma_{11}\gamma_{22} - 2\sigma_{12}\gamma_{12})(\alpha/V_c)^2 + (\gamma_{11}\gamma_{22} - \gamma_{12}^2)(\alpha/V_c)^4 = 0 \tag{56}$$

In equation (56), $V_c^2 = H/\tilde{\rho}$. All of the coefficients appearing in equation (56) are determined explicitly in terms of the microstructural properties of the rock we are considering through our equations (52)–(55). One can also obtain other results in Biot's original paper, but now expressed more explicitly. For example, by taking $U = ((1-\eta)/\eta)V$ we can obtain a more explicit form of Biot's so-called "dynamic compatibility condition" in which

$$\alpha^2 = V_c^2 = H/\tilde{\rho} = ((\lambda^* + 2\mu^*)/\tilde{\rho}) + ((Q+R)^2/R\tilde{\rho}) \tag{57}$$

Also, by putting $\rho_f = 0$ in equations (52)–(56), we obtain just one "effective wave speed," in which

$$\alpha^2 = V_c^2 = (\lambda^* + 2\mu^*)/\rho_s(1-\eta) \tag{58}$$

(*c*) **The Shear Wave Speeds.** Here, as earlier, we write $(<\mathbf{u}>, <\mathbf{v}>) = (\mathbf{U},\mathbf{V})e^{i\mathbf{k}\cdot\mathbf{x}}$, but now with $(\mathbf{U}, \mathbf{V}) = (U, V)\hat{k}_1$, where k_1, is perpendicular to the wave vector $\mathbf{k}$. From equations (50) one then obtains

$$\mathbf{V} = \frac{\eta(\chi - 1)\mathbf{U}}{\chi(1-\eta)} \tag{59}$$

Since $0<\eta<1$ and from Section 4 we have shown that $\chi \geq 1$, the fluid and solid displacement amplitudes in equation (59) are similarly directed. It then follows directly from equation (51) that there is always exactly one real *S*-wave speed $\beta = \omega/k$ given by

$$\beta^2 = \mu^* \Big/ \left[\rho_s(1-\eta) + \frac{\eta\rho_f(\chi-1)}{\chi}\right] \tag{60}$$

6 Conclusions

In this paper, equations have been derived in which both the constitutive coefficients and the dynamic coupling coefficients appearing in Biot's equations (1956), and hence also the effective longitudinal and shear wave speeds have been related explicitly to the microstructural properties of a fluid saturated rock. The rock was modeled by a random packing of identical, contacting spherical particles. The connected pore space was filled with an inviscid, compressible fluid. The constitutive coefficients are also expressed explicitly in terms of the bulk modulus of the pore fluid and the effective elastic moduli of the rock framework whose values are known from our earlier work (Digby, 1981; Walton, in press). The effect of the initial confining pressure to which the rock is subjected, as well as the effect of the nature of the contact regions between adjacent particles in the rock, can therefore also be included in our calculations. We believe that the work presented in this paper will provide a good theoretical basis for the study of more realistic problems in which finite frictional forces are also exerted across contact regions and the pore fluid is both compressible and viscous. Direct comparison of results from a study of this type with experiments should then be possible.

Acknowledgments

One of the authors, (P. J. Digby), would like to thank the Swedish Natural Research Foundation (NFR) for their financial support of the work described in this paper.

References

Biot, M. A., 1956, "The Theory of Elastic Waves in a Fluid Saturated Porous Rock – I. Low Frequency Range," *J. Acous. Soc. Am.*, Vol. 28, pp. 168–178.

Burridge, R., and Keller, J. B., 1981, "Poro-Elasticity Equations Derived from Micro-Structure," *J. Acous. Soc. Am.*, Vol. 70, pp. 1140–1146.

Digby, P. J., 1981, "The Effective Elastic Moduli of Porous Granular Rocks," A.S.M.E. Journal of Applied Mechanics, Vol. 48, pp. 803–808.

Walton, K., 1977, "Elastic Wave Propagation in Model Sediments – I," *Geophysical Journal of the Royal Astronomical Society*, Vol. 48, pp. 461–478.

Walton, K., "The Effective Elastic Moduli of a Random Packing of Spheres," *J. Mech. Phys. Solids*, to appear.

Walton, K., "Wave Propagation within Random Packings of Spheres," *Geophysical Journal of the Royal Astronomical Society*, to appear.

K. H. Parker

R. V. Mehta

C. G. Caro

Physiological Flow Studies Unit,
Imperial College,
London, SW7, England

Steady Flow in Porous, Elastically Deformable Materials

The steady, one-dimensional flow of an incompressible fluid through a deformable porous material is studied theoretically and experimentally. The theoretical model is essentially that of Biot. Assuming that the stiffness and permeability of the matrix are functions of the local strain gradient, the governing equations can be solved and analytical solutions are presented for several simple constitutive relationships. The stiffness and permeability properties of one particular foam are measured and then used to predict the rate of fluid flow and the distortion of the matrix as a function of the applied pressure difference across the material. Comparison of the predictions of the model with experimental observations indicates good qualitative agreement.

1 Introduction

When fluid flows through a porous, deformable medium there is coupling between the flow and the deformation of the medium which may lead to various interesting phenomena. For example, the strain distribution in a porous, deformable matrix under steady mechanical compression is uniform, but if the strain is produced by fluid flow, the strain distribution can be highly nonuniform (Caro et al. 1984) as can be seen in Fig. 1. In general, the nature of the flow and the resultant strain distribution depend upon the properties of the porous matrix and the fluid and their interactions as well as the boundary conditions: geometry, applied pressure, etc. In order to assess some of these effects, we have analyzed the behavior of the simple system: steady, one-dimensional flow through a slab of porous deformable material restrained at the downstream end by a freely draining rigid support. The results of this analysis have been compared with experimental results obtained for the flow of a glycerine-water mixture through a polyurethane foam.

The first theories of flow in porous, deformable materials were developed to explain the consolidation phenomena in soil mechanics (Terzaghi, 1925) culminating in the equations proposed by Biot (1955) which are widely accepted and used to explain observed behavior in geological materials. The fundamental basis of this theory has been studied both from the viewpoint of the thermodynamic theory of mixtures (Crochet and Naghdi, 1966); Rice and Cleary, 1976; and Kenyon, 1976a, 1976b) and from a two-scale analysis of the Navier-Stokes equations (Burridge and Keller, 1981). More recently, interest in flow in porous, deformable materials has been stimulated by problems in biomechanics. Both the response of articular cartilage to loading (Lai and Mow, 1980; Lai et al. 1981) and the flux of water through arterial tissue (Kenyon, 1979; Jayaraman, 1983) have been analyzed in these 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, September 22, 1986; final revision February 20, 1987.

2 Mathematical Model

(a) Basic Equations. As the starting point of this analysis we shall take the consolidation equations of Biot (1955) with the definitions and interpretations of Kenyon (1976a, 1976b), who has derived these equations as a special case of his theory of isothermal solid-fluid mixtures. In this continuum approach the solid matrix and the fluid coexist, both the solid stress and the fluid pressure are continuous properties representing some "pore averaged" value, and the fluid velocity is measured relative to the solid matrix. The solid is described by constitutive equations which relate the stiffness, Λ, and porosity, K, of the matrix to the local strain gradient, R. For one-dimensional problems, the balance of forces and Darcy's law lead to the dimensionless equations

$$\frac{\partial(\lambda R)}{\partial x} = \frac{\partial p}{\partial x} = -\frac{w}{k} \tag{2.1}$$

The height, stiffness, and permeability of the matrix in the reference state have been chosen as characteristic values so that $\lambda = \Lambda_c$ is the dimensionless stiffness, $k = K/K_c$ the dimensionless permeability, $p = P/\Lambda_c$ the dimensionless pressure, and $w = \mu H_c W/K_c \Lambda_c$ is the dimensionless fluid velocity where μ is the fluid viscosity. For an incompressible solid-fluid mixture $w = -\partial R/\partial t$ where $t = TK_c\Lambda_c/\mu H_c^2$ is the dimensionless time. Here and throughout, capital letters refer to dimensional parameters, small letters refer to the corresponding dimensionless parameters and the subscript c refers to characteristic dimensional values. Combining these equations gives the familiar diffusion equation for the local strain gradient

$$\frac{\partial}{\partial x}\left(k(R)\frac{\partial}{\partial x}\left(R\lambda(R)\right)\right) = \frac{\partial R}{\partial t} \tag{2.2}$$

(b) Boundary Conditions. In the following, we will consider the simplest case of a homogeneous matrix of undeformed thickness H_0 restrained at $x = 0$ by a rigid grid which

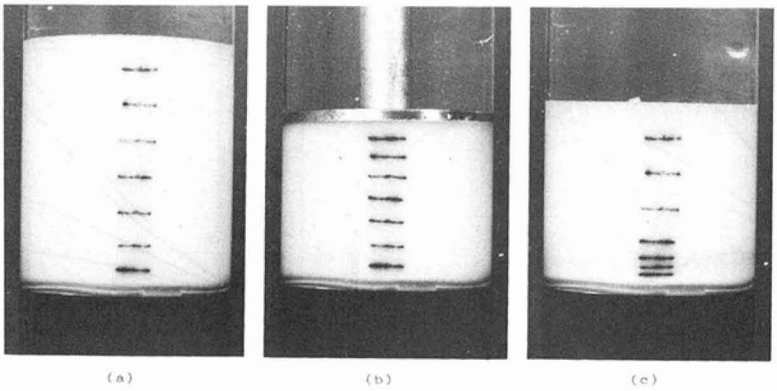


Fig. 1 Strain distribution in a porous, deformable matrix under different conditions: (a) no load, no flow, showing the foam in its relaxed state; (b) mechanical compression with no flow, showing uniform strain distribution; (c) the nonuniform strain distribution resulting from flow

does not impede the flow. We will assume the pressure in the fluid at $x=0$ is maintained at a constant level which, without loss of generality, we will take to be zero. At the top of the matrix, $x=h$, we will assume that both the pressure in the fluid and the contact stress on the solid matrix are prescribed. The dimensionless boundary conditions are

$$\begin{aligned} x &= 0: \quad \lambda R = s_1 + p_1 \\ x &= h: \quad \lambda R = s_1 \end{aligned} \tag{2.3}$$

where $p_1 = P_1/\Lambda_c$ and $s_1 = S_1/\Lambda_c$ are the pressure and the contact stress applied at the top of the matrix. The location of the top of the matrix is, of course, not known a priori and the calculation of h under different conditions is one of the primary goals of the analysis. From the definition of strain, however, we have the "matching" condition

$$\int_0^h R dx = 1 - h \tag{2.4}$$

These equations together with an initial condition describe a well posed, free boundary problem for the local strain, R.

(c) Steady State. If the applied stresses, p_1 and s_1, are constant, then the problem is steady and equations (2.1), (2.3), and (2.4) are sufficient to determine the constants w and h and the distribution of strain $R(x)$ as a function of the applied stresses p_1 and s_1.

3 Results

In principle, solutions are possible for any sufficiently well-behaved functions of permeability and stiffness, $k(R)$ and $\lambda(R)$; but in practice, analytical solutions are obtainable only for relatively simple constitutive relations. In order to explore some of the interesting phenomena of flow in deformable, porous matrices, we will look at the solutions of these equations for several simple relationships, in the knowledge that in real materials both k and λ can be quite complex functions of the local distortion.

(a) Zero Flow. If $p_1 = 0$, then $R = s_1$, is a solution of equations (2.1) and (2.3). Substituting into equation (2.4) and integrating we obtain

$$h = (1 + s_1/\lambda)^{-1} \tag{3.1}$$

In its dimensional form, this solution implies that the strain is uniform throughout the matrix thus providing a convenient way of measuring $\Lambda(R)$ experimentally. By applying different loads, S_1, and measuring the resultant height of the matrix, $\Lambda = S_1/R$ can easily be calculated since for constant strain, $R = (H_0 - H)/H$.

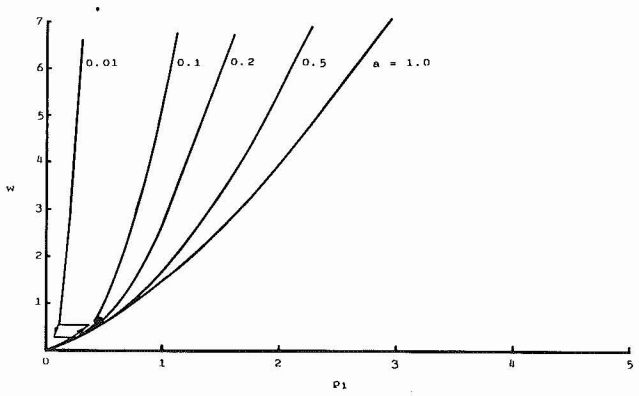


Fig. 2 w versus p_1 calculated for constant permeability and exponential stiffness, equation (3.6), with $b=1$ and various values of a. The curves for $a=0.01$ and 0.1 exhibit the hysteresis resulting from the multi-valued nature of the assumed stress-strain law. The limiting case, $a=1$, is the solution for constant stiffness and permeability, equation (3.8).

For the rest of this section, we will assume that $s_1 = 0$, i.e., there is no contact stress at the top of the matrix, and explore the nature of the solutions with flow.

(b) Constant Permeability: $k=1$. If the permeability of the matrix does not vary as the matrix is distorted, $k=1$. For this case equations (2.1) and (2.3) have the solution

$$\lambda R = p_1(1 - x/h) \tag{3.2}$$

and $w = p_1/h$. The thickness of the matrix is determined from the matching condition,

$$h = 1 + \int_0 R dy^{-1} \tag{3.3}$$

where $y = x/h$ is the distance scaled by the deformed matrix thickness.

(*i*) *Constant Stiffness.* For the special case of constant matrix stiffness, $\lambda = 1$, and the solution becomes

$$h = (1 + p_1/2)^{-1} \tag{3.4}$$
$$w = p_1(1 + p_1/2) \tag{3.5}$$

This solution corresponds to the incompressible case discussed by Keynon (1978) and is shown as the limiting case, $a=1$, in Fig. 2.

(*ii*) *Exponential Stiffness.* Consider the exponential stiffness law

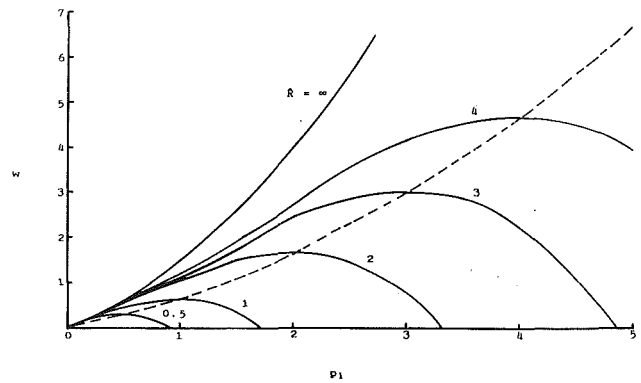


Fig. 3 w versus p_1 calculated for constant stiffness and linear permeability, equation (3.11), for different values of the critical strain $\hat{R}$. The curves have a maximum value $\hat{R}(1+\hat{R}/3)/2$ at $p_1=\hat{R}$. The limiting case, $\hat{R}=\infty$, is the solution for constant stiffness and permeability.

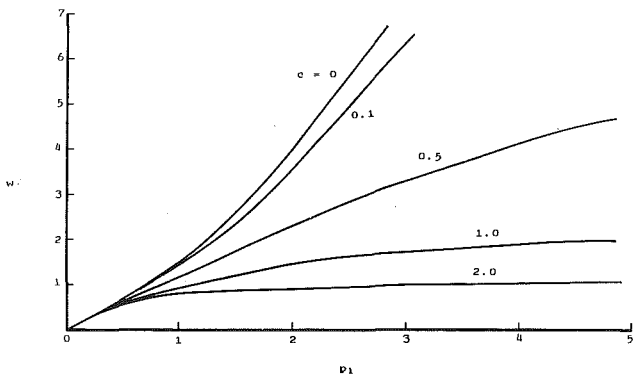


Fig. 4 w versus p_1 calculated for constant stiffness and the exponential permeability, equation (3.16), for different values of c. The curves are monotonic with the asymptote $(1+c)/c^2$. The limiting case, $c=0$. is the solution for constant stiffness and permeability.

$$\lambda = a + (1-a)\exp(-bR) \tag{3.6}$$

where a is the limiting stiffness at large compaction. From the boundary conditions, the strain varies from 0 at the top of the matrix to R_0 at the bottom, where R_0 is the solution of $R_0\lambda(R_0)=p_1$. For values of $a<(e^2+1)^{-1}$ care must be taken because there is a range of pressures for which there is more than one solution for R_0. Such behavior could give rise to a sudden "collapse" of the matrix as the pressure is increased beyond some critical value. It could also lead to hysteresis since the collapsed matrix would not recover as the pressure is lowered until another, lower, critical pressure was reached.

The integral in equation (3.3) has the solution

$$\int_0^1 R\,dy = \frac{1}{p_1}\left[\frac{1}{b^2}\left(\frac{p_1}{R_0}-1\right)-\frac{R_0}{b}\left(a-\frac{p_1}{R_0}\right)+R_0^2\left(\frac{p_1}{R_0}-\frac{a}{2}\right)\right] \tag{3.7}$$

The resultant velocity is plotted as a function of p_1 for different values of a in Fig. 2. The effect of the hysteresis is seen at the two lowest values of a.

(c) Constant Stiffness: $\lambda=1$. The results in the previous section were calculated for the special case of constant permeability. We now relax that condition and calculate the effect of nonconstant permeability under the simplifying assumption of constant stiffness, $\lambda=1$. For this case we must solve

$$k(R)\frac{dR}{dx} = -w \tag{3.8}$$

$$x=0,\ R=p_1 \tag{3.9}$$

$$x=h,\ R=0 \tag{3.10}$$

(*i*) *Constant Permeability.* This case was discussed above with h and w given by equations (3.4) and (3.5), respectively. The predicted flow as a function of the applied pressure is shown as the limiting behavior in Fig. 2.

(*ii*) *Linear Permeability.* Assume that the permeability varies linearly with local strain,

$$k(R) = 1 - R/\hat{R} \tag{3.11}$$

for $R<\hat{R}$, the critical strain at which the permeability becomes zero. The solution for this case is,

$$R = \hat{R}\left[1-\left(1-\frac{2wh}{\hat{R}}\left(1-\frac{x}{h}\right)\right)^{1/2}\right] \tag{3.12}$$

where

$$w = p_1\left(1+\frac{p_1(\hat{R}-1)}{2\hat{R}}-\frac{p_1^2}{3\hat{R}}\right) \tag{3.13}$$

and

$$h = \frac{p_1}{w}\left(1-\frac{p_1}{2\hat{R}}\right) \tag{3.14}$$

The fluid velocity for this case is shown in Fig. 3. One of the interesting features of this solution is that the fluid velocity goes through a maximum as the pressure is increased. The maximum velocity depends upon the critical strain, $\hat{R}$,

$$w_{\max} = \hat{R}(1+\hat{R}/3)/2 \tag{3.15}$$

and occurs at $p_1=\hat{R}$. As the pressure is increased further, the velocity decreases, eventually reaching zero. This limit is probably not very realistic since real materials will deviate from the linear permeability law at large deformation. However, the flow limitation properties implied by the maximum are more realistic since they occur at lower levels of distortion of the matrix where the linear variation of permeability with distortion may be a reasonable model of the behavior of the matrix.

(*iii*) *Exponential Permeability.* A constitutive relation for the permeability which avoids the problem of a critical strain at which the material becomes impermeable is the exponential relation

$$k = \exp(-cR) \tag{3.16}$$

This is the law proposed by Lai et al. (1981) for the flow of synovial fluid through articular cartilage. With this law, the solution is,

$$R = -\frac{1}{c}\ln\left(1-cwh(1-x/h)\right) \tag{3.17}$$

where

$$w = \frac{(1+c)}{c^2}\left(1-\exp(-cp)\right)-\frac{p_1}{c}\exp(-cp) \tag{3.18}$$

and

$$h = \left[\frac{c+1}{c}-\frac{p_1\exp(-cp)}{1-\exp(-cp)}\right]^{-1} \tag{3.19}$$

The variation of fluid velocity with pressure is shown in Fig. 4. In this case the curves are monotonic but approach an asymptotic maximum

$$w_{\max} = \frac{1+c}{c^2} \tag{3.20}$$

4 Experimental Measurements

The porous, elastically deformable material used in the ex-

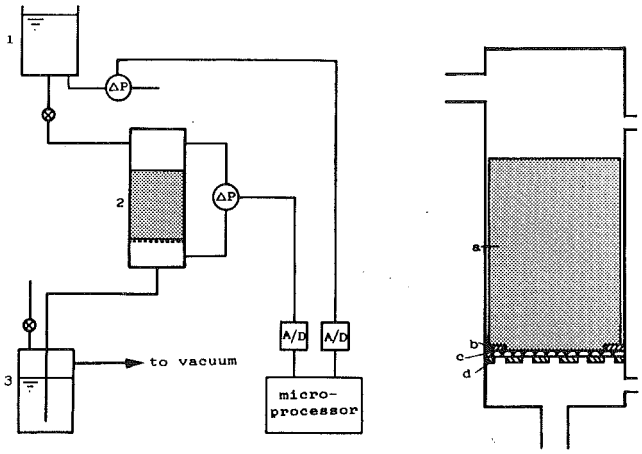


Fig. 5 A sketch of the experimental apparatus: (1) upstream reservoir; (2) test section; (3) downstream reservoir; (*a*) polyurethane foam; (*b*) plastic clay gasket; (c) wire mesh; (*d*) freely draining rigid grid

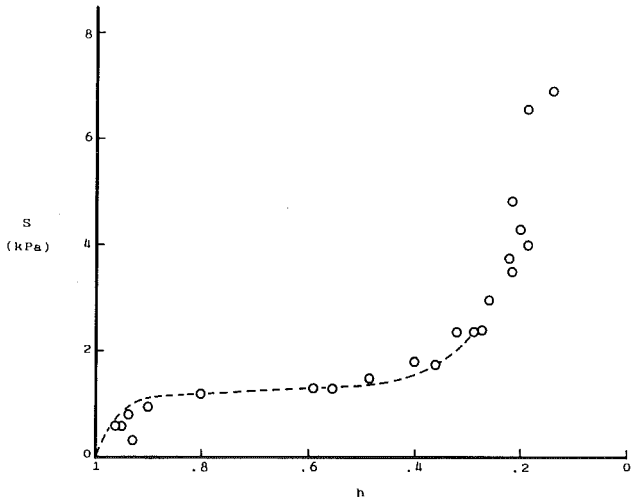


Fig. 6 Results of uniaxial compression tests on the polyurethane foam plotted as stress versus height. Broken line corresponds to data for similar polyurethane foams obtained by Beavers et al. (1981b).

periments was a highly flexible, open-celled polyurethane foam marketed by Declon, England. The foam had a nominal pore distribution of 100 pores per linear inch with a corresponding nominal pore diameter of 150 μm. The void fraction of the foam was about 0.97 and it had a bulk density of 27 kg/m^3. When observed under a dissecting microscope, the structure of the foam was seen to be lattice-like with polyurethane fibres as the members of the lattice and no free fibre ends internal to the structure. This foam was chosen as matrix material because it showed acceptable variations in its poroelastic properties from sample to sample. Although inhomogeneous, it was more uniformly porous than other candidate materials. An 80:20 glycerine-water mixture at room temperature was used as the fluid medium. The viscosity of this medium was 40 to 50 times higher than the viscosity of water at the same temperature. As a result, we were able to induce relatively large drag in the foam at much lower flow rates while operating in the Darcy flow regime. In this regime, the pore Reynolds number (that is, the Reynolds number based upon a typical pore dimension) is less than 1.

A schematic diagram of the experimental apparatus is shown in Fig. 5. A 650 cm^3 Perspex cylinder served as the feed reservoir which was connected to the test section by 9 mm ID Tygon tubing. The test section, Fig. 5 inset, was a 12 cm high Perspex cylinder 5.1 cm in diameter. It retained the polyurethane foam on a perforated plate having 7 mm diameter perforations on a 1 cm triangular pitch. In order to prevent any bulging of the foam through the perforations, an ordinary wire mesh, 2 mm square mesh size, was placed between the foam and the perforated plate. The pressure drop across the mesh and the plate was negligible compared to that across the foam itself under experimental conditions. Pressure taps were installed 1 cm above the upstream face and 1 cm below the downstream face of the uncompressed foam whose height was 7.6 cm. The diameter of the holes used for the taps was 6 mm. The taps were connected by Tygon tubing to a 175 mbar differential pressure transducer (Druck Ltd, England). The feed solution entered the test section at the same level as the upper pressure tap and left it at the same level as the lower pressure tap through 6 mm ID plastic tubes. The Reynolds number based on the test section diameter was less than 1 under experimental conditions and no entrance or exit effects were expected. On the one hand it was essential to minimize friction between the foam and the test section and on the other hand it was important to minimize flow around the sides of the foam. After careful experimentation, we used a slightly undersized foam and reduced the amount of flow around the foam to an acceptable minimum by placing a plastic clay gasket, 3 mm in width, at the downstream face of the foam (Fig. 5). The effluent liquid from the test section was collected in a 5000 cm^3 glass bottle which served as the downstream reservoir. The glass bottle was connected via a variable leak to a vacuum pump so that the downstream pressure could be adjusted to the desired level.

The feed flow rate or its equivalent, the rate of change of the height of the feed solution in the upper reservoir, was monitored by measuring the pressure at the base of the reservoir with another 175 mbar differential pressure transducer. The outputs from the two pressure transducers were passed through signal conditioning modules to a BBC microprocessor for on-line data collection. Signals were sampled every 100 msec and displayed simultaneously on the screen. For a given experiment, 180 such samples were collected for each signal and stored. The stored data were processed later to obtain the corresponding pressure drop, ΔP, and flow rate, Q.

Preliminary tests indicated that when a new foam was used, its poroelastic properties varied with the number of times it had been compressed. After it had been compressed several times, however, its properties became independent of its history reflecting a mature, unchanging state. Similar findings were reported by Beavers et al. (1981b). We have used such a well conditioned foam in all our experiments to eliminate history dependence. Three different sets of experiments were performed to characterize the foam and test the mathematical model. These were (1) measurement of stiffness, Λ, as a function of strain, R; (2) measurement of permeability, k, as a function of strain; and (3) experiments concerned with the steady behavior of the foam involving measurements of flow rate, Q, and overall compaction, H/H_c as a function of pressure drop, ΔP. The principal features of each set of experiments are described below.

(a) Stress Versus Strain. The foam was placed in the test section and immersed in the test fluid. In order to ensure that the matrix was air-free, the air filled test section together with the foam was placed under vacuum and the degassed glycerine solution was then allowed to rise slowly through the foam. Once the foam was completely filled with the solution, the vacuum was released and care was taken not to expose the foam to air. A rigid grid was placed on the top of the foam and the deflection of the foam under different loads was measured. The poroelastic nature of the system was reflected by the finding that the deflection for a fixed load increased slowly to its equilibrium value over a very long period of time

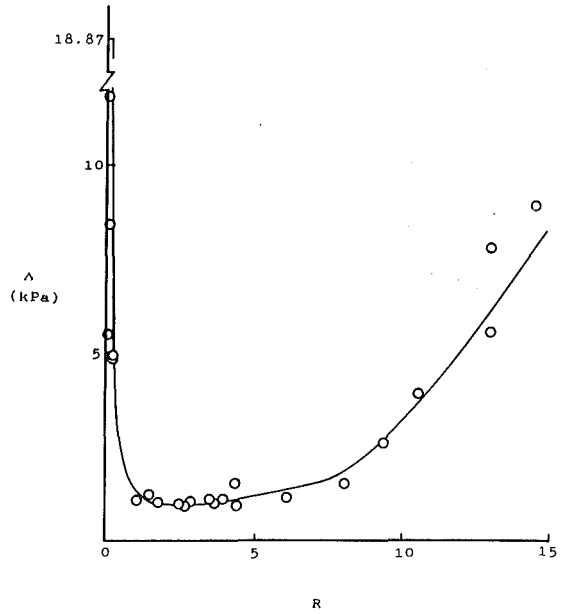


Fig. 7 The uniaxial compression data of Fig. 6 plotted as stiffness versus strain. The solid line is the fitted curve used in the numerical calculations (equation (5.1)).

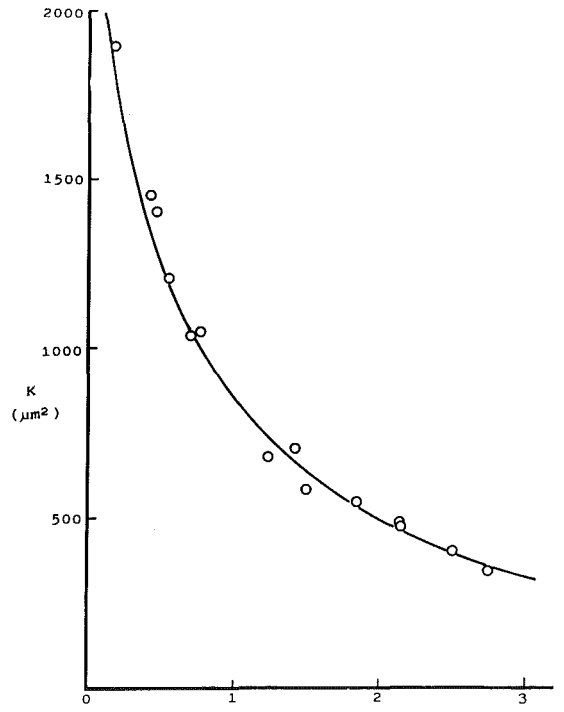


Fig. 8 The measured permeability of the polyurethane foam used in the experiments versus strain. The solid line is the fitted curve used in the numerical calculations (equation (5.2)).

even though it increased quite rapidly in the initial period after loading. We found that at least 80 percent of the equilibrium deflection was achieved within the first four hours. The stress versus height data shown in Fig. 6 correspond to values at the end of that four hour period. The stress values are accurate to within ±2 percent and the heights to within ±1 percent. Also included in the figure are data reported by Beavers et al. (1981b) for a similar foam in water. The same data are shown in Fig. 7 in the form Λ versus R along with the fitted curve

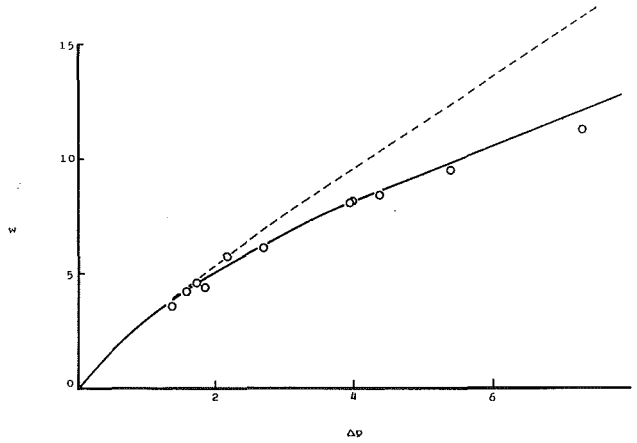


Fig. 9 The predicted flow versus applied pressure. The symbols represent experimental data; solid line corresponds to prediction for experimental stiffness and broken line corresponds to prediction for hypothetical stiffness (equation (6.1)).

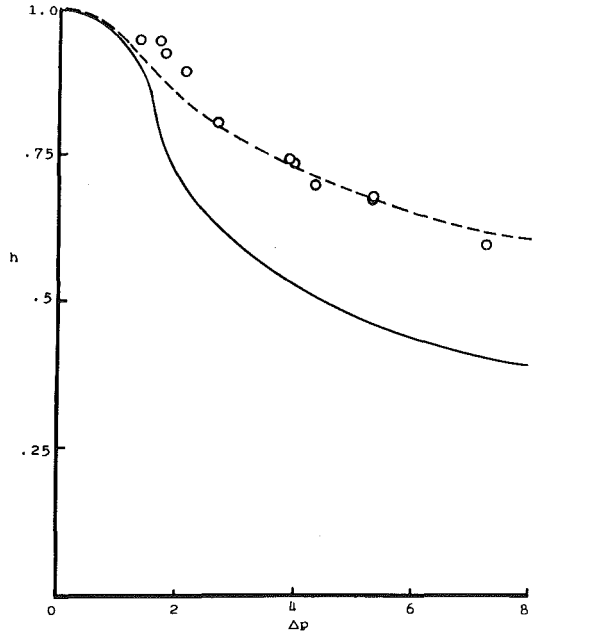


Fig. 10 The predicted height of the matrix versus applied pressure. The symbols represent experimental data; solid line corresponds to prediction for experimental stiffness and broken line corresponds to prediction for hypothetical stiffness (equation (6.1)).

which was used in the related numerical solution of equation (3.1).

(*b*) Permeability Versus Strain. The measurement of the permeability of the foam involved constraining the foam between two porous grids at the desired strain and measuring the flow rate as a function of the pressure drop across the foam. Since the flow induced a nonuniform distribution of local strain within the foam matrix even at small pressure drops, measurements were made at several different pressure drops and the permeability was calculated from the initial slope of the measured flow rate versus the pressure drop curve. The permeability measured thus was accurate to within ±6 percent. The results, plotted as a function of the imposed strain, are shown in Fig. 8 in the form K versus R along with the fitted curve which was used in the numerical calculations.

(*c*) Steady State Behavior. The steady state behavior of

the foam, constrained only at the bottom end, was studied at different flow rates. The flow rate was varied by changing the pressure at the downstream end by applying a suitable vacuum. The height of the foam as well as the flow rate and pressure drop were continuously monitored until steady state conditions were attained. At least two pore volumes of solution were allowed to flow through the foam before data collection was begun. The measured steady flow rate for several pressure drops is plotted in Fig. 9 in nondimensional coordinates. The corresponding steady state height is shown against the applied pressure drop in Fig. 10.

5 Theoretical Predictions

Both the stiffness, Λ, and permeability, K, had strong and complex dependence on the strain as shown in Figs. 7 and 8. Therefore, an analytical solution of equations (3.1) - (3.3) was not possible and appropriate analytical expressions were fitted to the poroelastic properties data for use in the numerical solution. Polynominals were fitted to the stress-strain data over three piecewise continuous ranges of strain:

$$\begin{aligned} 0 \le R \le 0.065,\ \lambda &= 18.87/\Lambda_c \\ 0.065 \le R \le 7,\ \lambda &= (0.2\mathrm{R}^2 - 0.1\mathrm{R} + 1.23)/(R\Lambda_c) \\ 7 \le R,\ \lambda &= (1.44\mathrm{R}^2 - 17.52\mathrm{R} + 62.41)/(R\Lambda_c) \end{aligned} \quad (5.1)$$

The following expression was fitted to the permeability data:

$$k(R) = (3103/K_c)\ \exp(-1.289\sqrt{R}) \quad (5.2)$$

The undeformed height of the matrix, H_0, was used as the characteristic length. For better accuracy, Λ_c and K_c were chosen to correspond to their values at $\mathrm{R} = 1$ ($\Lambda_c = 1.33$ kPa, $K_c = 855\ \mu\mathrm{m}^2$). The model equations were then solved numerically using a second order Runge-Kutta integration scheme. Since it was a free boundary problem, solutions were found by iterating on height until all the boundary conditions were satisfied.

The solid lines in Figs. 9 and 10, respectively, represent the predicted fluid velocity and predicted matrix height while the symbols represent experimental data.

6 Discussion

The agreement between experiments and theoretical predictions of fluid velocity is good considering the fact that there are no free parameters in the theory. On the other hand, the predicted height of the matrix is as much as 35 percent below the experimental measurements. This may be an experimental artifact or could be an indication of the limitation of the proposed theory. This point merits more discussion.

It is possible that the pressure-flow relationship was dominated by the most compressed region of the foam very near to its constrained end while the overall height of the matrix was largely determined by the less strained upper section within which a nonuniform distribution of strain existed. Since the stress-strain curve for this foam shows a wide plateau from $R = 0.065$ to about $R = 3.000$, a small discrepancy in local stress would correspond to a large error in estimated R. The cumulative error over the entire matrix height could well account for the extent of disagreement with the experimental results. To test this hypothesis, we changed the stress-strain curve in the plateau region and beyond to the following:

$$\begin{aligned} 0.065 \le R \le 7,\ \lambda &= 0.2(0.2\mathrm{R}^2 - 0.1\mathrm{R} + 1.23)/(R\Lambda_c) \\ 7 \le R\ ,\ \lambda &= 0.2(1.44\mathrm{R}^2 - 17.52\mathrm{R} + 62.41)/(R\Lambda_c) \end{aligned} \quad (6.1)$$

The resultant prediction of matrix height, shown by the broken line in Fig. 10, is in close agreement with the experimental results. The corresponding predicted fluid velocity, however, is much higher than the experimental result. Apparently, it is difficult to reconcile theory and experiments completely. We speculate that incorporation of the strain gradient, dR/dX, as an additional independent parameter might shed more light on the complex interaction of matrix and flow seen in the present experiments.

There are few previous studies with which this work can be compared. Manins and Roberts (1975) used a similar model with a power law for the permeability to calculate numerical solutions for the unsteady, two-dimensional problem of a Hookean elastic matrix confined in a rectangular box with one porous side. They studied the case of loading from the top with no through flow which makes comparison to this work difficult. More directly comparable are the studies of Beavers et al. (1975, 1981a, 1981b) who investigated the flow of air and of water through polyurethane foams. Our results for the elastic properties of the foam in the glycerine solution are very similar to their results on "conditioned" foams in water. They also observed that the polyurethane foams they used tended to undergo reduction in cross-sectional area with increasing compression. We did not detect such behavior in our experiments. The major difference between our study and theirs is in the modelling and analysis. They assumed a nonlinear Darcy law in which the pressure drop had a quadratic dependence on the fluid velocity and included a threshold pressure gradient below which no flow could occur. We have assumed a linear Darcy law without any threshold pressure gradient and yet our model is able to predict the nonlinear pressure flow relationship of the flexible polyurethane foam very well.

7 Concluding Remarks

The simple model presented in this paper is essentially Biot's theory (Biot, 1955) in Kenyon's formalism (Kenyon, 1976a, 1976b) and it appears to provide a reasonable description of the pressure-flow relationship in porous, deformable media. The model prediction of overall strain in the medium lacks quantitative accuracy although it is qualitatively consistent with experimental observations.

The present findings suggest that depending on the kinematic properties of the fluid and the poroelastic properties of the deformable matrix, the system can be used to perform some useful tasks. For example, the nonuniform distribution of strain and permeability within the matrix could be useful in some separation processes, perhaps using the flexibility of the matrix material to modify the filtration properties or to allow for more efficient cleaning of the filter medium. The analytical solutions presented in Section 3 also suggest useful flow-control applications such as flow limitation and, given suitable stress-strain properties, the possibility of hysteresis in the pressure-flow relationship. Finally, fluid drag may not always lead to compaction. If the solid matrix is prestressed by constraining it between two freely draining rigid grids, the porous structure will expand on the upstream end and be further compressed downstream when a flow passes through the matrix. This suggests yet another useful application of the drag induced behavior of a porous, deformable medium.

Acknowledgment

This work was supported by the B.P. Venture Research Unit and British Petroleum, plc.

References

Beavers, G. S., Wilson, T. A., and Masha, B. A., 1975, "Flow Through a Deformable Porous Material," ASME Journal of Applied Mechanics, Vol. 42, pp. 598-602.

Beavers, G. S., Hajji, A., and Sparrow, E. M., 1981, "Fluid Flow Through a Class of Highly-Deformable Porous Media: Part I—Experiments With Air," *ASME Journal of Fluid Engineering*, Vol. 103, pp. 432-439.

Beavers, G. S., Wittenberg, K., and Sparrow, E. M., 1981, "Fluid Flow Through a Class of Highly-Deformable Porous Media: Part II—Experiments With Water," *ASME Journal of Fluid Engineering*, Vol. 103, pp. 440–444.

Biot, M. A., 1955, "Theory of Elasticity and Consolidation for a Porous Anisotropic Solid," *J. Appl. Phys.*, Vol. 26, pp. 182–185.

Burridge, R., and Keller, J. B., 1981, "Poroelasticity Equations Derived from Microstructure," *J. Acoust. Soc. Am.*, Vol. 70, pp. 1140–1146.

Caro, C. G., Lever, M. J., Morgan, G. E., Parker, K. H., and Winlove, C. P., 1984, "Studies of Transport in Deformable Porous Materials With Relevance to Connective Tissue Properties," *J. Physiol.*, Vol. 358, p. 3.

Crochet, M. J., and Naghdi, P. M., 1966, "On Constitutive Equations for Flow of Fluid Through an Elastic Solid," *Int'l. J. Engng. Soc.*, Vol. 4, pp. 383–401.

Jayaraman, G., 1983, "Water Transport in the Arterial Wall—A Theoretical Study," *J. Biomechanics*, Vol. 16, pp. 833–840.

Kenyon, D. E., 1976a, "Thermostatics of Solid-Fluid Mixtures," *Arch. Ration. Mech. Analysis*, Vol. 62, pp. 117–129.

Kenyon, D. E., 1976b, "The Theory of an Incompressible Solid-Fluid Mixture," *Arch. Ration. Mech. Analysis*, Vol. 62, pp. 131–147.

Kenyon, D. E., 1979, "A Mathematical Model of Water Flux Through Aortic Tissue," *Bull. Math. Biol.*, Vol. 41, pp. 79–90.

Lai, W. M., and Mow, V. C., 1980, "Drag-Induced Compression of Articular Cartilage During a Permeation Experiment," *Biorheology*, Vol. 17, pp. 111–123.

Lai, W. M., Mow, V. C., and Roth, V., 1981, "Effects of Nonlinear Strain-Dependent Permeability and Rate of Compression on the Stress Behavior of Articular Cartilage," *J. Biomech. Eng.*, Vol. 103, pp. 61–66.

Manins, P. C., and Roberts, B. W., 1975, "Compression of an Elastoporous Medium," *Int. J. Non-Linear Mechanics*, Vol. 13, pp. 75–93.

Rice, J. R., and Cleary, M. P., 1976, "Some Basic Stress Diffusion Solutions for Fluid-Saturated Elastic Porous Media With Compressible Constituents," *Rev. Geophys. Space Phys.*, Vol. 14, pp. 227–241.

Terzaghi, K., 1925, *Erdbaumechanik auf Bodenphysikalischen Grundlagen*, Wien, Deuticke.

J. S. Patton
George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332
Assoc. Mem. ASME

C. E. Brennen
Mem. ASME

R. H. Sabersky
Fellow ASME

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

Shear Flows of Rapidly Flowing Granular Materials

Shear flows of granular materials are studied in an open channel. The wall shear is calculated from an open channel momentum equation which includes the density variations in the flow. An experimental technique was developed that allowed the measurement of the average density of the flow at different longitudinal locations in the channel. Two sizes of glass beads are examined and results show the variations in the wall shear as a function of various dimensionless parameters.

I Introduction

In both industrial and commercial applications, large amounts of material are transported in granular form. This includes the handling of such substances as coal, metal ores, shale, dry chemicals, and grain. In addition, flowing granular streams are being considered for some advanced concepts for solar power plants and fusion reactor chambers. In order to design the equipment for these varied applications in an effective and economical way, it is necessary to obtain a thorough understanding of the flow characteristics of granular materials. These design needs have already motivated extensive analytical and experimental investigations of granular flows. At this time, however, there is still no clear understanding of the constitutive relations that govern the motion of granular materials. The general field is still in a stage of development comparable to that of fluid mechanics before the advent of the Navier-Stokes relations. The present work was designed to contribute information which may prove helpful in the eventual formulation of a constitutive law. The data also provides practical information on the friction due to the flow of granular materials over a smooth wall.

II Review of Related Work

R. A. Bagnold is credited with the development of the modern research in granular material flows with his experiments and theories dating to the early 1950's (Bagnold, 1954, 1956, 1966). Recent progress has been described in an excellent review by Savage (1984). The reader is referred to this work and no attempt will be made here to offer any extensive account of the background material. For our purposes, perhaps the most pertinent recent work is that performed by Savage (1979, 1983), Sayed and Savage (1983), Hanes and Inmann (1985), Craig et al. (1986), and Bailard (1978) all of whom used Couette flow devices or open channels to experimentally study the shear of granular materials. With the aid of Couette flow devices, the behavior of the shear stress as a function of the shear rate, the normal pressure, and the particle size and density was studied. Sayed and Savage (1983), Augenstein and Hogg (1974, 1978), and Bailard (1978) were able to compute velocity and density profiles for the flow along an inclined chute based on certain constitutive assumptions. Also applicable is the work by Campbell et al. (1985a) on granular flow in an inclined chute. This experimental work yielded some preliminary results on the shear stress in open channel flows; however, their study did not account for the density changes of the flow in the channel. The present study shows that these density changes play an important role in adequately describing a granular material flow. The computational work carried out by Campbell (1982), and Campbell and Brennen (1985b), involved a statistical analysis of an assembly of particles flowing down a chute. The results give some indication of the magnitude and distribution of velocity and density, as well as the fluctuational components of the flow field, and may serve as background information for the eventual formulation of the constitutive laws.

III Experimental Installation

The present investigation was designed to obtain further information on the parameters that influence the shear in a flowing granular material. For simplicity of analysis, an open channel was proposed as the test section. With this purpose in mind, a large installation was constructed which would allow for continuous flow in a relatively wide open channel. A wide channel was selected so that the effect of the side walls could be minimized and a continuous operating loop allowed proper adjustment of the flow as well as more accurate measurements of the flow quantities by eliminating the time constraints on

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April 17, 1986; final revision May 7, 1987.

Paper No. 87-WA/APM-27.

the measurements. An installation was built consisting of a 3 m long, 150 mm wide channel which could be placed at angles up to 40 deg from the horizontal position. A high-speed mechanical conveyor delivers the material to an upper hopper from which the material flows into the channel. The discharge from the channel is collected in a lower hopper which feeds the material to the conveyor (See Fig. 1).

The measurements taken were the mass flow rate, profiles of the depth of the flowing stream, and the local density. The mass flow rate was determined by monitoring the rate of depletion of the upper feed hopper through a graduated transparent panel. Some confirmation of these flow rates were obtained by collecting and weighing the material discharging from the chute in a given amount of time. The depth of the flow in the channel was measured at several points along the chute by means of point depth probes identical to those commonly used in open-channel hydraulics. A simple yet effective method was developed to measure an average density of the flowing material. A device consisting of two plates connected by a handle was suddenly pushed into the flow thereby trapping the flowing material in the space between the plates. The trapped material was then collected and weighed, and the average density of the original stream was computed from this weight, the measured depth of the original stream, and the dimensions of the trap. For any given flow, this procedure could be repeated at different locations along the channel which allowed the density gradients to be evaluated in the direction of flow. From such density measurements, meaningful average velocities and densities could be computed. The ability to obtain an indication of the average density in this way was an important factor in allowing a more realistic interpretation of the data for open channel granular material flows. A drawing of the channel, point probes, and density device may be seen in Fig. 2.

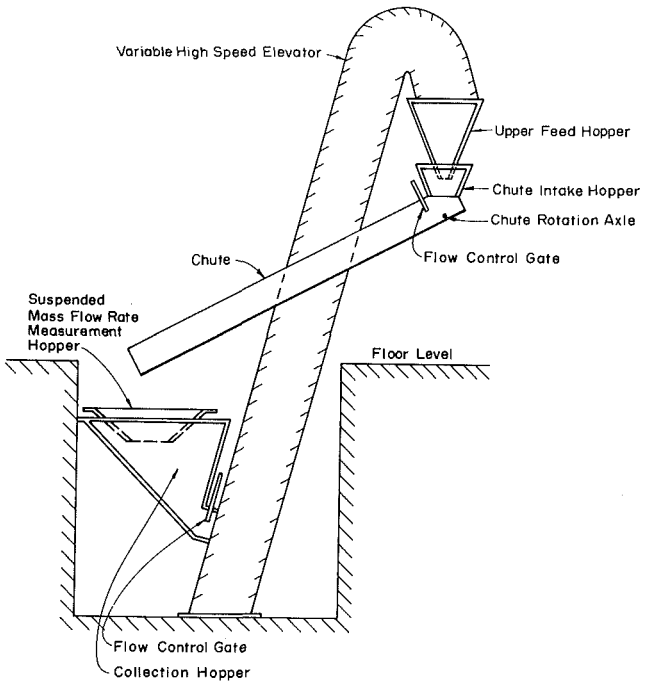


Fig. 1 Schematic of the experimental facility

IV Computation of Shear

In addition to the flow quantities just mentioned, the shear on the bottom of the channel was determined by means of the following considerations. First the momentum equation was written for the flow in the channel. The flow was assumed to be steady and one-dimensional and the pressure distribution was taken to be hydrostatic. It should be noted, however, that the density is a variable and these density changes must be accounted for in the momentum equation. With these assumptions the resulting equation may be written as:

$$\frac{\tau_w P}{\rho_p \nu g h \cos\theta} = \tan\theta + \frac{dh}{dx}(\mathrm{Fr}^2 - 1) + \frac{h}{\nu}\frac{d\nu}{dx}(\mathrm{Fr}^2 - 1/2) \tag{1}$$

where τ_w represents the shear at the channel bottom, ρ_p the particle density, ν the solid fraction, h the depth of flow, θ the channel inclination, x the distance along the channel, and g is the gravitational acceleration. The Froude number, Fr, was defined as $U/(gh\cos\theta)^{1/2}$, where U is the average velocity. The first term on the right-hand side corresponds to the component of the gravitational force in the direction of flow. The last two terms account for the acceleration of the flow and depend on the gradient of the depth of flow, (dh/dx), and on the gradient of the density, $(d\nu/dx)$. The accuracy of this equation could be improved by the inclusion of profile parameters which depend on the shape of the velocity and density profiles over the depth of flow. However, in the absence of this information, these profile parameters are assumed to be unity. The factor P on the left-hand side accounts for the friction effect of the side walls and may be expressed as

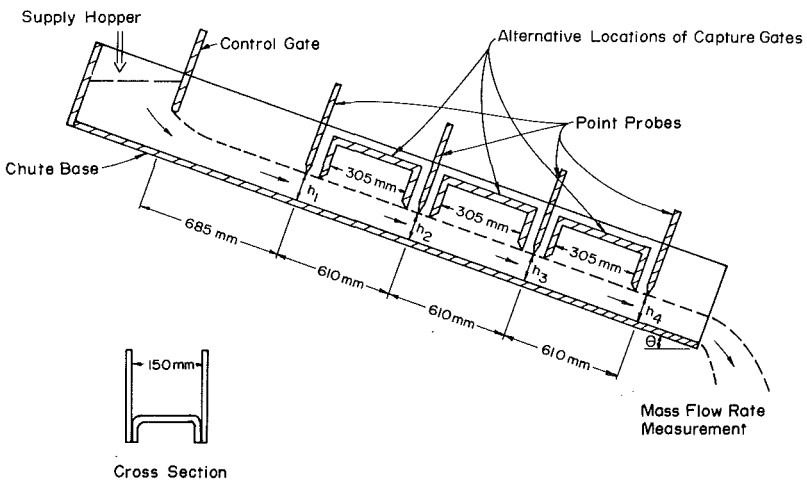


Fig. 2 Schematic of the test channel and density device

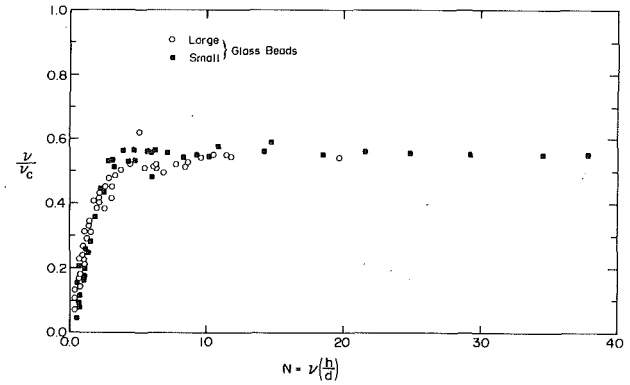


Fig. 3 The variation of the solid fraction, ν, as a function of the characteristic number of particle layers in the flow, N, for both sizes of glass beads

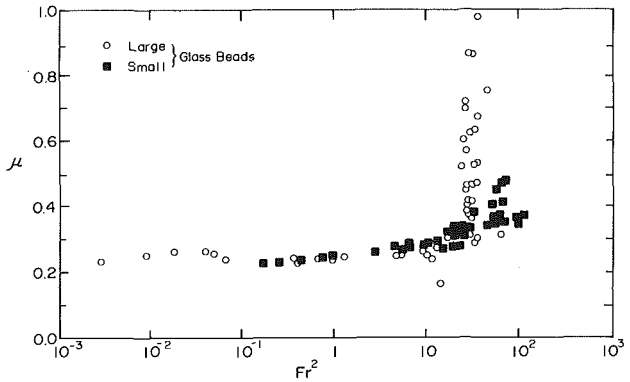


Fig. 4 The variation of the friction coefficient, μ, as a function of Froude Number Fr^2. (Both 3 mm and 0.3 mm glass beads are shown.)

Table 1 Material properties

Average size of small glass beads	0.26 mm
Average size of large glass beads	2.94 mm
Bulk specific gravity (both sizes)	1.5
Wall friction angle against aluminum (both sizes)	15 to 18 deg
Internal friction angle (both sizes)	18 to 26 deg

$$P = (1 + \beta h/b)$$

where b is the width of the channel, h is the flow depth, and β is a constant which is selected as discussed below. Once this selection has been made, the shear τ_w (or the ratio of τ_w to the normal stress $\rho_p \nu gh\cos\theta$) may be computed from measured quantities $h(x)$, $\nu(x)$ and the total mass flow rate. The ratio $\tau_w/\rho_p \nu gh\cos\theta$ is called the friction coefficient, μ.

The accuracy of the determination of τ_w from equation (1) depends on the accuracy with which the variables can be measured. Perhaps the largest source of uncertainty comes from the measurement of h, particularly for flows of very low densities. This measurement in turn influences the values of the density ν. Under these conditions we have allowed for errors as high as ± 30 percent. Fortunately, the product νh, which enters the computation of τ_w, is obtained directly from the trapped material and this quantity is, therefore, more accurately assessed than either ν or h alone. The terms involving the derivatives dh/dx and dv/dx were, in all cases, small compared to the body force term, and did not materially contribute to errors in the computation of τ_w.

A comment should be added regarding the constant β. A value of $\beta = 2$ results from the assumption that the shear against the side walls is equal to that at the channel bottom. A value of $\beta = 0$, on the other hand, represents frictionless side walls. For the granular materials tested, the data suggest a value for β somewhere in between. Furthermore, at very low velocities, one would expect the shear to correspond to that produced by Coulomb friction between two solids. With this consideration in mind, β was selected to be unity. Fortunately the selection of β is not critical to the determination of the shear τ_w, since the chute was designed to be sufficiently wide with the express purpose of minimizing the wall friction effect.

V Experimental Results

A large number of experiments were conducted using two sizes of glass beads measuring 0.3 mm and 3.0 mm in diameter. The material properties of these glass beads are given in Table 1. For each material the channel was set at a number of different angles and for each angle a range of mass flow rates and flow depths were covered. The flow was controlled by gates at the inlet section and in some cases also at the discharge section of the channel.

The experimental results may be presented in terms of a number of possible dimensionless parameters. The two parameters which seem to be particularly suitable are the friction coefficient ($\tau_w/\rho_p \nu gh\cos\theta$) and the Froude number $U/(gh\cos\theta)^{1/2}$. Additional parameters, such as h/d, may be considered when results for different sizes of particles are compared.

Before considering the variations in the friction coefficient with various parameters, it is useful to examine the effects of the density on the flow characteristics. To obtain further insight into the relation between the density and the flow characteristics, one may define a quantity

$$N = \frac{\nu h}{d}$$

One may think of N as the number of particle layers in the flow. By plotting N against ν, one can examine the extent to which the dilation of the material depends on the number of layers of particles in the flow. Fig. 3 is such a graph and contains the data for both the 0.3 and 3 mm glass beads. Note that for both sizes, the depth-averaged solids fraction, ν, is essentially constant for flows with N greater than approximately four. On the other hand, when N decreases below 4 the material dilates substantially. These data appear to be almost independent of shear rate.

The relationship between the friction coefficient, μ, and the Froude number is shown in Fig. 4. The data for both sizes of glass beads indicate that the friction coefficient remains fairly constant up to a certain Froude number. Beyond that, the data shows a rapid increase in μ with a further increase in the Froude number. Even though the individual points may be subject to large errors, as pointed out earlier, the measured increase in friction coefficient is so large as to leave no doubt about the reality of this sudden rise. The deviation from this constant value occurs at different Froude numbers for the two sizes. The smaller glass beads show a more gradual increase with Froude number than the larger glass beads, but it is apparent that a definite deviation from the Coulomb friction coefficient occurs. For both sizes of beads the value of the friction coefficient at low Froude numbers is essentially equal to the Coulomb friction coefficient for glass beads against aluminum, which implies Coulomb friction governs these flows in this range. At higher flow velocities, the increase in the friction coefficient represents a deviation from this simple Coulomb behavior. This behavior of the friction coefficient has an interesting consequence. It implies that a solid material in granular form may reach uniform flow (non-accelerating flow) at various chute angles.

The question arises as to the mechanism responsible for the increase in the friction coefficient at high Froude numbers. At these high Froude numbers it is observed that the particles are highly agitated and contact with the wall consists of collisions

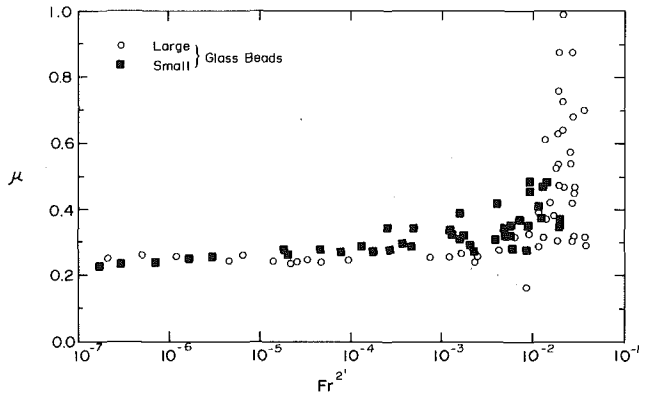


Fig. 5 The variation of the friction coefficient, μ, as a function of Froude Number, $Fr^{2\prime} = Fr^2(d/h)^2$. (Both sizes of glass beads are shown.)

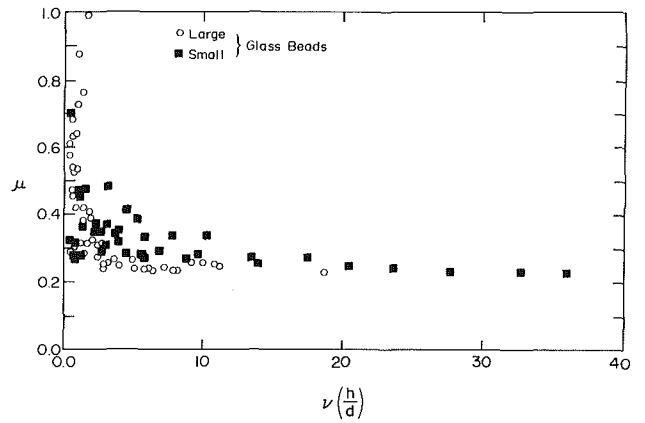


Fig. 6 The variation of the friction coefficient, μ, as a function of $\nu(h/d)$. (Both sizes of glass beads are shown.)

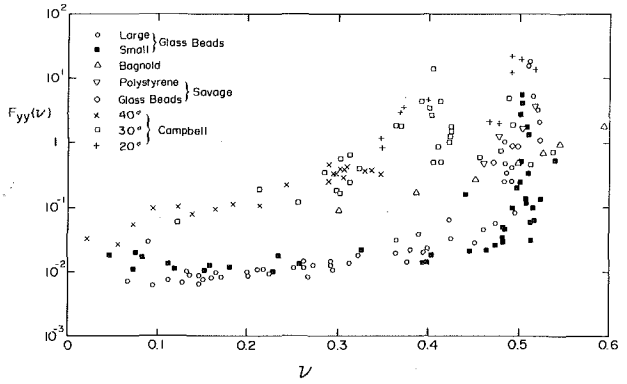


Fig. 7 The variation of $F_{xy}(\nu)$ as a function of ν as compared with previous works

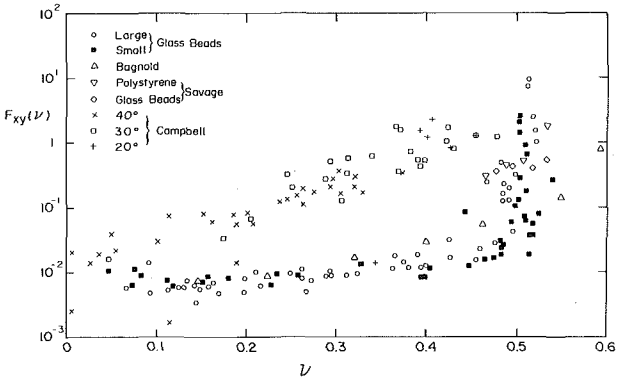


Fig. 8 The variation of $F_{yy}(\nu)$ as a function of ν, as compared with previous works

rather than sliding. These different means for the exchange of momentum might explain the deviation from Coulomb friction seen in these flows. The precise mechanism responsible for an increase in the effective friction angle, however, is unclear. The interaction between the granules and the channel surface is likely to be quite complex and may depend on additional parameters including the size of the particles, the thickness of the channel bottom, the materials involved and the condition of the surfaces.

As we have already observed, the experimental data of the friction coefficient for the two different sizes of beads fall on two different curves when plotted as a function of the Froude number. This fact seems to indicate that a geometrical factor such as h/d is influencing the results. In an attempt to include this factor in a simple way, a modified Froude number is introduced.

$$Fr' = Fr(d/h)$$

The results for both sizes of glass beads are shown in Fig. 5 in which the friction coefficient is plotted against this new Froude number. The two sets of data appear to fall on approximately the same curve. The generalization of this result will depend on future experiments with a greater variety of sizes and materials. Nevertheless, the data in Fig. 5 indicate the possibility of obtaining reasonable correlations by means of simple modified parameters.

So far the friction coefficient, μ, has been presented as a function of the Froude numbers. One might also consider the average density ν as a parameter. In order to account also for the factor h/d, several combinations of ν and h/d were tried and a rather acceptable correlation was established between ν and $\nu(h/d)$, as shown in Fig. 6. In general, low densities coincide with high Froude numbers. Therefore, the friction coefficient reaches its largest values as $\nu(h/d)$ becomes small.

VI Implications for the Constitutive Relations

In a model originally proposed by Bagnold (1954, 1956, 1966) and further developed by Lun et al. (1984), Jenkins and Savage (1983), as well as others, the stresses in a rapidly flowing granular material may be expressed in the form so that

$$\sigma_{ij} = \rho_p F_{ij}(\nu) d^2 (du/dy)^2 \qquad (2)$$

where the functions F_{ij} depend only on the solid fraction ν. The available experimental data do not allow accurate computations or local determinations of these functions. However, since the formulation of stresses in terms of the functions F_{ij} represents a very fundamental aspect in the understanding of granular flows, an attempt to obtain some rough estimates of those functions seems justified. Therefore, since details of the velocity profile and any possible slip at the wall have not been measured, the velocity gradient du/dy will be characterized by U/h where U is the average velocity and h the depth of flow. The stress σ_{yy} and σ_{xy} are taken to be the pressure and shear on the channel bottom and the corresponding density functions are given the symbols F_{yy} and F_{xy}.

The function F_{xy} computed in this way for the present study is shown in Fig. 7. The data for the two sizes of glass beads fall essentially on the same curve. The data from the experiments by Bagnold (1954, 1956, 1966) and by Savage and Sayed (1982) are shown which were taken using Couette flow devices. Considering the rough approximations used in the computation of F_{xy}, the data from these experiments correlate rather well with the present results and adds further support to the significance of Bagnold's original formulation. The data for the normal stress function F_{yy} are shown in Fig. 8 together

with those of Bagnold and Savage. The same conclusions may be drawn about the normal stress function.

It should be mentioned that values for the functions F_{xy} and F_{yy} may also be derived from the computational work of Campbell (1982) and others. The values derived by Campbell, although showing the same trends, are numerically higher than those derived experimentally. These computational works, however, are based on a number of assumptions concerning the details of the collision mechanics which, at this time, cannot be compared to the actual physical conditions. The data developed by Campbell (1982) for the density functions, F_{xy} and F_{yy}, are also shown in Figs. 7 and 8.

VIII Summary and Conclusions

An extensive experimental study of the flow of granular materials in an open channel was performed. The materials involved in the study were glass beads of two different sizes. A technique was developed which allowed the measurement of the average density of the flowing material. The results clearly show that for the materials studied in this work the friction coefficient, μ, is not a constant as in the case of two solids in sliding contact but rather that the friction coefficient increases with increases in the Froude number. This result implies that for the flows of granular materials in open channels, it may be possible to obtain uniform flow for a range of different angles of channel inclination.

The experimental data was also used to assess the validity of the analytic expressions proposed by Bagnold for the stresses developed in a granular material flow. In particular, Bagnold's expressions for the shear and normal force, which define the functions F_{xy} and F_{yy} as functions of only the density, were tested. The results support such a relationship and are in qualitative agreement with other experimental and theoretical works for simple shear flows of granular materials.

Acknowledgments

The authors would like to thank the Union Carbide Corporation and the National Science Foundation Grant # CME 7915132 for their interest and support in this study. The authors would also like to thank Susan Rosik and Gerald Zeininger for their help in the collecting of the data. A special thanks is in order for Professor Charles S. Campbell for his encouragement and numerous conversations concerning the procedures and results that are presented.

References

Augenstein, D. A., and Hogg, R., 1974, "Friction Factor for Powder Flow," *Power Technology*, Vol. 10, p. 43.

Augenstein, D. A., and Hogg, R., 1978, "An Experimental Study of the Flow of Dry Powders on Inclined Surfaces," *Powder Technology*, Vol. 19, p. 205.

Bagnold, R. A., 1954, "Experiments on a Gravity Free Dispersion of Large Solid Spheres in a Newtonian Fluid," *Proc. Roy. Soc.*, Vol. 225, p. 49.

Bagnold, R. A., 1956, "The Flow of Cohesionless Grains in Fluids," *Phil. Trans. R. Soc. London*, Ser. A., Vol. 249, p. 235.

Bagnold, R. A., 1966, "The Shearing and Dilation of Dry Sand and the Singing Mechanism," *Proc. Roy, Soc. A.*, Vol. 295, p. 219.

Bailard, J., 1978, "An Experimental Study of Granular - Fluid Flow," Ph.D. Thesis, University of California, San Diego.

Campbell, C. S., Brennen, C. E., and Sabersky, R. H., 1985a, "Flow Regimes in Inclined Open Channel Flows of Granular Materials," *Powder Technology*, Vol. 41, p. 77.

Campbell, C. S., 1982, "Shear Flows of Granular Materials," Ph.D. Thesis, California Institute of Technology, Pasadena, CA.

Campbell, C. S., and Brennen, C. E., 1985b, "Chute Flows of Granular Materials: Some Computer Simulations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, p. 172.

Craig, K., Buckholz, R. H., and Domoto, G., 1986, "An Experimental Study of the Rapid Flow of Dry Cohesionless Metal Powder," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, p. 935.

Hanes, D. M., and Inman, D. L., 1985, "Observations of Rapidly Flowing Granular Fluid Materials," *J. Mech.*, Vol. 150, p. 357.

Jenkins, J. T., and Savage, S. B., 1983, "A Theory for the Rapid Flow of Identical Smooth, Nearly Elastic Particles," *J. Fluid Mech.*, Vol. 130, p. 187.

Lun, C.K.K., Savage, S. B., Jeffrey, D. J., and Chepurniy, N., 1984, "Kinetic Theories for Granular Flow: Inelastic Particles in a Couette Flow and Slightly Inelastic Particles in a General Flow Field," *J. Fluid Mech.*, Vol. 140, p. 223.

Sayed, M., and Savage, S. B., 1983, "Rapid Gravity Flow of Cohesionless Granular Materials down Inclined Chutes," *J. App. Math. Phys.*, Vol. 34, p. 84.

Savage, S. B., 1984, "The Mechanics of Rapid Granular Flows," *Adv. Appl. Mechanics*, Vol. 24, p. 289.

Savage, S. B., and McKeown, S., 1983, "Shear Stress Developed During Rapid Shear of Dense Concentrations of Large Spherical Particles Between Concentric Cylinders," *J. Fluid Mechanics*, Vol. 127, p. 453.

Savage, S. B., 1979, "Gravity Flow of Cohesionless Granular Material in Chutes and Channels," *J. Fluid Mech.*, Vol. 92, Part 1, p. 53.

Savage, S. B., and Sayed, M., 1982, "Stress Developed by a Dry Cohesionless Granular Material Sheared in an Annular Shear Cell," *J. Fluid Mech.*, Vol. 142, p. 391.

Analytical Characterization of Shear Localization in Thermoviscoplastic Materials

A. Molinari
Laboratoire de Physique et Mécanique des Matériaux, U.A. No. 1215 Associée au C.N.R.S, U.E.R. des Sciences, Ile du Saulcy, 57000 Metz, France

R. J. Clifton
Division of Engineering, Brown University, Providence, RI 02912

Critical conditions for shear localization in thermoviscoplastic materials are obtained in closed form for idealized models of simple shearing deformations. The idealizations, which include the neglect of heat conduction, inertia, and elasticity, are viewed as quite acceptable for many applications in which shear bands occur. Explicit results obtained for the idealized, but fully nonlinear problem show the roles of strain-rate sensitivity, thermal softening, strain hardening, and initial imperfection on the localization behavior. Numerical solutions for two steels are shown to exhibit the principal features reported for torsional Kolsky bar experiments on these steels. Mathematically exact critical conditions obtained for the fully nonlinear problem are compared with critical conditions obtained by means of linear perturbation analysis. Use of relative changes instead of absolute changes in the linear perturbation analysis gives better agreement with predictions of the fully nonlinear analysis.

1 Introduction

Shear instabilities in the form of shear bands are commonly observed in metals and polymers subjected to large deformations. The formation of a shear band is often an immediate precursor to rupture of the material. Even when rupture does not occur, the development of shear bands generally reduces the performance of the material. Thus, improved understanding of shear band formation is critical to the development of improved materials and components made from these materials.

Shear bands can be divided into two types: those in which thermal softening plays a negligible role in their formation and those in which thermal softening plays a primary role. In the former case the shear bands, sometimes called isothermal shear bands, form as a result of strain softening due, for example, to material damage, to the development of soft textures, or to phase transformations. In the latter case the shear bands, often called adiabatic shear bands, form as a result of an autocatalytic process: an increase in strain rate in a weaker zone causes a local increase in temperature which in turn, for a thermal softening material, causes a further increase in strain rate.

In this paper we consider both types of shear bands. We limit our attention to simple shearing deformations. Two fundamental questions regarding the critical conditions for shear band formation are addressed.

1. For a given constitutive law, will shear localization occur for a sufficiently large shear?
2. If so, what is the nominal critical shear for which the catastrophic process occurs?

As background for this study we note that an analysis of the stability of homogeneous simple shearing deformations has been presented by Clifton (1978) for a quasi-static deformations and Bai (1982) for dynamic deformations. They used a classical, linear perturbation analysis in which the coefficients in the linear differential equations for the perturbations were assumed to vary sufficiently slowly that these variations could be neglected in estimating the rate of growth or decay of fluctuations from the homogeneous solution. This procedure determines a critical strain at which fluctuations begin to grow; however, this initial growth may or may not lead to instablity depending on the neglected effects of the time dependence of the coefficients and the nonlinearity of the complete system of equations. Molinari and Clifton (1983) and Molinari (1984, 1985) have presented some analytical solutions of the fully nonlinear problem under quasi-static and adiabatic (no heat conduction) conditions. With these solutions available for measuring the reliability of more simple approaches for determining the onset of instability, Molinari (1985) and Fressengeas and Molinari (1987) developed a so-called relative linear perturbation analysis that accounts, in part, for the nonsteadiness of the homogeneous solution by linearizing in the relative perturbation defined as the perturbation divided by the corresponding unperturbed quantity. This approach has been shown to give predictions, as to whether or not shear bands will form, that are more in agreement with the fully nonlinear theory than are predictions based on classical

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 29, 1986.

Paper No. 87-WA/APM-26.

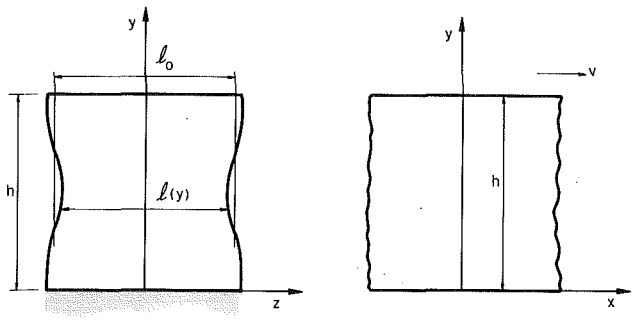


Fig. 1 Specimen geometry and loading condition

linear perturbation analysis. Dafermos and Hsiao (1983) obtained a priori estimates of the asymptotic behavior of the solution of the nonlinear problem (including inertia, but not heat conduction) for the case of the Newtonian fluid with temperature-dependent viscosity. Tzavaras (1984) extended these results to the case of non-Newtonian fluids with temperature-dependent viscosities.

Numerical solutions of the fully nonlinear system of equations have been presented by several authors: Shawki et al. (1983), Shawki (1986), Wright and Batra (1985), and Molinari (1985). From these solutions one can conclude that dynamical effects and heat conduction are relatively unimportant for steel specimens, with lengths of 5–10 mm, subjected to shearing rates of $10^3 s^{-1}$ as in the torsional Kolsky bar experiments of Costin et al. (1979) and Hartley (1986). Thus, in this paper we neglect dynamical effects and heat conduction in order to present an analytical approach to the fully nonlinear problem of thermoviscoplastic localization in simple shear. Our aim is to obtain simple analytical formulae for determining whether or not a shear strain localization instability will occur and, if so, the critical strain γ^c at which the localization becomes catastrophic. The boundary conditions will, in some cases, be general whereas in others they will be restricted to a constant imposed shear stress or a constant imposed velocity. Isothermal shear bands are considered in Section 2 and adiabatic shear bands are considered in Sections 3 and 4.

2 Isothermal Problem

We consider a simple shearing deformation of strain hardening material with strain-rate sensitivity. For illustrations, we consider the following constitutive law:

$$\tau = f(\gamma)\dot{\gamma}^m \qquad (m>0) \tag{1}$$

where τ is the shear stress, γ is the shear strain, and $\dot{\gamma}$ is the shear rate. The function $f(\gamma)$ takes account of the strain hardening. This function is not necessarily monotonicaly increasing in order to account for possible strain softening.

Suppose that, for a constant applied strain rate $\dot{\gamma}$, the shear stress τ passes through a maximum. Will strain localization occur? By localization we mean that in some narrow region, the strain becomes much larger than elsewhere. More precisely, we define localization as follows:

L_∞-Localization. If for every point A different from B, the ratio γ_B/γ_A tends to infinity with increasing time, then L_∞-localization of the deformation at the point B is said to occur.

The analysis of localization in this section is performed in two different ways. First we derive an analytical solution of the fully nonlinear problem. Then an absolute and a relative linear perturbation analysis are performed and the corresponding predictions are compared with the exact solution.

2.1 The Nonlinear Theory. We consider a slab with a geometrical defect. The width $\ell(y)$ is nonuniform as shown in Fig. 1. Using the same approach as Hutchinson and Neale (1977) for the uniaxial tension of a bar, we get from the equilibrium equation written at two different points A and B:

$$\ell_A \tau_A = \ell_A f(\gamma_A)(\dot{\gamma}_A)^m = \ell_B \tau_B = \ell_B f(\gamma_B)(\dot{\gamma}_B)^m. \tag{2}$$

Taking the power $1/m$ of each term, we get after integration:

$$\ell_A^{1/m} \int_{\gamma_A^o}^{\gamma_A} (f(\zeta))^{1/m} d\zeta = \ell_B^{1/m} \int_{\gamma_B^o}^{\gamma_B} (f(\zeta))^{1/m} d\zeta \tag{3}$$

where γ_A^o and γ_B^o are the initial strains at points A and B. If $(f(\zeta))^{1/m}$ is integrable at infinity, then the values of the integrals are finite. While maintaining the equality (3), let γ_B and γ_A be increased until the strain becomes infinite at one of the two points, say B. Then there exists a finite strain γ_A for which equation (3) is satisfied. Hence, we have L_∞-localization of strain if and only if the function $(f(\zeta))^{1/m}$ is integrable at infinity.

Assume that $f(\zeta)$ has power law behavior at infinity of the form

$$f(\zeta) \sim a\zeta^{-p} \text{ as } \zeta \to \infty \tag{4}$$

where a and p are positive constants. Then, from the integrability condition (3), the deformation exhibits L_∞ localization if and only if

$$-p+m<0 \tag{5}$$

This condition illustrates the stabilizing effect of the strain-rate sensitivity for $m>0$. Even if the material is strain softening ($p>0$), localization will occur only if m is sufficiently small ($m<p$). For $m-p>0$, L_∞-localization does not occur. Indeed, for $m>p$ it is readily shown that (Molinari and Clifton, 1986)

$$\lim_{\gamma_A \to \infty} (\gamma_B/\gamma_A) = (\ell_A/\ell_B)^{1/(m-p)} \tag{6}$$

so that γ_B/γ_A remains bounded as $\gamma_A \to \infty$.

To appreciate the strong stabilizing effect of strain rate sensitivity, consider a material with no strain hardening at large strains (i.e., $p=0$) and with a 1 percent geometrical defect (i.e., $\ell_B/\ell_A = 0.99$). Values of $(\gamma_B/\gamma_A)_\infty = (\ell_A/\ell_B)^{1/m}$ are given in Table 1 for different values of m. A value $m=0.01$ is sufficient to prevent pronounced localization as $\gamma \to \infty$. The stabilizing effect of strain-rate sensitivity has been shown by Pan (1983) using a similar approach.

Table 1 Influence of strain rate sensitivity on the asymptotic behavior of plastic flow (geometrical defect: $\ell_B/\ell_A = 0.99$)

m	0.2	0.1	0.01	0.006	0.001
$(\gamma_B/\gamma_A)_\infty$	1.052	1.106	2.732	7.464	2.316×10^4

2.2 Linear Perturbation Analysis. For comparison of the results of the fully nonlinear theory to the predictions of a linear stability analysis, consider a block of uniform thickness $\ell(y)=\ell_o$ undergoing homogeneous simple shearing deformation $\gamma_o(t)$. Let the perturbation $\delta\gamma$ be the difference $\gamma(y, t)-\gamma_o(t)$ where $\gamma(y, t)$ is the shear strain for the same block subjected to the same boundary conditions, but having a fluctuation in strain and strain rate beginning at some time t_o. Let the relative perturbation $\Delta\gamma$ be defined as

$$\Delta\gamma = \frac{\delta\gamma}{\gamma_o}. \tag{7}$$

Using the constitutive law (1) and considering the problem as quasistatic (i.e., $\delta\tau=0$) we obtain:

$$\frac{\delta\dot{\gamma}}{\delta\gamma} = -\frac{f'(\gamma_o)}{f(\gamma_o)}\frac{\dot{\gamma}_o}{m} \tag{8}$$

$$\frac{\Delta\dot{\gamma}}{\Delta\gamma} = \frac{\delta\dot{\gamma}}{\delta\gamma} - \frac{\dot{\gamma}_o}{\gamma_o} = -\left(\frac{\gamma_o f'(\gamma_o)}{mf(\gamma_o)} + 1\right)\frac{\dot{\gamma}_o}{\gamma_o} \tag{9}$$

when the respective perturbations are sufficiently small. Equation (8) shows that, at least initially, the strain difference $\delta\gamma$ grows when $f'(\gamma) < 0$, i.e., when strain softening occurs. The relative perturbation $\Delta\gamma$ tends to grow more slowly than the absolute perturbation $\delta\gamma$ and may even decay as the perturbation grows. If $f(\gamma_o)$ has the behavior (4) for large values of γ_o, then as $\gamma_o \to \infty$ the relative perturbation $\Delta\gamma$ continues to grow for $-p+m<0$ and decays for $-p+m>0$. These conditions are, respectively, the same as the critical conditions for L_∞ – localization to occur or not. This parallellism between predictions of the linear relative perturbation analysis and the exact results for the nonlinear theory suggests that linear relative perturbation analysis may be more widely useful in predicting the stability of deformations than is commonly used linear perturbation analysis represented by equation (8). However, we emphasize that the localization analysis in the nonlinear theory and the linear relative perturbation analysis address different problems and there is not a priori reason to expect that the critical conditions for L_∞-localization are, in general, the same as the critical conditions for predicted unbounded growth of a relative perturbation.

3 Adiabatic Case

We consider next the influence of temperature on localization. As discussed in the introduction, the deformation is assumed to be adiabatic and quasistatic. We consider the constitutive equation

$$\tau = \tau(\gamma, \dot{\gamma}, \theta), \tag{10}$$

the equation of equilibrium

$$\ell(y)\tau(y, t) = \ell(h)\tau(h, t), \tag{11}$$

the compatibility equation

$$\dot{\gamma} = \frac{\partial v}{\partial y}, \tag{12}$$

and the energy equation

$$\rho C\frac{\partial\theta}{\partial t} = \beta\tau\dot{\gamma}. \tag{13}$$

In these equations ρ is the mass density, C is the heat capacity per unit mass, θ is the absolute temperature, v is the particle velocity, and β is the Taylor-Quinney coefficient which characterizes the fraction of plastic work that is converted into heat; usually β is taken constant and equal to 0.9. Equations (10)–(13) constitute four equations in the four unknowns γ, θ, τ, v. With elasiticity effects neglected these equations are applicable to arbitrarily large deformations in simple shear.

In the following, we present a discussion of localization for different constitutive laws and different boundary conditions. We consider the cases of constant velocity boundary conditions:

$$v(0, t) = 0$$
$$v(h, t) = v_o \tag{14}$$

or constant stress boundary conditions

$$\ell(h)\tau(h, t) = \ell(0)\tau(0, t) = \text{const.} \tag{15}$$

3.1 Materials Without Strain-Hardening. An exact solution of the fully nonlinear problem has been presented by Molinari and Clifton (1983) for the case in which the material is not strain hardening and equation (10) has the form

Table 2 Localization results for visco-plastic, thermal softening materials without strain hardening ($m>0$, $\mu_O>0$, $a>0$)

	Constitutive law	L_∞-Localization
L_1	$\tau = \mu_o\theta^\nu\dot{\gamma}^m$	$\nu+m<0$
L_2	$\tau = \mu_o\exp/(-\alpha\theta)\dot{\gamma}^m$	$\alpha>0$
L_3	$\tau = \sup(a+b\theta, 0)\dot{\gamma}^m$	$b<0$
L_4	$\tau = \mu_o\exp(-\alpha/\theta)\dot{\gamma}^m$	never exhibits L_∞-localization

$$\tau = \mu(\theta)\dot{\gamma}^m. \tag{16}$$

In order to obtain this exact solution we write equations (11) and (13) at two different points A and B. Substitutions of equation (16) into (11) and use of equation (13) to eliminate $\dot{\gamma}_A/\dot{\gamma}_B$ gives

$$\ell_A^{(m+1)/m}\mu(\theta_A)^{1/m}d\theta_A = \ell_B^{(m+1)/m}\mu(\theta_B)^{1/m}d\theta_B \tag{17}$$

which, after integration, becomes

$$\ell_A^{(m+1)/m}\int_{\theta_A^o}^{\theta_A}\mu(\zeta)^{1/m}d\zeta = \ell_B^{(m+1)/m}\int_{\theta_B^o}^{\theta_B}\mu(\zeta)^{1/m}d\zeta \tag{18}$$

where θ_A^o and θ_B^o are the initial temperatures at points A and B.

From equation (18) it appears that L_∞-localization of temperature occurs at B if, and only if, $\mu(\theta)^{1/m}$ is integrable at infinity. Localization will occur at the point B where the following quantity, defined for each point M of the slab,

$$M \to \ell_M^{(m+1)/m}\int_{\theta_M^o}^{\infty}\mu(\zeta)^{1/m}d\zeta \tag{19}$$

is a minimum. At localization the temperature θ_A^c at any point A is given by

$$\ell_A^{(m+1)/m}\int_{\theta_A^o}^{\theta_A^c}\mu(\zeta)^{1/m}d\zeta = \ell_B^{(m+1)/m}\int_{\theta_B^o}^{\infty}\mu(\zeta)^{1/m}d\zeta. \tag{20}$$

It is easy to show that L_∞-localization of temperature tends to result in strain localization. Indeed, from the equilibrium condition (11) and the constitutive law (16), we have

$$\ell_A\mu(\theta_A)\dot{\gamma}_A^m = \ell_B\mu(\theta_B)\dot{\gamma}_B^m.$$

Then, assuming temperature localization, we get

$$\lim_{\theta_A\to\theta_A^c}\left(\frac{\dot{\gamma}_A}{\dot{\gamma}_A}\right) = \lim_{\theta_A\to\theta_A^c}\left(\frac{\ell_A\mu(\theta_A)}{\ell_B\mu(\theta_B)}\right)^{1/m} = +\infty$$

since $\lim_{\theta_A\to\theta_A^c}\mu(\theta_B) = 0$ from the integrability condition. This L_∞-localization of the strain rate essentially ensures L_∞-localization of the strain although various pathological cases must be excluded in a rigorous analysis. We henceforth consider constitutive equations and loading conditions for which such pathological cases are excluded.

Localization results obtained from the integrability condition are summarized in Table 2 for several constitutive laws.

3.1.1 Calculation of the Critical Strain. In this section we obtain explicit results for the critical strain at localization. We illustrate the approach by considering the constitutive law L_2. Substituting this law in equation (18), we obtain the following expression for the temperature at a point B as a function of the temperature at a point A

$$\theta_B = -\frac{m}{\alpha}\log\left[\left(\frac{\ell_A}{\ell_B}\right)^{(1+m)/m}\exp(-\alpha\theta_A/m) + C_1\right] \tag{21}$$

where

$$C_1 = \exp(-\alpha\theta_B^o/m) - \left(\frac{\ell_A}{\ell_B}\right)^{(1+m)/m} \exp(-\alpha\theta_A^o/m). \quad (22)$$

From equation (21), a necessary condition for localization to occur is $C_1 < 0$ since for $C_1 \geq 0$ the logarithm cannot tend to infinity. Let us identify the point B as the point where the quantity $\ell_B{}^{(m+1)/m}\exp(-\alpha\theta_B^o/m)$ is a minimum. Then, if the initial temperature (or width ℓ) is nonuniform, C_1 is strictly negative.

For thermal softening material, i.e., $\alpha > 0$, the quantity $\exp(-\alpha\theta_A/m)$ decreases to zero as $\theta_A \to \infty$. For some critical temperature θ_A^c, the temperature θ_B will become infinite. From equation (22), this value is

$$\theta_A^c = \theta_A^o - \frac{m}{\alpha}\log\left[1 - \left(\frac{\ell_B}{\ell_A}\right)^{\frac{1+m}{m}} \exp(-\alpha(\theta_B^o - \theta_A^o)/m)\right]. \quad (23)$$

If the material is thermal hardening, i.e., $\alpha < 0$, then the term $\exp(-\alpha\theta_A/m)$ grows and, from equation (21), it is obvious that L_∞-localization is impossible. Indeed, for sufficiently large θ_A, the difference

$$\theta_B - \theta_A \simeq \log\left[\left(\frac{\ell_A}{\ell_B}\right)^{(-1-m)/\alpha}\right]$$

becomes small compared to the absolute temperature, say θ_A.

To calculate the critical strain from the critical temperature we consider first the case in which a constant stress τ is applied at the boundary. Then, from the energy equation (13), θ can be calculated as a function of the strain γ

$$\theta = \frac{\beta\tau}{\rho C}\gamma + \theta^o. \quad (24)$$

The critical strain, γ_A^c, at A is obtained by substituting the critical temperature θ_A^c given by equation (23) into (24) to obtain

$$\gamma_A^c = -\frac{\rho C m}{\alpha\beta\tau}\log\chi \quad (25)$$

where

$$\chi = 1 - \left[\frac{\ell_B}{\ell_A}\right]^{\frac{1+m}{m}} \exp(-\alpha(\theta_B^o - \theta_A^o)/m) \quad (26)$$

is a defect parameter that includes both geometrical and temperature defects. The stabilizing effects of increased strain rate sensitivity (i.e., larger m) and decreased thermal softening (i.e., smaller α) are evident in this expression. The critical, nominal strain at which the temperature becomes infinite at B is obtained from the integration of the critical strain $\gamma^c(y_A) \equiv \gamma_A^c$ over the height of the block. Thus, the critical, nominal strain is

$$\gamma_c = \frac{1}{h}\int_0^h \gamma^c(y_A)\,dy_A. \quad (27)$$

Numerical integration of equation (27) is straightforward as long as the thickness $\ell(y_A)$ varies sufficiently slowly near the point (s) B at which the strain becomes infinite.

We consider next the calculation of the critical strain for the case of the velocity boundary conditions (14). An exact solution does not appear to be possible in this case. However, a good approximate solution can be obtained for the case of weak strain-rate sensitivity (i.e., $m \ll 1$). Such weak strain rate sensitivity is commonly observed in metals at room temperature for strain rates up to $10^3\ s^{-1}$. Typical values of m are of the order of $m = 0.01$. In order to obtain an approximate solution for small m we introduce the mean constant strain rate

$$\dot\gamma_o = V/h \quad (28)$$

For small values of m we can approximate the stress τ by (Molinari and Clifton, 1986)

$$\tau \simeq \mu_o \exp(-\alpha\theta)\dot\gamma_o^m. \quad (29)$$

Substitution of the approximate stress (29) into the energy equation (13) leads to

$$\rho C\frac{\partial\theta}{\partial t} = \beta\mu_o\exp(-\alpha\theta)\dot\gamma_o^m\dot\gamma. \quad (30)$$

This equation can be integrated by separation of the variables θ and γ to give

$$\theta(\gamma) = \theta^o + \frac{1}{\alpha}\log\left[1 + \frac{\alpha\beta\mu_o\dot\gamma_o^m}{\rho C}\exp(-\alpha\theta^o)\gamma\right]. \quad (31)$$

where θ^o is the initial temperature.

With the relationship between θ and γ given by equation (31), the critical strain, γ_A^c, at A can be obtained by integration of the equilibrium equation (11). Such integration gives

$$\ell_A^{1/m}\int_o^{\gamma_A}\exp(-\alpha\theta_A(\gamma)/m)\,d\gamma = \ell_B^{1/m}\int_o^{\gamma_B}\exp(-\alpha\theta_B(\gamma)/m)\,d\gamma \quad (32)$$

with $\theta_A(\gamma)$ and $\theta_B(\gamma)$ given by equation (31). At localization, γ_B becomes infinite and the critical strain γ_A^c at point A becomes, for $m \ll 1$,

$$\gamma_A^c = \frac{\rho C}{\alpha\beta\tau_A^o}\left\{\left[1 - \left(\frac{\ell_B}{\ell_A}\right)^{1/m}\exp(-\alpha(1-m)(\theta_B^o - \theta_A^o)/m)\right]^{-m/(1-m)} - 1\right\} \quad (33)$$

where

$$\tau_A^o = \mu_o\dot\gamma_o^m\exp(-\alpha\theta_A^o)$$

is the shear stress at A in an isothermal deformation at the same strain rate. If ℓ and θ^o are both uniform, then equation (33) implies that γ_A^c is infinite and localization does not occur. The critical strain decreases as (ℓ_B/ℓ_A) decreases and $\theta_B^o - \theta_A^o$ increases. The energy measure, $\tau_A^o\gamma_A^c$, of the critical strain increases with increasing strain rate sensitivity (i.e., increasing m) and decreasing thermal softening (i.e., decreasing α). Again, the nominal critical strain is obtained by the substitution of equations (33) into (27). Comparison of equation (33) and (25) indicates that, for $m \ll 1$, the defect parameter χ of equation (26) again characterizes the effects of both temperature and geometric defects.

Identical calculations can be performed for a power law dependence of the flow stress on the temperature. Analogous results for the constitutive law L_1 of Table 2 are

$$\gamma_A^c = \frac{\rho C\theta_A^o}{\beta\tau_A}\left\{\left[1 - \left(\frac{\ell_B}{\ell_A}\right)^{\frac{1+m}{m}}\left(\frac{\theta_B^o}{\theta_A^o}\right)^{\frac{\nu+m}{m}}\right]^{\frac{m}{\nu+m}} - 1\right\} \quad (34a)$$

for constant stress boundary conditions, and

$$\gamma_A^c = \frac{\rho C\theta_A^o}{(1-\nu)\beta\tau_A^o}\left\{\left[1 - \left(\frac{\ell_B}{\ell_A}\right)^{\frac{1}{m}}\left(\frac{\theta_B^o}{\theta_A^o}\right)^{\frac{\nu+m(1-\nu)}{m}}\right]^{\frac{(1-\nu)m}{\nu+m(1-\nu)}} - 1\right\} \quad (34b)$$

for constant velocity boundary conditions; equation (34*b*) holds only for $m \ll 1$.

3.2 Materials With Strain Hardening. Strain hardening cannot be ignored for most materials. In this section we derive analytical localization criteria for constitutive laws of the form

$$\tau=\mu(\theta)(\gamma+\gamma^o)^n\dot{\gamma}^m \tag{35}$$

where γ^o is the initial strain. The approach is similar to that used in Section 3.1.

3.2.1 Stress Boundary Condition. Elimination of $\dot{\gamma}$ between the constitutive law (35) and the energy equation (13) leads to

$$\frac{d\theta}{dt}=\frac{\beta\tau}{\rho C}^{\frac{m+1}{m}}\mu(\theta)^{-1/m}(\gamma+\gamma^o)^{-n/m}. \tag{36}$$

We write equation (36) at two different points A and B, take account of the equilibrium equation (11), use equation (24) to eliminate γ, and integrate to obtain

$$\ell_A^{\frac{m+1}{m}}\int_{\theta_A^o}^{\theta_A}\mu(\zeta)^{1/m}\left((\zeta-\theta_A^o)\frac{\rho C}{\beta\tau_A}+\gamma_A^o\right)^{n/m}d\zeta$$

$$=\ell_B^{\frac{m+1}{m}}\int_{\theta_B^o}^{\theta_B}\mu(\zeta)^{1/m}\left((\zeta-\theta_B^o)\frac{\rho C}{\beta\tau_B}+\gamma_B^o\right)^{n/m}d\zeta. \tag{37}$$

Equation (37) is a generalization of equation (18) to strain hardening materials. Analysis of equation (37) analogous to that of (18), shows that L_∞-localization occurs if and only if the function

$$\theta\to\mu(\theta)^{1/m}\left((\theta-\theta^o)\frac{\rho C}{\beta\tau}+\gamma^o\right)^{n/m}$$

is integrable at infinity. For the constitutive law

$$\tau=\mu_1\theta^\nu(\gamma+\gamma^o)\dot{\gamma}^m \tag{38}$$

L_∞-localization occurs if and only if

$$\nu+n+m<0. \tag{39}$$

The inequality (39) provides a good illustration of the competition between the stabilizing effects of strain hardening ($n>0$) and positive strain-rate sensitivity ($m>0$), and the destabilizing effects of thermal softening ($\nu<0$). The localization criteria (39), obtained by Molinari and Clifton (1983), has also been obtained by Fressengeas and Molinari (1986) as the criterion for the initial growth of a fluctuation based on a linear relative perturbation analysis. The inequality (39) differs from the condition

$$\nu+n<0 \tag{40}$$

that must be satisfied for the initial growth of a fluctuation according to absolute linear perturbation analysis. The difference between the conditions (39) and (40) illustrates the tendency for absolute linear perturbation analysis to predict growth of fluctuations under some conditions for which the full nonlinear analysis predicts that localization will not occur.

For the constitutive law

$$\tau=\mu_o e^{-\alpha\theta}(\gamma+\gamma^o)^n\dot{\gamma}^m \tag{41}$$

a similar analysis shows that L_∞-localization occurs if and only if $\alpha>0$ (thermal softening). For this constitutive equation the critical strain γ_A^c at A when the strain at point B becomes infinite is obtained by the substitution of equation (41) into (37) to obtain, after a change of variable,

$$K_A\int_{\tilde{\gamma}_A^o}^{\tilde{\gamma}_A^c}e^{-\nu}\nu^{n/m}d\nu=K_B\int_{\tilde{\gamma}_B^o}^{\infty}e^{-\nu}\nu^{n/m}d\nu \tag{42}$$

where

$$\tilde{\gamma}_A^c=\frac{\alpha\beta\tau}{m\rho C}(\gamma_A^c+\gamma_A^o),\quad \tilde{\gamma}_A^o=\frac{\alpha\beta\tau}{m\rho C}\gamma_A^o$$

$$K_A=\ell_A^{\frac{1+n+m}{m}}$$

with A replaced by B for $\tilde{\gamma}_B$ and K_B. Localization will occur at the point B, where the quantity on the right side of equation (42) is a minimum. Equation (42) has the same form as that obtained by Hutchinson and Neale (1977) in the study of the rupture of a viscoplastic bar in tension although the physical effects being modeled are different—their analysis included necking, but did not include the thermal softening which is included here.

3.2.2 Velocity Boundary Conditions. As in Section 3.1 we consider constant velocity boundary conditions and assume that the strain rate sensitivity of the material is small (i.e., $m\ll 1$). To calculate the temperature from the energy equation (13) we replace $\dot{\gamma}$ by $\dot{\gamma}_o=V/h$ in the constitutive equation (35) and integrate to obtain

$$\int_{\theta^o}^{\theta}\frac{1}{\mu(\zeta)}d\zeta=\frac{\beta\dot{\gamma}_o^m}{\rho C(n+1)}\left((\gamma+\gamma^o)^{n+1}-(\gamma^o)^{n+1}\right). \tag{43}$$

Substitution into equation (43) of functions $\mu(\theta)$ that model the temperature dependence of the flow stress gives the required relationship between the temperature θ and the shear strain γ. For $\mu(\theta)=\mu_1\theta^\nu$ we obtain

$$\theta(\gamma)=\theta^o\left[1+(1-\nu)\frac{\beta\mu_1\dot{\gamma}_o^m}{\rho C(n+1)(\theta^o)^{1-\nu}}\right.$$
$$\left.[(\gamma+\gamma^o)^{n+1}-(\gamma^o)^{n+1}]\right]^{\frac{1}{1-\nu}} \tag{44}$$

and for $\mu(\theta)=\mu_o e^{-\alpha\theta}$ we obtain

$$\theta(\gamma)=\theta^o+\frac{1}{\alpha}\log\left[1+\frac{\alpha\beta\mu_o\dot{\gamma}_o^m e^{-\alpha\theta^o}}{\rho C(n+1)}\right.$$
$$\left.\left((\gamma+\gamma^o)^{n+1}-(\gamma^o)^{n+1}\right)\right]. \tag{45}$$

These equations provide an approximate relationship between the temperature and the strain at each position as long as the exponent m is sufficiently small for the dependence of the shear stress on strain rate to be represented by $\dot{\gamma}_o^m$, where $\dot{\gamma}_o$ is the nominal strain rate, instead of by $\dot{\gamma}^m$, where $\dot{\gamma}$ is the local strain rate.

In order to investigate the critical conditions for localization, we substitute the functions $\theta(\gamma)$ obtained from equations (44) or (45) into the equation

$$\ell_A^{1/m}\int_o^{\gamma_A}\mu(\theta_A(\zeta))^{1/m}(\zeta+\gamma_A^o)^{n/m}d\zeta$$

$$=\ell_B^{1/m}\int_o^{\gamma_B}\mu(\theta_B(\zeta))^{1/m}(\zeta+\gamma_B^o)^{n/m}d\zeta. \tag{46}$$

As before, L_∞-localization occurs if, and only if, the integral on the right side of equation (46) remains bounded as $\gamma_B\to\infty$. For $\theta(\gamma)$ given by equation (44), the condition for L_∞-localization becomes

$$\nu+n+m(1-\nu)<0 \tag{47}$$

for $0<m\ll 1$, and $\nu<1$. This condition is slightly more restrictive than the condition (39) obtained for stress boundary conditions. That is, the tendency for localization is slightly stronger for stress boundary conditions than for velocity boundary conditions in that the localization condition (39) is

Table 3 Thermomechanical properties of CRS 1018 and HRS 1020 steels

Steel Parmeter	CRS 1018	HRS 1020
ν	−0.38	−0.51
n	0.015	0.12
m	0.019	0.0133
ρ	7800 kg/m^3	7800 kg/m^3
C	500 J/kg K	500 J/kg K
μ_1	3579×10^6 S.I.	7587×10^6 S.I.

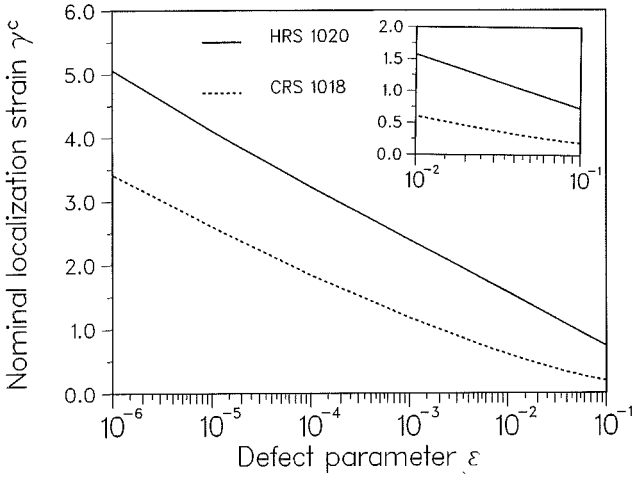


Fig. 2 Dependence of the nominal critical strain on the initial imperfection for CRS 1018 and HRS 1020

satisfied by all ν, m, and n which satisfy (47); however, for $m \ll 1$, the terms involving m in both equations (39) and (47) are often so small that, effectively, the localization conditions (39) and (47) are the same. For $\theta(\gamma)$ given by equation (45) the condition for L_∞-localization is satisfied for all $\alpha > 0$ provided that m, n satisfy $m > 0$, $n > -1$.

4 Numerical Example

Dynamic torsion experiments for investigating shear localization have been performed by Hartley et al. (1986) on two different types of steel: CRS 1018 and HRS 1020. At the strain rates ($10^3 s^{-1}$) and temperatures ($\theta^o = 300$ K) of these experiments the behavior of these materials can be represented reasonably well by a constitutive equation of the form (38). Numerical values of the various parameters in the model are given in Table 3 (Shawki, 1986). The strain γ^o is taken to have the value 0.01 for both steels. More detailed fitting of the plastic response of these steels has been presented by Klepaczko (1986). The length h of the specimens is 2.5 mm.

Variations $\ell(y)$ in the wall thickness of the specimens were not reported by Hartley (1986). Subsequently, Duffy (1986) has sectioned specimens used in such experiments to determine the variation in wall thickness, both along the length of the specimen and around its circumference. For CRS 1018 the wall thickness is relatively uniform around the circumference, but strong variation—up to 10 percent—occurs along the length of the specimen. For the purpose of this numerical example we take the geometrical factor (ℓ_B/ℓ_A) in the preceding analysis to be a parameter that varies from 0.9 to 0.999999. In order to relate the critical strains γ_A^c to the nominal strain γ^c at localization (see equation (27)), the variation in wall thickness $\ell_A = \ell(y_A)$ must be prescribed over the entire length of the specimen. Based on the general appearance of the sectioned specimens we take this variation to have the form

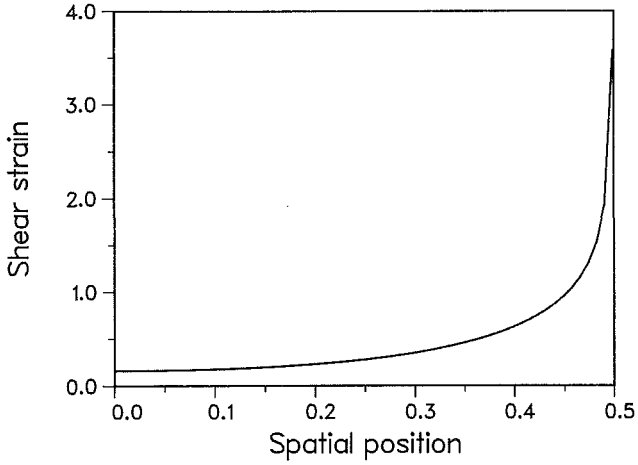


Fig. 3 Strain distribution at localization for CRS 1018 ($\epsilon = 0.02$)

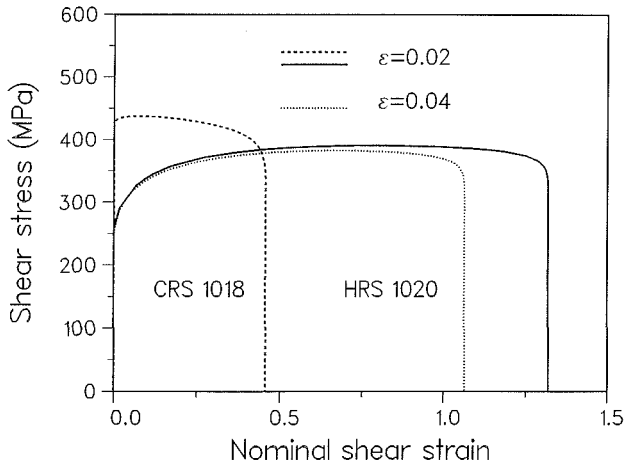


Fig. 4 Nominal stress-strain curve for simple shear of CRS 1018 and HRS 1020

$$\frac{\ell(y)}{\ell_A} = 1 + \frac{\epsilon}{2}\left(\cos\frac{2\pi y}{h} - 1\right) \tag{48}$$

where ϵ is a geometrical parameter that is taken to vary from 10^{-1} to 10^{-6} to give the range of values of 0.9 to 0.999999 for ℓ_B/ℓ_A.

Boundary conditions for the dynamic torsion ("torsional Kolsky bar") experiment are effectively those of imposed constant velocity a the ends of the specimen. Hence, we use the solution for velocity boundary conditions given by equations (44) and (46). The restriction to $m \ll 1$ that is required in obtaining equation (44) is well satisfied by the values $m = 0.019$ for the CRS and $m = 0.0133$ for the HRS. Evaluation of $\gamma^c(y_A)$ from equation (46) and integration over the length of the specimen, according to equation (27), gives the dependence of the critical nominal strain γ^c on the geometrical imperfection parameter ϵ that is shown in Fig. 2. For small ϵ the nominal critical strain varies approximately as log ϵ, as predicted for the local critical strain by equation (25); the corresponding equation, equation (33), for constant velocity boundary conditions also gives the logarithmic dependence for $0 < \epsilon \ll m \ll 1$ (Molinari and Clifton, 1986). The insert provides an expanded scale of the region of primary interest in the interpretation of torsional Kolsky bar experiments. For one value of ϵ ($\epsilon = 0.02$), the strain distribution at localization for CRS 1018 is shown in Fig. 3. The width of the band of intense shear (say, the region for which $\gamma(y) > 3\gamma(0)$) is approximately 20 percent of the length of the specimen. Such relatively wide bands are observed in CRS 1018 (Hartely et al., 1985).

The predicted nominal stress-strain curves for the two steels, with the geometrical imperfection parameter ϵ equal to 0.02, are shown in Fig. 4; a corresponding curve for $\epsilon = 0.04$ is included for HRS 1020. The general features of the curves include a slowly rising segment during which the shearing is quite uniform, a slowly falling segment during which a broad band of enhanced shearing develops, and a sharply falling segment during which the shearing becomes intensely localized in a band. These general features are characteristic of the experimental records obtained in such experiments (Hartley et al., 1986). Numerical values for the strain at the peak of the stress-strain curve and the strain at the beginning of the sharp decline in stress are comparable to values obtained in experiments. However, the predicted rate of sharp decline is greater than normally measured. This rate of decline is affected by the detailed geometry of the initial imperfection which probably was not modeled adequately by the generic form (48). Other difficulties with comparisons between theory and experiment for the steeply falling part of the curve include: (*i*) the inadequacy of the assumption that the stress obtained using the nominal strain rate can be used in calculating the local rate of energy dissipation; (*ii*) the likelihood that the final localization varies so strongly around the circumference of the specimen that a one-dimensional analysis is inappropriate; and (*iii*) the lack of constant velocity boundary conditions when the stress decreases strongly in torsional Kolsky bar experiments.

5 Conclusions

By assuming the deformation to be adiabatic and quasi-static, and by neglecting elasticity effects, we have characterized, analytically, the critical conditions for shear strain localization in simple shear. The assumed conditions are good approximations for the specimen sizes and strain rates that are commonly used in torsional Kolsky bar experiments on shear band formation in steels.

We assume the existance of initial inhomogeneities which are either geometrical defects or nonuniform fields of initial temperature or strain. The localization strain is obtained as a function of these defects, the material parameters and the boundary conditions. Two types of boundary conditions have been considered:

-constant applied stress,
-constant applied velocity.

In the latter case, the analytical results are restricted to materials with weak strain rate sensitivity.

The results are particularly simple for materials without strain hardening. In this case, explicit expressions are obtained for the dependence of the critical strain on a defect parameter that characterizes the geometrical defect and the nonuniformity of the intial temperature. For materials with weak strain rate sensitivity the critical strain depends weakly (essentially logarithmically) on the amplitude of the imperfection for small imperfections.

Comparison of predictions of the theory with experimental results for a cold-rolled steel shows good agreement in the qualitative features of the response.Quantitative comparisons require detailed descriptions of the geometrical defects of the specimens used in the experiments. Preliminary comparisons based on approximate representations of the geometrical imperfections of the specimens suggest that good quantitative agreement may be obtained once the defects are modeled accurately.

Acknowledgments

This research was supported by the Army Research Office through Contract No. DAAG29-85-K-0003 with Brown University as well as by la Direction des Recherches Etudes Techniques through Contract No. 86-814. Calculations reported here were performed in the Computational Mechanics Facility at Brown University. The facility was made possible by grants from the National Science Foundation (Solid Mechanics Program), the General Electric Foundation, and the Digital Equipment Corporation.

References

Bai, Y. L., 1982, "Thermo-Plastic Instability in Simple Shear," *Journal of the Mech. and Phys. of Solids*, Vol. 30, pp. 195–207.

Clifton, R. J., 1978, "Adiabatic Shear," in Report NMAB-356 of the NRC Committee on Material Response to Ultrasonic Loading Rates.

Costin, L. S., Crisman, E. E., Hawley, R. H., and Duffy, J., 1979, "On the Localization of Plastic Flow in Mild Steel Tubes Under Dynamic Torsional Loading," *Second Conf. on Mechanical Properties at High Rates of Strain*, Harding, J., ed., Oxford, England, pp. 90–100.

Dafermos, C. M., and Hsiao, L., 1983, "Adiabatic Shearing of Incompressible Fluids with Temperature Dependent Viscosity," *Q. Appl. Math.*, Vol. 41, pp. 45–58.

Duffy, J., 1986, Private communication.

Fressengeas, C., Molinari, A., 1987, "Instability and Localization of Plastic Flow in Shear at High Strain Rates," *Journal of the Mech. and Phys. of Solids*, Vol. 35, pp. 185–211.

Hartley, K. A., Duffy, J., and Hawley, R. H., 1986, "Measurement of the Temperature Profile during Shear Band Formation in Steels Deforming at High Strain Rates," Technical Report DAAG-29-85-K-000312, Division of Engineering, Brown University, Providence, RI.

Hartley, K. A., 1986, "Temperature Profile Measurement During Shear Band Formation in Steels at High Strain Rates," Ph.D. Thesis, Brown University, Providence, RI.

Hutchinson, J. W., and Neale, K. W., 1977, "Influence of Strain Rate Sensitivity on Necking under Uniaxial Tension," *Acta Metallurgica*, Vol 25, pp. 839–846.

Klepaczko, J., 1987, "A Practical Stress-Strain Rate Temperature Constitutive Relation of the Power Form," to appear in the *Journal of Mechanical Working Technology*, Vol. 12.

Litonski, J., 1977, "Plastic Flow of a Tube Under Adiabatic Torsion," *Bulletin de l'Académie Pololnaise des Sciences, Sciences Tech.*, Vol. 25, pp. 7–14.

Molinari, A., and Clifton, R. J., 1983, "Localisation de la Deformation Viscoplastique en Cisaillement Simple: Résultats Exacts en Théorie Non Linéaire," *Comptes, Rendus de l'Académie des Sciences*, Paris, Série II, Vol. 296, pp. 1–4.

Molinari, A., 1984, "Instabilité Viscoplastique en Cisaillement Simple," *5th Congrēs de Métallurgie et Mécanique de Tarbes*, ed. by ADISTA-ATS.

Molinari, A., 1985, "Instabilité Thermoviscoplastique en Cisaillement Simple," *Journal of Theoretical and Applied Mechanics*, Vol. 4, pp. 659–684.

Molinari, A., and Clifton, R. J., 1986, "Analytical Characterization of Shear Localization in Thermoviscoplastic Materials," Technical Report DAAG29-85-K-0003/4, Division of Engineering, Brown University, Providence, RI.

Pan, J., "Perturbation Analysis of Shear Strain Localization in Rate Sensitive Materials," *Int. J. Solids, Structures*, Vol. 19, 1983, pp. 153–164.

Shawki, T. G., Clifton, R. J., and Majda, G., 1983, "Analysis of Shear Strain Localization in Thermal Visco-Plastic Materials," Technical Report DAAG-29-81-0121/3, Division of Engineering, Brown University, Providence, RI.

Shawki, T. G., 1986, "Analysis of Shear Band Formation at High Strain Rates and the Visco-Plastic Response of Polycrystals," Ph.D. Thesis, Brown University, Providence, RI.

Tzavaras, A. E., 1984, "Shearing of Materials Exhibiting Thermal Softening or Temperature Dependent Viscosity," Technical Report LCDS #84-22, Division of Applied Mathematics, Brown University, Providence, RI.

Wright, T. W., and Batra, R. C., 1985, "The Initiation and Growth of Adiabatic Shear Bands," *Int. J. Plasticity*, Vol. 1, pp. 205–212.

Thermal Conditions for Inhibition of Stress Induced Slip in Zinc-Blende Crystals in Czochralski Growth Configuration

S. Motakef
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
Cambridge, MA 02139

The methodology for calculation of thermal conditions required for control of thermal stresses at below the plasticity limit during Czochralski growth of crystals with the zinc-blende structure is developed, and quantitative results are obtained for GaAs as the model crystal.

Introduction

Increased performance and yield of electronic devices based on compound semiconductors such as Gallium Arsenide (GaAs), and Indium Phosphide (InP), require improvement of the chemical and crystalline perfection of the presently available substrate materials. The primary crystalline defect in III-V compound semiconductors is the presence of a large density of dislocation networks which are associated with excessive thermal stresses experienced by the solid during growth. Whereas quantitative study of the plastic deformation mechanism of compound semiconductor crystals relating the absolute levels of thermal stresses in the growing solid and the observed dislocation densities requires a (presently unavailable) model for plastic strain release mechanism in compound semiconductors, modelling (Jordan et al., 1980); Kobayashi et al., 1985; Motakef and Witt, 1987; and Motakef, 1987b) and experimental (Seki et al., 1978; Uemura et al., 1981; and Jordan et al., 1984) results indicate that the stresses in III-V crystals grown by the liquid encapsulated Czochralski (LEC) technique significantly exceed the plasticity limit of the matrix.

Thermoelastic study of GaAs crystals grown by the LEC technique in conventional furnaces was first pioneered by Jordan (1980) and later improved and modified by Kobayashi and Iwaki (1985) through the exact solution of the thermoelastic equations and by Motakef and Witt (1987) and Motakef (1987a) by identification of the influence of the encapsulant on thermal stresses in the growing solid.

The present study is concerned with the development of a methodology for calculation of thermal conditions required for control of stresses in crystals with preferred orientations for plastic deformation. Results are obtained for GaAs crystals, characterized by the primary slip system $\{111\}\langle 1\bar{1}0\rangle$, grown by the LEC technique. The numerical calculations are conducted for a 75 mm (3 in.) diameter crystal with uniform thermal and elastic properties (Table 1).

Problem Formulation

The growing solid is modelled as an axisymmetric cylinder of uniform radius r_c with shoulder and growth interface geometries $S_n(r,z)$ and $S_i(r,z)$, respectively (Fig. 1(*a*)).

The temperature distribution in the crystal, used as the input to the thermoelastic equations, is obtained by the solution of the Fourier equation, $\nabla^2\Theta = 0$, subject to the following boundary conditions:

$$\partial\Theta/\partial r = 0. \quad \text{at } r = 0 \tag{1}$$

$$\partial\Theta/\partial r = q_s(z) \quad \text{at } r = 1 \tag{2}$$

$$\partial\Theta/\partial n = q_n \quad \text{at } S_n \tag{3}$$

$$\Theta = 1 \quad \text{at } S_i, \tag{4}$$

where the length scales are nondimensionalized by the crystal radius, Θ is the nondimensional temperature variable T/T_f (T_f is the freezing point temperature of the matrix), and q is the nondimensional heat flux variable $= \tilde{q}r_c/k_cT_f$ (the superscript $\sim$ denotes a dimensional quantity).

The stress distribution associated with the axial and radial temperature gradients in the crystal is obtained by solution of the thermoelastic equations. The radial and axial displacement components, u and w, respectively, in an axisymmetric and isotropic body with constant thermoelastic properties are (Boley and Weiner, 1960):

Table 1 Listing of thermoelastic parameters

GaAs	
Thermal Conductivity (W/cm K)	0.08
Surface Emissivity	0.5
Isotropic Thermoelastic Constant ($\alpha E/1-\nu$) (Dynes/cm^2 – K)	1.1×10^7
CRSS (Dynes/cm^2 – K)	$10^{(5.83+1382/T)}$
Encapsulant (Boric Oxide)	
Thermal Conductivity (W/cm K)	1.9×10^{-2}

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS DIVISION. Manuscript received by ASME Applied Mechanics Division, February 19, 1987; final revision May 27, 1987.

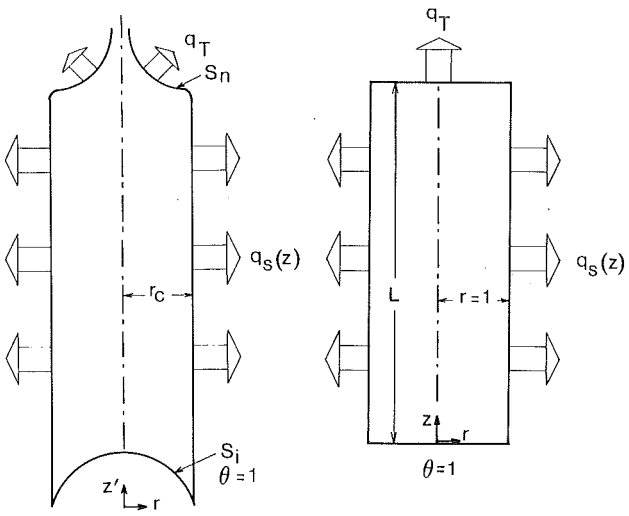


Fig. 1 Geometric model of (a) axisymmetric crystal including the cone and growth interface morphologies, and (b) cylindrical model of the crystal

$$\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}-\frac{u}{r^2}+\frac{\partial^2 u}{\partial z^2}+\frac{1}{1-2\nu}\frac{\partial e}{\partial r}=\frac{2(1+\nu)}{1-2\nu}\alpha T_f\frac{\partial \theta}{\partial r} \tag{5}$$

$$\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}+\frac{\partial^2 w}{\partial z^2}+\frac{1}{1-2\nu}\frac{\partial e}{\partial z}=\frac{2(1+\nu)}{1-2\nu}\alpha T_f\frac{\partial \theta}{\partial z} \tag{6}$$

where ν is the Poisson ratio, α is the thermal expansion coefficient of the crystal, and e is the dilatation parameter:

$$e=\frac{\partial u}{\partial r}+\frac{u}{r}+\frac{\partial w}{\partial z} \tag{7}$$

The thermal stress tensor, $\mathbf{T}$, for the geometry under consideration is:

$$\mathbf{T}=\begin{bmatrix}\sigma_r & 0 & \tau_{rz}\\ 0 & \sigma_\varphi & 0\\ \tau_{zr} & 0 & \sigma_z\end{bmatrix}. \tag{8}$$

The elements of $\mathbf{T}$ are calculated from Hooke's strain-stress relationship:

$$\tilde{\sigma}_r=\frac{E}{1+\nu}\left[\frac{\partial u}{\partial r}+\frac{\nu}{1-2\nu}e-\frac{1+\nu}{1-2\nu}\alpha T_f(\theta-1)\right] \tag{9}$$

$$\tilde{\sigma}_\phi=\frac{E}{1+\nu}\left[\frac{u}{r}+\frac{\nu}{1-2\nu}e-\frac{1+\nu}{1-2\nu}\alpha T_f(\theta-1)\right] \tag{10}$$

$$\tilde{\sigma}_z=\frac{E}{1+\nu}\left[\frac{\partial w}{\partial z}+\frac{\nu}{1-2\nu}e-\frac{1+\nu}{1-2\nu}\alpha T_f(\theta-1)\right] \tag{11}$$

$$\tilde{\tau}_{rz}=\frac{E}{2(1+\nu)}\left(\frac{\partial w}{\partial r}+\frac{\partial u}{\partial z}\right) \tag{12}$$

In the above E is the Young's modulus of the crystal. Equations (9)–(12) are subject to the boundary conditions of traction free surfaces and zero radial displacement at the centerline.

The geometry of the crystal is reduced to a cylinder with flat top and bottom surfaces while the influence of the shoulder and growth interface geometries is retained by the introduction of a radially nonuniform heat flux distribution, $q_T(r)$, at the cylinder top, a radially nonuniform temperature distribution, $\Theta_B(r)$, at the cylinder's bottom, and radially nonuniform traction boundary conditions $\tilde{\mathbf{X}}_T(r)$ and $\tilde{\mathbf{X}}_B(r)$ at the cylinder's top and bottom. Thus, the thermal and traction boundary conditions for the cylindrical model of the crystal are:

$$\partial\Theta/\partial r=0, \quad u=0 \quad \text{at } r=0 \tag{13}$$

$$\partial\Theta/\partial r=q_s(z), \quad \tilde{\sigma}_r=\tilde{\tau}_{rz}=0 \quad \text{at } r=1 \tag{14}$$

$$\Theta=\Theta_B(r), \quad \tilde{\mathbf{X}}=\tilde{\mathbf{X}}_B(r) \quad \text{at } z=0 \tag{15}$$

$$\partial\Theta/\partial z=q_T(r), \quad \tilde{\mathbf{X}}=\tilde{\mathbf{X}}_T(r) \quad \text{at } z=L \tag{16}$$

The above boundary conditions indicate that the stress field in the crystal is controlled by: (1) the heat flux distribution at the crystal periphery, equation (14); (2) growth interface morphology, equation (15); and (3) the shoulder geometry, equation (16). Although the primary influence of the shoulder and growth interface geometries on the stress levels in the crystal is restricted to the areas close to the two regions, generation of excessive stresses at these areas can significantly influence the crystalline perfection of the growing solid independent of the following stresses generated by the heat loss at the crystal periphery: (*a*) plastic deformation of the crystal shoulder results in generation of dislocations that penetrate into the crystal core, and (*b*) the low values of CRSS close to the growth interface provide the potential for plastic deformation of the crystal at low stress levels. Therefore, equations (14)–(16) suggest that the results of thermoelastic studies of crystals which exclude the thermal and traction boundary conditions of equations (15) and (16) are inaccurate close to the growth interface and the shoulder area.

The present study is concerned with maintaining the thermal stresses in the growing solid at below the plasticity limit through control of heat losses from the crystal periphery and, thus, the influences of the shoulder and interface morphologies on the stress distribution in the crystal are excluded. Therefore, the thermal and elastic boundary conditions, equations (14)–(17), are modified to exclusively reflect the influence of heat loss from the crystal periphery (Fig. 1(*b*)): the crystal top and bottom are assumed to be flat and traction free, the growth interface is taken to be at the freezing point, and the heat flux at the crystal top is radially uniform (a radially uniform heat flux at the cylinder's top results in a constant axial temperature gradient in the solid which, ignoring variations in the thermoelastic properties of the matrix with temperature, does not result in generation of thermal stresses in the crystal).

The new boundary conditions are:

$$\partial\Theta/\partial r=0, \quad u=0 \quad \text{at } r=0 \tag{19}$$

$$\partial\Theta/\partial r=q_s(z), \quad \tilde{\sigma}_r=\tilde{\tau}_{rz}=0 \quad \text{at } r=1 \tag{20}$$

$$\Theta=1, \quad \tilde{\sigma}_z=\tilde{\tau}_{rz}=0 \quad \text{at } z=0 \tag{21}$$

$$\partial\Theta/\partial z=q_T, \quad \tilde{\sigma}_z=\tilde{\tau}_{rz}=0 \quad \text{at } z=L \tag{22}$$

Solutions

The locus of thermal boundary conditions resulting in generation of stresses smaller than the plasticity limit of the crystal is obtained, for the geometry defined in the previous section, through calculation of the maximum allowable heat fluxes at the crystal periphery, $q_s(z)$. Characterization of the permissible thermal boundary conditions in terms of $q_s(z)$, as opposed to the calculation of allowable axial temperature distribution at the crystal's surface ($T_s(z)$), is advantageous for two reasons. First, $q_s(z)$ respresents radial temperature

gradients at the crystal periphery and, thus, is directly related to thermal stresses in the crystal providing for significant simplification of the solution methodoly. Second, calculation of allowable families of $T_s(z)$ rquires Taylor expansion of $T_s(z)$ to quantify the nonlinear axial variation of the crystal surface temperature which significantly increases the complexity of the solution scheme.

Thermal stresses in the growing solid are generated by axial and radial temperature gradients and, ignoring variations of thermoelastic properties of the crystal with temperature, are not affected by the absolute levels of temperature in the crystal. Furthermore, as linear axial temperature gradients do not generate thermal stresses, the stresses are invariant with respect to the magnitude of a radially uniform heat flux at the crystal top, q_T. Therefore, the temperature distribution required for the calculation of thermal stresses in the crystal is obtained by imposing adiabatic boundary conditions at the solid's top, $q_T = 0$:

$$\theta = 1 - q_{so} \sum_n \left[\frac{I_o(\alpha_n r) \sin(\alpha_n z)}{I_1(\alpha_n) \dfrac{\alpha_n L}{2}} \int_o^L \frac{q_s(z)}{q_{so}} \sin(\alpha_n z) dz \right] \quad (21)$$

where

$$\alpha_n = \frac{(2n+1)\pi}{2L} \quad n = 0,1 \ldots, \quad (22)$$

$$q_{so} = \frac{1}{L} \int_o^L q_s(z) dz, \quad (23)$$

and I_i is the modified Bessel function of the first kind of order i. The temperature distribution in the crystal may be succinctly written as:

$$\Theta = 1 - q_{so} f(r,z;L,q_s(z)/q_{so}). \quad (24)$$

Plastic deformation in GaAs is presently considered to be associated with glide in the $\{111\}\ <1\bar{1}0>$ slip system which has twelve permissible operations. Once the resolved shear stress in each of the twelve slip systems exceeds a critical value, the so-called critical resolved shear stress (CRSS), plastic deformation of the crystal is to be expected. The resolved shear stress in each of the slip operations, $\tilde{\sigma}^i_{rs}$, is obtained by projection of the principal stresses onto the appropriate crystallographic planes and, thus, it is a linear combination of the principal stresses:

$$\tilde{\sigma}^i_{rs} = L_i(\tilde{\sigma}_r, \tilde{\sigma}_\phi, \tilde{\sigma}_Z, \tilde{\tau}_{rz}) \quad i = 1,12, \quad (25)$$

where L_i is the linear operator associated with the ith slip operation. The thermal stress tensor is obtained by the solution of the thermoelastic equations in conjunction with the calculated temperature distribution, equation (24). As the thermoelastic equations are linear in temperature (for a constant property system) the resovled shear stress for each slip direction in the crystal can be written as:

$$\tilde{\sigma}^i_{rs} = \frac{\alpha E T_f}{1-\nu} L_i H(\Theta) \qquad i = 1,12 \quad (26)$$

In the above $H(\Theta)$ represents the linear operation associated with the solution of the thermoelastic equations which yields the thermal stress tensor $\mathbf{T}$. Equation (26) suggests nondimensionalization of the stresses as $\sigma = \tilde{\sigma}/P$ and P is the property parameter $\alpha E T_f/(1-\nu)$. Using equation (24) and recognizing that a uniform temperature field does not result in generation of thermal stresses ($H(\Theta = 1) = 0$), equation (26) becomes:

$$\sigma^i_{rs} = -q_{so} L_i H\{f(r,z: L,q_s(z)/q_{so})\} \quad i = 1.12, \quad (27)$$

where φ is the azimuthal direction in the radial plane of the crystal and reflects the anisotropy of the slip system. Thus, the resolved shear stress distribution in the crystal is uniquely defined by the solid's aspect ratio L, the periphereal heat flux distribution $q_s(z)$, and the property of parameter P.

The maximum allowable heat flux at the crystal periphery, $q_s^*(z)$, may be calculated by constraining the magnitude of resolved shear stress in each of the twelve slip operations to be less than the plasticity limit σ_Y. This large number of constraints can be reduced to one by an averaging approximation whereby the sum of the absolute values of the resolved shear stresses at each location, σ_T, is less than $12\sigma_Y$:

$$\sigma_T = \sum_{i=1}^{12} |\sigma^i_{rs}| = 12\sigma_Y, \quad (28)$$

In certain locations in the crystal where the resolved shear stress exceeds the plasticity limit in *some* of, but *not in all* of, the 12 slip operations, the above averaging approximation underestimates the actual driving forces for plastic deformation of the crystal. Therefore, equation (28) will result in the overestimation of maximum allowable heat fluxes at the crystal periphery.

In the radial plane the maximum value of σ_T occurs at the periphery of the solid ($r = 1$), and (for the $<100>$ growth direction) at the angle $\phi_o = n\pi/2$, $n = 0-3$. The requirement for maintaining the stress levels in the crystal to below the plasticity limit is satisfied by constraining the maximum stresses in the radial plane (occuring at $r = 1$, $\varphi = 0$) to be less that $12\sigma_Y$:

$$\sigma_T(r = 1,z,\varphi = 0) \leq 12\sigma_Y \quad (29)$$

The heat flux distribution $q_s^*(z)$ is related to the plasticity limit by combining equations (26), (28), and (29):

$$q^*_{so} = 12\sigma_Y/F, \quad (30)$$

where

$$F = \sum_{i=1}^{12} |L_i H\{f(r,z,\varphi;L,q_s^*(z)/q^*_{so})\}|_{r=1,\,\varphi=0} \quad (31)$$

Calculation of $q_s^*(z)$ requires quantification of the plasticity limit. Presently two different types of plasticity limits are considered: (1) an athermal constraint where at all axial locations $\sigma_T(r = 1,z,\phi = 0)$ is restrained to be equal to a constant value, and (2) a temperature-dependent constraint where $\sigma_T(r = 1,z,\varphi = 0)$at each axial location is contained by the local magnitude of the temperature-dependent plasticity limit. The temperature-dependent values of CRSS (Jordan, 1980) are used as the temperature-dependent plasticity limit, and the minimum values of CRSS (occurring at the melting point of the matrix) is chosen as the athermal constraints (CRSS values are nondimensionalized, similar to the stresses, by the property parameter P). As CRSS is a decreasing function of temperature the athermal constraint underestimates the resistance of the crystal to plastic deformation at locations away from the growth interface and results in the calculation of a lower bound to $q_s^*(z)$ (denoted as $q_s^{**}(z)$). The choice of the two plasticity limits has significant implications for the practical implementation of the calculated heat fluxes and is later discussed.

It must be noted that although the presented methodology is for a plasticity model based on glide in the slip system, any criterion for nucleation of dislocations which is a linear function of the principal thermal stresses can be used to arrive at a formulation similar to equations (30) and (31).

Calculation of $q_s^{**}(z)$. The heat flux distribution $q_s^{**}(z)$ is obtained by contraining $\sigma_T(r = 1,z,\varphi = 0)$ at all axial locations to be equal to ECRSS (Effective CRSS = 12CRSS) evaluated at the freezing point of the matrix:

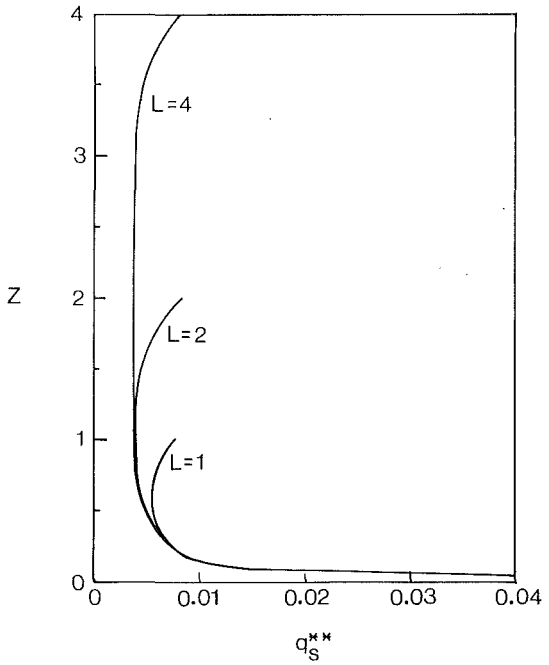


Fig. 2 The profiles of $q_s^{**}(z)$ for a GaAs crystal at aspect ratios 1, 2, and 4

$$\sigma_T(r=1,\varphi=0,z)=\text{ECRSS}(\Theta=1) \quad \forall z \tag{32}$$

As constraint (32) is independent of the temperature distribution in the crystal it is (similar to the thermal stress distribution in the crystal) invariant with respect to the magnitude of the radially-uniform heat loss at the crystal top (q_T). Therefore, considering equations (30) and (31), $q_s^{**}(z)$ is, for a given material, only a function of the crystal's aspect ratio.

The calculation of $q_s^{**}(z)$ is achieved by the simultaneous solution of equations (30) and (31) in conjunction with equations (21) and (32), where the function F is calculated by the numerical solution of the thermoelastic equations and the tensor transformations relating resolved shear stresses to the principal stresses.

The axial distribution of $q_s^{**}(z)$ for various aspect ratios of a GaAs crystal growing in the $<100>$ direction is presented in Fig. 2. The results indicate that the influence of the end effects persists up to approximately one-half of the crystal radius and, thus, for crystals with aspect ratio of less than two the growth interface and the cone morphologies control the stress levels in the crystal. The increase of $q_s^{**}(z)$ in the direction towards the growth interface reflects the assumed conditions of a flat growth interface ($\partial\Theta/\partial r=\mathbf{X}_B=0$ at $z=0$), and suggests that if such conditions can be maintained (during growth) the maximum values of $q_s^{**}(z)$ occur close to the growth interface. In practice, however, as a result of the thermal coupling between the solid and the melt large heat losses from the crystal surface at $z \sim 0$ result in a nonplanar growth morphology and the assumed boundary conditions (planar solidification interface) can not be sustained. Therefore, the observed increase in $q_s^{**}(z)$ towards $z \sim 0$ in an artifact of the assumed boundary conditions and suggests that the accurate calculation of $q_s^{**}(z)$ close to the solidification front requires incorporation of the growth interface morphology in the geometric model of the crystal. Towards the crystal top $q_s^{**}(z)$ also increases to a large value. Nevertheless, $q_s^{**}(z=L) < q_s^{**}(z=0)$ reflecting the difference in the imposed thermal boundary conditions at the solid's two ends: the assumed uniform temperature distribution at the crystal bottom results in zero radial temperature gradients and an associated reduction of local thermal stresses whereas the imposed condition of uniform heat flux at the crystal top does permit presence of radial thermal gradients (generated by the peripheral heat losses) in that region. The calculation results also indicate that, consistent with the imposed constraint of uniform total resolved shear stress at the crystal periphery, $q_s^{**}(z)$ away from the two ends of the crystal is approximately constant.

Calculation of $q_s^*(z)$. The maximum allowable heat flux at the crystal periphery, $q_s^*(z)$, is calculated by requiring $\sigma_T(r=1,z,\varphi=0)$ at every axial location along the crystal be equal to ECRSS evaluated at the local temperature:

$$\sigma_T(r=1,\varphi=0,z)=\text{ECRSS}(\Theta(r=1,z)) \tag{33}$$

The above constraint is directly tied to the absolute levels of temperature in the crystal and is, therefore, influenced by the magnitude of heat loss at the crystal top. The temperature distribution in the solid subject to the thermal boundary conditions of equations (17)–(20) ($q_T=0$) is obtained by the superposition of a linear axial temperature drop onto equation (21):

$$\theta=1-q_{so}\sum_n\left[\frac{I_o(\alpha_n r)\,\sin(\alpha_n z)}{I_1(\alpha_n)\,\frac{\alpha_n L}{2}}\int_o^L\frac{q_s(z)}{q_{so}}\sin(\alpha_n z)dz\right]-q_T z \tag{34}$$

In order to closely relate the calculations in this work to the Czochralski growth conditions, the crystal top is taken to radiate to a black environment at temperature Θ_T^*:

$$q_T=Rn_c(\Theta_T{}^4-\Theta_T^{*4}) \tag{35}$$

where Θ_T is the average temperature at the crystal top:

$$\Theta_T{}^4=\int_0^1\Theta^4(r,z=L)\;2rdr \tag{36}$$

and

$$Rn_c=\frac{\epsilon\sigma^*T_f^3r_c}{k_c} \tag{37}$$

In equation (37), σ^* is the Stefan-Boltzmann constant and ϵ is the surface emissivity of the crystal. The heat flux at the crystal top, equation (35), is, consistent with the thermal boundary conditions of equation (21), radially uniform. By considering the parameters which influence the temperature distribution in the crystal, equation, (30) is reformulated into the following form:

$$q_{so}^*=\text{ECRSS}\{\Theta(r=1,z;L,q_s^*(z)/q_{so}^*,\Theta_T^*,Rn_c)]\}/F \tag{38}$$

The above indicates that $q_s^*(z)$ is parametrically dependent on, in addition to the crystal aspect ratio L and the property parameter P (implicit in the ECRSS term), the magnitude of heat loss for the crystal top.

Calculation of $q_s^*(z)$ is achieved by the iterative solution of equations (31), (34), and (38) in conjunction with the temperature-dependent plasticity limit for GaAs:

$$\text{CRSS}=[10^{(5.83+.92/\Theta)}]/P \tag{39}$$

The calculated profiles of $q_s^*(z)$ for 3 crystal aspect ratios of 1, 2, and 4 over a range of Θ_T^* are presented in Figs. 3–5. Beyond the zone of influence of the end effect at the crystal bottom the magnitude of $q_s^*(z)$ increases towards the crystal top, indicating that as the temperature along the crystal decreases (increasing ECRSS) the maximum magnitude of heat flux that can be removed from the crystal surface while

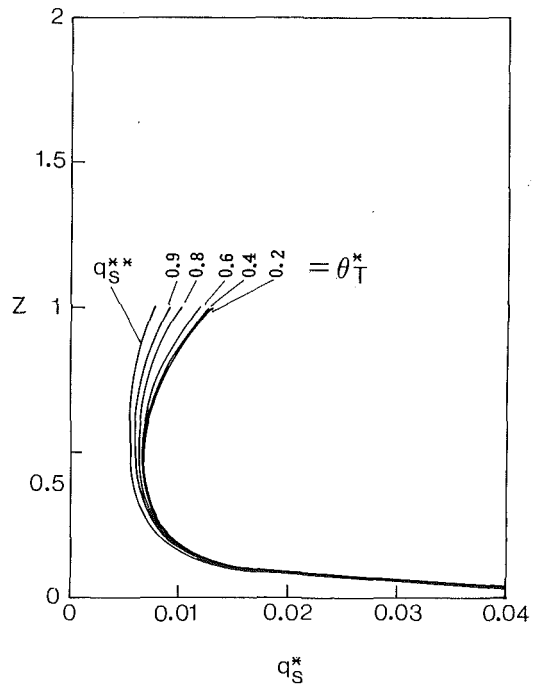


Fig. 3 The profiles of $q_S^*(z)$ for a GaAs crystal with $L=1$ at different values of Θ_T^*

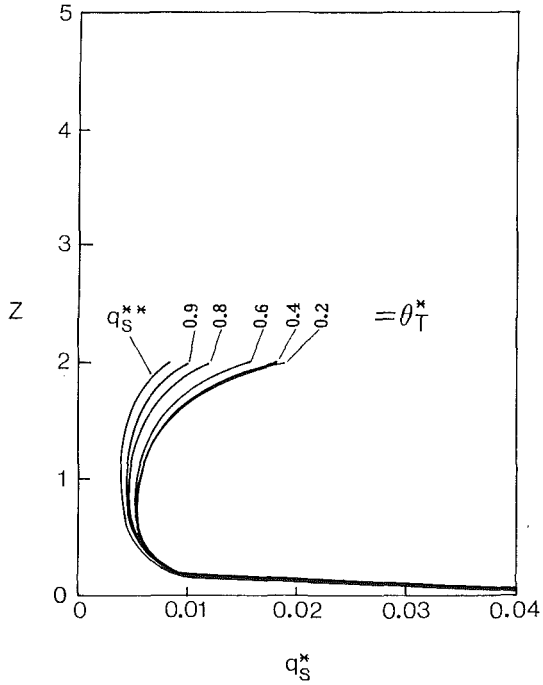


Fig. 4 The profiles of $q_S^*(z)$ for a GaAs crystal with $L=2$ at different values of Θ_T^*

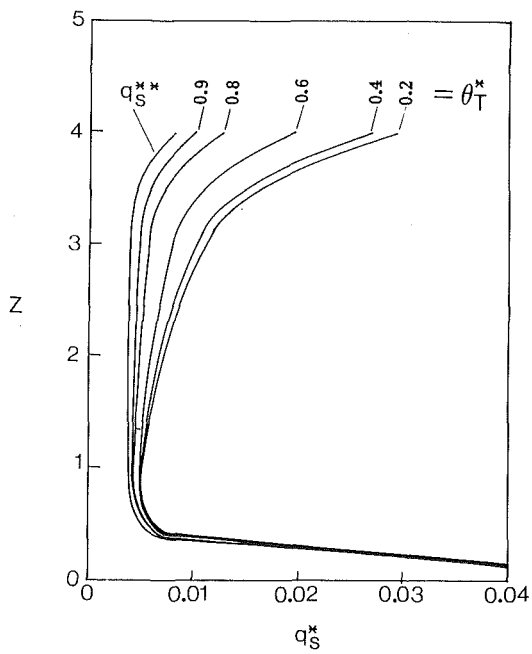


Fig. 5 The profiles of $q_S^*(z)$ for a GaAs crystal with $L=4$ at different values of Θ_T^*

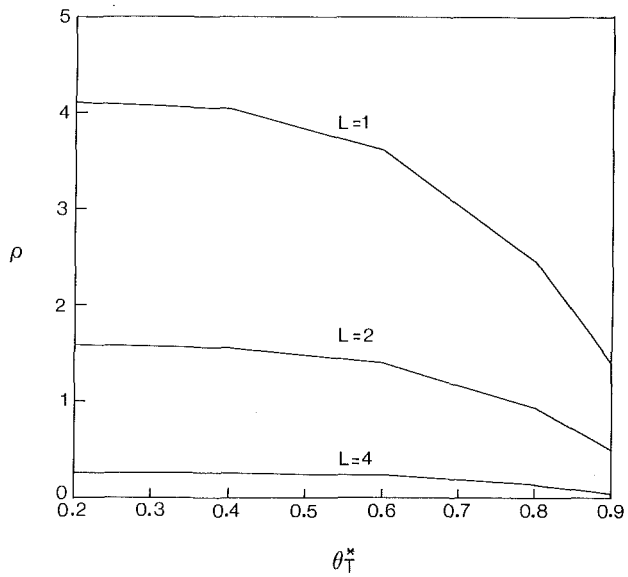


Fig. 6 The ratio of heat loss from the crystal top to the heat loss from its periphery for different lengths of GaAs crystal versus Θ_T^*

maintaining the peripheral stresses at the plasticity limit increases. The increase in $q_s^*(z)$ above its lower bound, $q_s^{**}(z)$, essentially reflects the temperature dependence of CRSS and is a function of the environmental temperature that the crystal top radiates to. It can be readily observed that a reduction in Θ_T^* leads to an overall reduction of the crystal temperature resulting in lower values of ECRSS and $q_s^*(z)$. The reduction in the rate of increase of $q_s^*(z)$ with decreasing values of Θ_T^* reflects the nonlinear dependence of q_T on Θ_T^* (equation (35)) which asymptotically becomes independent of Θ_T^* as $\Theta_T^* \to 0$.

The relative magntiude of heat loss at the crystal top to that at the crystal periphery is investigated in Fig. 6, where the ratio of the heat loss at the crystal top to the peripheral losses, $\rho = \int_0^1 q_T 2\, rdr / \int_0^1 q_s^*(z) 2\, dz$, is plotted versus Θ_T^*. The results indicate that during the early stages of growth the primary heat loss mechanisms from the crystal is from its top. Therefore, reduction of the temperature to which the crystal top radiates is an important factor in maximization of total heat loss from the crystal. Under such conditions the temperature distribution in the crystal is essentially determined by the magnitude of heat loss at its top. It must be noted, however, that the axial variation of the crystal diameter in the cone area results in generation of temperature gradients and associated thermal stresses in that region which limits the maximum rate of heat transfer at the crystal top. At $L=1$, the centerline temperature of a crystal radiating to $\Theta_T^*=0.2$ exhibits a linear axial variation (Fig. 7) indicating that the temperature distribution in the crystal is one-dimensional and determined by the magnitude of radiative losses at its top. The effect of heat loss at the

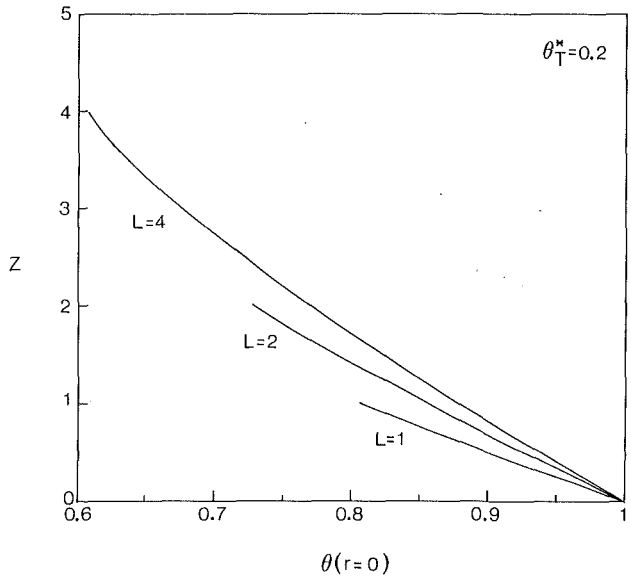


Fig. 7 The axial distribution of the centerline temperature in a GaAs crystal at aspect ratios of 1, 2, and $\Theta_T^* = 0.2$

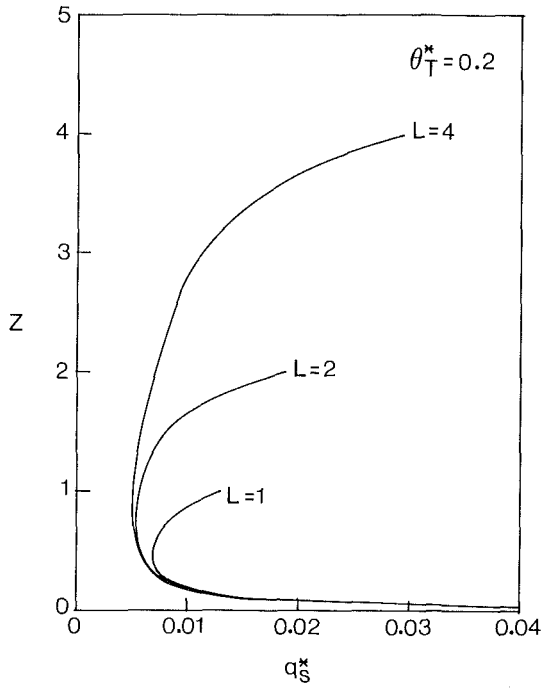


Fig. 8 The axial distribution of $q_S^*(z)$ for a GaAs crystal at aspect ratios of 1, 2, and 4, $\Theta_T^* = 0.2$

crystal periphery on the temperature distribution at $L = 2$ appears as slight deviations of the centerline temperature from linearity, and at $L = 4$ the axial variation of the centerline temperature does not exhibit a linear variation; thus, only as the crystal length approaches 4 radii the peripheral heat losses influence the overall temperature distribution in the solid. During LEC growth of GaAs the crystal top radiates to temperatures which are significantly higher than 0.2 indicating that the deviation of the heat transfer in the crystal from quasi one-dimensional to two-dimensional occurs at crystal lengths smaller than 4 radii. With increasing crystal length the conductive path to the crystal top increases and the solid's temperature at fixed axial positions increases (Fig. 7), resulting in the reduction of local values of ECRSS and $q_s^*(z)$. The axial variations of $q_s^*(z)$, plotted for the three crystal

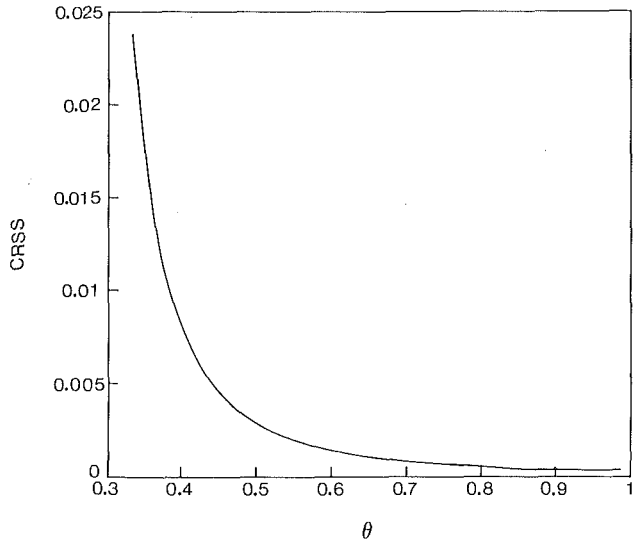


Fig. 9 Temperature dependence of CRSS for GaAs (Motakef and Witt, 1986)

aspect ratios (at $\theta_T^* = 0.2$) in Fig. 8, demonstrate that at fixed axial positions the magnitude of maximum allowable heat flux from the crystal surface decreases with increasing crystal length.

The large calculated values of $q_s^*(z)$ (as well as those of $q_s^{**}(z)$) close to the solidification interface is a result of the assumed planarity of the growth interface. In practice, however, removal of such large fluxes close to the solidification interface can be expected to lead to a concave growth interface and an associated generation of stresses in that region. Nevertheless, the heat flux from the periphery of the crystal constitutes a minor portion of the total heat transfer for crystals with $L < 4$ (Fig. 7) and for $L = 4$ the heat transfer from the crystal's lower one-eighth does not influence the temperature distribution in the crystal significantly. In order to investigate the influence of the large calculated fluxes at $z \sim 0$ on the temperature distribution in the crystal, and consequently on the magnitude of $q_s^*(z)$ beyond $z = 0$ through the temperature dependance of CRSS, the temperature distribution in the solid and $q_s^*(z)$ were recalculated (for the 3 crystal aspect ratios of 1, 2, and 4) by constraining the heat flux close to the growth interface to be a constant value and equal to 0.005 (a representative value of $q_s^*(z)$ beyond $z = 0$). The results indicate that the reduction of heat losses from the crystal at $z \sim 0$ does not appreciably alter the temperature distribution in the crystal and the calculated profiles of $q_s^*(z)$ are essentially unchanged. Therefore, it can be concluded that the assumed condition of planar growth interface does not influence the magnitude of calculated maximum allowable heat fluxes beyond $z \sim 0$.

Critical Crystal Length

The temperature dependence of CRSS for GaAs, Fig. 9, indicates that CRSS is relatively constant at temperatures close to the melting point of the matrix and increases exponentially with decreasing temperature for $\Theta < 0.6$. Beyond a critical length of the crystal, L_C, the temperature of the upper portion of the solid enters the region of rapid variation of CRSS with temperature and, thus, as the crystal length exceeds L_c the calculated magnitude of $q_s^*(z)$, associated with the local values of CRSS, increases significantly. Such a behavior suggests that beyond L_C control of heat transfer from the crystal is not required and that the growing solid can be safely allowed to thermally interact with the walls of the growth furnace.

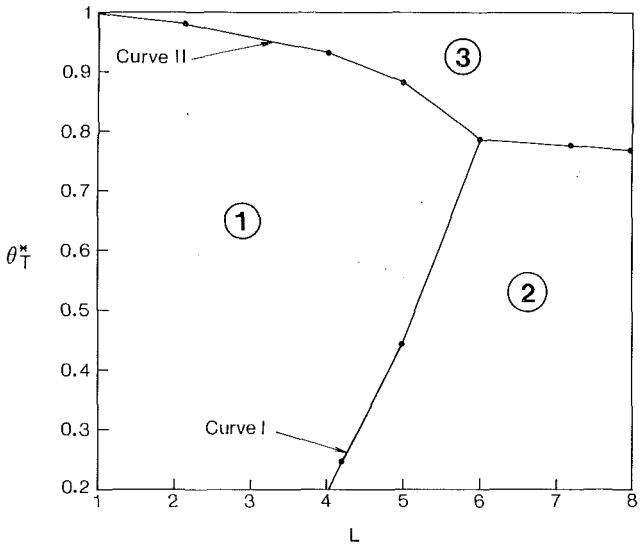


Fig. 10 Theee regions in the $\Theta_T^* - L$ plane for GaAs: (1) where control of heat transfer at the crystal periphery is required; (2) where the crystal can be allowed to thermally interact with the environment at any temperature; (3) where the crystal is heated at its top by the environment

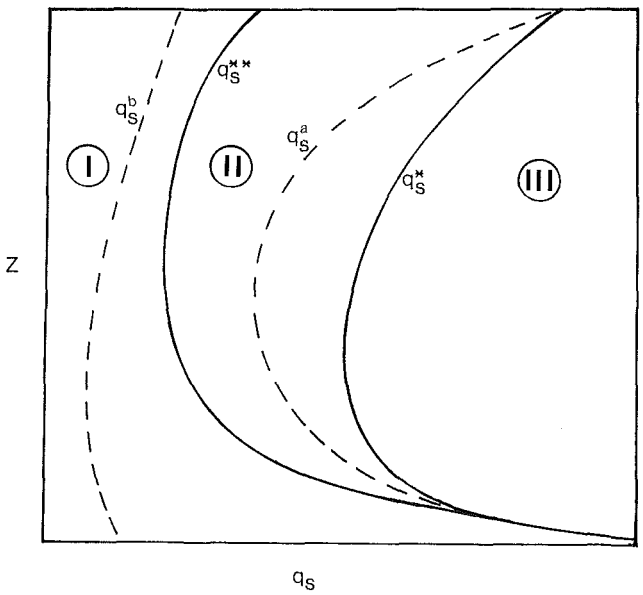


Fig. 11 The regions of operation in the q_s—z plane

The magnitude of the critical crystal length is determined by the heat loss at the solid's periphery and top. As the maximum allowable heat transfer at the periphery of the crystal ($q_s^*(z)$) is a function of the temperature to which the crystal top radiates (equation (38)), the critical crystal length is uniquely determined by Θ_T^*. The calculated locus of (L_c, Θ_T^*)) is shown in Fig. 10 as curve *I*. The results indicate that the smallest critical length is about 4.1 (corresponding to the minimum value of $\Theta_T^* = 0.2$) and L_c increases with Θ_T^* (decreasing heat transfer at the crystal top).

Radiative cooling of the crystal requires that the average temperature at the crystal top (Θ_T) be larger than the environmental temperature to which it is radiating (Θ_T^*). Therefore, the envionmental temperature Θ_T^* must be always smaller than the value of Θ_T calculated for a crystal with an adiabatic top and experiencing a heat transfer rate of $q_s^*(z)$ at its periphery. The locus of (L,Θ_T^*) associated with the calculated profiles of $q_s^*(z)$ is shown as curve *II* in Fig. 10. As $L \to 6$ curves *I* and *II* intersect indicating that the crystal has reached such a length that with no heat transfer at its top ($\Theta_T^* = \Theta_T$) it is sufficiently cooled so that there are no upper limits on the heat transfer from its periphery. The region bounded by curves *I*, *II* and the two axes (region 1 in Fig.10) is the locus of $L - \Theta_T^*$ where the reduction of the stress levels in the crystal requires control of heat transfer at the solid's periphery. The region to the right of curve *I* and below curved *II* (region 2 in Fig. 13) respresent the locus of $L - \Theta_L^*$ where the heat transfer from the crystal does not need to be controlled. For example, at an environmental temperature of $\Theta_T^* = 0.5$ control of heat transfer at the crystal periphery must be maintained only for the first 5 radii of the crystal; as the crystal aspect ratio exceeds 5 the growing solid may be safely allowed to interact with the environment at any temperature. In the region curve *II* (region 3 in Fig. 10) the crystal is radiatively heated by the environment ($\Theta_T^* > \Theta_T$) and is, therefore, in a nonoptimal region of operation.

Regions of Operation

During growth of semiconductor crystals, maximization of total heat transfer from the crystal surface, subject to the thermal stress constraints, is advantageous from the point of view of increased growth rate as well as decreased interaction and clustering of point defects through reduction of the time-temperature integral of the growing matrix. Therefore, the heat flux distribution at the crystal periphery must be maintained at the calculated values of $q_s^*(z)$ and dynamically controlled as the crystal length increases. However, during growth the establishment and dynamic control of such distributions may be impractical and deviations from the prescribed profiles of $q_s^*(z)$ appear unavoidable. In this section the effect of deviations from the calculated profiles of $q_s^*(z)$ on thermal stresses in the crystal will be investigated.

For a crystal subject to some heat flux distribution ($q_s(z)$) at its periphery, the two heat flux distributions $q_s^*(z)$ and $q_s^{**}(z)$ define, for each value of L and Θ_T^*, three possible regions of operation, (Fig. 11):

Region I; $0 < q_s(z) < q_s^{**}(z)$

Region II: $q_s^{**}(z) < q_s(z) < q_s^*(z)$

Region III: $q_s^*(z) < q_s(z)$

The first two regions define the families of allowable heat flux distributions and operation in the third region must be avoided. The heat flux distribution $q_s^*(z)$ is calculated using a temperature-dependent plasticity limit (i.e., local values of CRSS) and, therefore, its values at any axial location along the crystal is dependent on the temperature distribution in the solid, and consequently on the entire axial profile of $q_s^*(z)$. Thus, an imposed heat flux distribution in region II which deviates from the calculated profile of $q_s^*(z)$ will result in a higher temperature and lower CRSS distribution in the crystal than the one associated with $q_s^*(z)$, and the plasticity limit in the crystal may be reduced to below the prevailing stress levels in the crystal resulting in generation of dislocations. In this context, the hypothetical distribution $q_s^a(z)$ located in region II of Fig. 11 (which is smaller than $q_s^*(z)$ close to the solidification interface and approaches it towards the crystal top) results in a temperature distribution in the crystal which at all locations is higher than the one associated with $q_s^*(z)$, leading to lower values of ECRSS throughout the solid; close to the crystal top as $q_s^a(z)$ approaches $q_s^*(z)$, the generated thermal stresses exceed the local values of ECRSS. On the other hand, the heat flux distribution $q_s^{**}(z)$ is calculated subject to an athermal plasticity limit (CRSS evaluated at the melting point of the matrix) and is invariant with respect to the absolute levels of temperature in the crystal. Stresses generated in a crystal experiencing any heat flux distribution

less than $q_s^{**}(z)$ (i.e., located in region I, such as $q_s^b(z)$ in Fig. 11) are everywhere smaller than the local plasticity limit. Hence, a significant distinction between operation in regions I and II exists: in region I any heat flux distribution can be safely removed from the crystal surface, whereas operation in region II is contingent on the availability of the necessary control hardware to closely emulate the calculated axial variations of $q_s^*(z)$. The complete definition of the characteristics of allowable families of $q_s(z)$ in region II, where the total heat transfer from the crystal is not maximized but the stress levels are maintained below the plasticity limit, is complicated by the nonlinear dependence of CRSS on temperature. However, as a broad guideline it may be noted that a heat flux distribution that is lower than $q_s^*(z)$ over some length of the crystal cannot be equal to it at other axial locations.

Discussion

The present study provides the dynamic conditions necessary for inhibition of plastic deformation in crystals with the $\{111\}\ <1\bar{1}0>$ slip system in general and GaAs in particular from after shouldering to the critical crystal length, beyond which control of heat transfer from the growing solid is not required. Results indicate that the morphologies of the cone, during the early stages of growth, and the solidification front, at all stages of growth, strongly influence the crystalline quality of the growing solid. Thus, reduction of the temperature gradients associated with the nonplanarity of the growth interface appears as a primary control objective; identification of the optimum cone geometry cannot be, however, readily deduced from the present analysis. In this context it must be noted that in the early stages of growth the cone influences, through controlling the radial distribution of heat transfer from the crystal top, the morphology of the growth interface. Furthermore, whereas results of the thermoelastic analysis of Szabo (Szabo, 1985), conducted for a crystal with a planar growth interface, suggest that the tendency for plastic deformation at the early stages of growth ($z < 0.2$) are largest at the crystal centerline and not at its periphery, the observed nonplanar geometry of the solidification interface in grown crystals indicates that identification of the mechanism of dislocation generation and the necessary conditions for minimization of stresses in this region require calculation of the growth interface morphology. Therefore, complete characterization of the optimum thermal environment for Czochralski growth of dislocation-free crystals mandates development of a macroscopic thermal model of the growth furnace, whereby accurate calculation of the heat transfer from the crystal cone and the shape of the solid-liquid interface can be achieved (Motakef, 1987b).

During growth, heat transfer from the crystal surface and top is controlled by the thermal coupling of the crystal (radiative in low pressure systems and radiative and convective in high pressure systems) to the crucible, encapsulant and melt surfaces, and the furnace walls. In conventional growth furnaces three primary controllers are available: (*i*) the crystal pull rate, (*ii*) the power input to the heater, and (*iii*) the location of the crucible containing the charge inside the heater. Of these the last two may be adjusted to control the "environmental" temperature distribution of the crystal, and the first to control the crystal radius during growth. Calculation of the required dynamic control trajectory of the crucible position and the heater power to insure operation in regions 1 or 2 of Fig. 11 is complicated by the transient (batch) nature of the growth process, the physical complexity of the growth furnace, and the need for accurate quantification of the evolution (during growth) of thermophysical properties such as the emissivity of the interacting surfaces, and transparency of the liquid encapsulant to infrared radiation. There are no indications at present to suggest that dynamic control of heater power and crucible position is sufficient to provide for control of thermal stresses at below the plasticity limit of the matrix: as of yet such a parameter space has not been empirically identified and modelling results (Motakef, 1987b) suggest that the optimum thermal conditions cannot be achieved by the adjustment of heater power. Elimination of the transient characteristics of the growth environment and the thermal coupling of the crystal with its environment can be achieved by isolation of the crystal from the crucible and furnace walls through installation of a coaxial heat exchange system around the growing solid. The calculation results indicate that the temperature distribution of the inner surface of such a system must be dynamically controlled with increasing crystal length. Furthermore, in addition to controlling the heat transfer from the crystal periphery, heat losses from the crystal top must be enhanced. In this context use of conical heat-reflectors, placed on top of the crystal, is not an optimal approach, for the heat transfer from both the side and top of the crystal is reduced. The calculated narrow operating range suggests that use of heat reflectors or uniform-temperature after heaters cannot provide the required spatial and temporal controllability of the growth process, and new approaches to heat transfer control during LEC growth of GaAs must be considered.

The reported thermoelastic analysis of GaAs in the literature [Jordan et al., 1980, 1984; Kobayahsi and Iwaki, 1985; and Szabo, 1985) are based on characterization of the growth furnace with a uniform Biot number and/or uniform temperature. Whereas such studies do provide qualitatively correct indications of the extent of the plastic deformation driving forces, protraction of such results to the calculation of thermal conditions necessary for inhibition of slip in the growing crystal requires more accurate thermal characterization of the growth furnace. For example, Jordan at al. (1984) have calculated a uniform environmental temperature distribution which would result in inhibition of plastic deformation of the growing GaAs crystal. These results are based on models of heat tranfer mechanisms from the crystal to its environment (a uniform environmental temperature, thermally transparent encapasulant, and grossly inaccurate natural convection heat transfer from the crystal to the encapsulant) which have been shown to be incorret (Motakef and Witt, 1987; Motakef, 1987a; Ostogorsky, 1987). Calculation of an "environmental temperature" based on the present results of maximum allowable heat transfer rates requires quantification of the thermal coupling between the crystal and the so-called "environment" which, considering the complex geometry of the growth system and the transient nature of the growth process, cannot be readily identified.

The present results are based on an averaging approximation of the resolved shear stresses in the growing solid which overestimates the maximum allowable heat transfer rates at the crystal periphery. An alternative approach to calculation of $q_s^*(z)$ and $q_s^{**}(z)$ is to restrict the maximum resolved shear stress at each location to be smaller than the plasticity limit. Extension of the present study to an analysis based on this criterion is in progress.

This analysis is based on the currently accepted, but as of yet unquantified, relationship between excess stress, onset of plastic deformation, and dislocation generation in GaAs. Although the methodology developed in this paper is not restricted to glide in the slip system and can be extended to other plastic deformation mechanisms, the calculated values of maximum allowable heat fluxes at the crystal periphery are based on the magnitudes of CRSS and other thermophysical properties cited in the literature. Therefore, further investigations into the mechanisms of dislocation generation in GaAs is essential for enhanced interaction between experimental and modelling efforts aimed at improvement of the LEC growth process.

Acknowledgments

The author wishes to express appreciation to the support provided by DARPA and AFWAL/MLPO (contract No., F33615-83-C-5089) for their support of this work.

References

Boley, B. A., and Weiner, J. H., 1960, *Theory of Thermal Stresses*, Wiley, p. 251.

Jordan, A. R., 1980, "An Evaluation of the Thermal and Elastic Constants Affecting GaAs Crystal Growth," *J. Cryst. Growth*, Vol. 49, pp. 631–642.

Jordan, A. S., VonNeida, A. R., Caruso, R., 1980, "A Thermoelastic Analysis of Dislocation Generation in Pulled GaAs Crystals," *Bell Sys. Tech. J.*, Vol. 59, pp. 593–677.

Jordan, A. S., VonNeida, A. R., Caruso, R., 1984, "The Theory and Practice of Dislocation Reduction in GaAs and InP," *J. Cryst. Growth*, Vol. 70, pp. 555–572.

Kobayashi, M., and Iwaki, T., 1985, "A Thermoelastic Analysis of the Thermal Stress Produced in Semi-Infinite Cylindrical Single-Crystal During the Czochralski Growth," *J. Cryst. Growth*, Vol. 73, pp. 96–110.

Motakef, S., and Witt, A. F., 1986, "Thermoelastic Analysis of GaAs in LEC Growth Configuration, Part I: Effect of Liquid Encapsultant on Thermal Streses," *J. Cryst. Growth*, Vol. 80, pp. 37–50.

Motakef, S., 1987a, "Effect of Natural Convection and Thermal Transparency of Liquid Encapsulant on Thermal Stresses During LEC Growth of GaAs," *Int. J. of Heat and Mass Trans.*, in press.

Motakef, S., 1987b, "Thermoelastic Analysis of GaAs in LEC Growth, Configuration, Part II: Temporal Evolution of the Stress Field," submitted to *J. Cryst. Growth*.

Ostrogorsky, A. G., Yao, K. H., and Witt, A. F., 1987, "Infra-Red Absorbance of B_2O_3 at Temperatures to 1250°C," *J. Cyrst. Growth*, in press.

Seki, Y., Watanabe, H., and Matsui, J., 1978, "Impurity Effect on Grown-In Dislocation Density in InP and GaAs Crystals," *J. Appl. Phys.*, Vol. 49, pp. 822–828.

Szabo, G., 1985, "Thermal Strain During Czochralski Growth,' *J. Cryst. Growth*, Vol. 73, pp. 131–141.

Uemura, C., et al., 1981, "LEC Growth and Characterization of Undoped InP Crystals,' *J. Cryst. Growth*, Vol. 52, pp. 591–596.

A Micromechanical Theory of High Temperature Creep

G. J. Weng
Professor,
Department of Mechanics
and Materials Science,
Rutgers University,
New Brunswick, NJ 08903
Mem. ASME

Based on the mechanism of dislocation climb-plus-glide, a micromechanical theory is developed for the high-temperature creep of polycrystals. This model assumes that dislocation climb is responsible for the release of dislocations and whose subsequent glide provides the only significant contribution to the overall creep strain. Taking into consideration the forces acting on both dislocation climb and dislocation glide, a microconstitutive equation is introduced to describe the transient and steady-state creep of slip systems. Together with the self-consistent relation, the creep property of a polycrystal is determined by an averaging process over the behavior of its constituent grains. The developed micromechanical theory is then applied to model the creep behavior of lead at 0.56 T_m, under both tension and shear. Based on these micromechanical analyses, a macroscopic multiaxial theory—involving an effective normal stress to reflect the climb force on the microscale as well as the usual effective shear stress—is also developed. It is found that the effective normal stress, which is independent of the hydrostatic pressure, depends primarily on the second invariant of the deviatoric stress, and only weakly so on the third invariant. Thus despite the distinct presence of two types of microstress, the constitutive equations on the macroscale can still be reasonably described by the second invariant alone even at high temperature.

1 Introduction

In their first and subsequent reports on the deformation-mechanism maps, Ashby (1972) and Frost and Ashby (1982) have clearly demonstrated that there are several inelastic deformation mechanisms for a polycrystalline metal, and that, depending on the stress and temperature level, one mechanism is usually more dominant than the other. Over the entire stress and temperature range, three principal fields have appeared in the map. In the low-temperature, high-stress range, dislocation glide is the most important one, leading to the low-temperature plasticity. But as temperature increases, the ability of dislocations to climb as well as to glide introduces the so-called power-law creep. Also at the high-temperature, but low-stress range, diffusional flow is more dominant. Although several other mechanisms are also potentially active, power-law creep—as was first systematically demonstrated by Dorn (1954)—appears to be the primary source of creep when both temperature and the applied stress are sufficiently high.

On the microcontinuum scale, dislocation glide has often been modelled as crystallographic slip and, by which, the plasticity of single crystals and polycrystalline metals has been fruitfully studied. While the inelastic behavior of metals is more deeply rooted in the behavior of dislocations and other microstructural activities, the behavior of a slip system often serves as a convenient point of departure for the study of micro-macro transition. This scale is small enough that the collected behavior of dislocations can still be reasonably represented, but large enough to allow one to characterize its local stress and strain state under an externally applied stress. For the dislocation glide, the constitutive equations then involve principally the resolved shear stress, shear strain, and strain rate. The constitutive equations of the slip systems involving the climb-plus-glide process, on the other hand, would involve the resolved normal stress for the dislocation climb, and resolved shear stress for the released dislocations to glide, in addition to the usual strain and strain-rate. The latter equation appears not to have been introduced in the study of inelastic behavior of polycrystals from those of its constituent grains, and this we shall attempt to do.

As in Frost and Ashby (1982), dislocation glide and power-law creep will be treated as alternative mechanisms and, within the intended stress and temperature range, creep strain will be calculated solely from the power-law creep. While power-law creep and point diffusional flow are independent mechanisms involving different defects, dislocation glide and power-law creep basically involve the same dislocations moving under different conditions, the fast one being dominating. At high stress, the strain contributed by diffusional flow is relatively small; especially with the transient creep to be included, its fractional contribution is really negligible. The calculation of creep strain based on the dislocation climb-plus-glide process, therefore, should be a reasonable approach at high temperature.

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, 1986; final revision June 11, 1987.

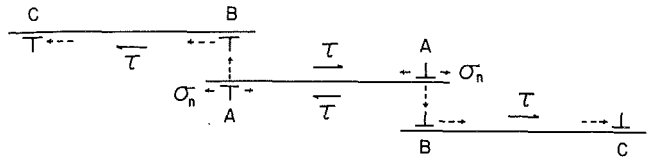


Fig. 1 A schematic representation of dislocation climb-plus-glide

Taking into account both active and latent hardening of slip systems, we shall incorporate the dislocation climb-plus-glide process to establish the constitutive equations of slip systems. This would allow one to calculate the creep strain of a constituent grain and, with the help of a self-consistent scheme, to determine the transient and steady-state power-law creep of a polycrystal by an averaging process.

While our principal goal is to develop a micromechanical theory at a constant temperature, we shall at the end use its basic features to construct a macroscopic multiaxial one suitable for engineering design and analysis.

2 Constitutive Equations of Slip Systems

The earliest theory of power-law creep was probably that developed for the steady state by Weertman (1955), whose basic premise that climb produces the release of dislocation loops whose glide provides the only significant contribution to the overall creep strain will be adopted in our theory. A schematic model of the dislocation climb-plus-glide process is depicted in Fig. 1. The motion of dislocations first involves the release from the locked position at A by dislocation climb to position at B, and then the glide from B to another locked position at C.

The creep strain is generated by the glide process from B to C. Since dislocations dwell mostly in the locked position the creep rate during the subsequent glide from B to C must first depend on the release rate of dislocations from A to B, and then the area (or distance) they sweep out in the subsequent glide. In this fashion the shear creep rate of a slip system may be written as (Evans and Knowles, 1977)

$$\dot{\gamma}^c = \dot{N}\phi b, \tag{2.1}$$

where $\dot{N}$ is the rate of dislocations released to position B—the gliding position—by dislocation climb, ϕ is the average area swept out by the released dislocations, and b the usual Burgers vector.

To establish the stress dependence of $\dot{\gamma}^c$ from this equation, it may be recognized that the rate of release $\dot{N}$ depends on both the velocity of dislocation climb and the dislocation density on the slip plane. The climb velocity is controlled by the climb force, whereas the dislocation density on a glide plane and the slipped area ϕ generally depend on the resolved shear stress and creep strains. We therefore expect a general dependence of $\dot{\gamma}^c$ on the climb force, resolved shear stress, and slip strains.

For instance, based on a model of dislocation network, Evans and Knowles (1977) found that the rate of release by lattice diffusion can be written as

$$\dot{N} \sim \frac{v_\ell}{\bar{\ell}} \cdot \frac{1}{\bar{x}^3}, \tag{2.2}$$

where v_ℓ is the climb velocity of dislocation link, $\bar{\ell}$ is the average link length and $\bar{x}$ the average mesh size. The area swept out ϕ, is approximately equal to $\bar{\ell} \cdot s$, or the product of link length and slip distance. In a three-dimensional network both $\bar{\ell}$ and s are proportional to $\bar{x}$; thus

$$\phi \sim \bar{x}^2, \text{ and } \dot{N} \sim v_\ell / \bar{x}^4. \tag{2.3}$$

The mesh-size in a crystal is in general inversely proportional to the applied stress ($\bar{x} \sim 1/\sigma$, see also Evans and Knowles, 1977). While it is appreciated that the formation of a three-dimensional network is primarily a result of multislip and that the applied stress here cannot be simply identified with τ or σ_n of a particular slip system alone, such an identification would become more justifiable if it is carried out for all systems. Then the activity of multislip leading to a network structure can be accounted for. Furthermore, recognizing that slip activity on the glide plane or the slip distance is indeed more dependent on τ, than on σ_n, we may set $\bar{x} \sim 1/\tau$ for every system as an approximation. Since the velocity of the diffusive, stress-assisted dislocation climb is proportional to the climb force, which includes both the chemical force due to the nonequilibrium vacancy concentration and the mechanical one contributed by the resolved normal stress σ_n, one has

$$v_\ell \sim \psi + \sigma_n, \tag{2.4}$$

where the chemical force—represented ψb—is equal to $(kTb/a_v)\ell n(c_v/c_o)$, with k, T, a_v, c_v, and c_o being the Boltzmann constant, absolute temperature, atomic volume, actual vacancy, and equilibrium vacancy concentrations, in turn (Weertman, 1965; Hirth and Lothe, 1982). The stress dependence of creep rate can therefore be cast as

$$\dot{\gamma}^c \sim (\psi + \sigma_n) \cdot \tau^n, \tag{2.5}$$

with $n = 2$. Experimental data show that the value of n may lie from 2 to 6.

As dislocation climb and the subsequent glide are interdependent and the climb process is the slower of the two, dislocation climb is rate-controlling in this process.

With this stress dependence we may write the constitutive equation of slip systems for a power-law creep. The dislocation climb-plus-glide is capable of producing both transient and steady creep, and may also cause active and latent hardening. At a constant temperature the steady creep rate of, say system k, may take the form

$$\dot{\gamma}_s^{(k)_c} = \kappa \left(\psi + \sigma_n^{(k)}\right) \tau^{(k)_n}. \tag{2.6}$$

Its transient creep rate, on the other hand, decreases with increasing active and latent hardening; assuming that there is a combined isotropic and kinematic hardening for the slip systems, it may be written as

$$\dot{\gamma}_t^{(k)_c} = \frac{\eta \left(\psi + \sigma_n^{(k)}\right) \tau^{(k)_n}}{\zeta + \sum_\ell \left[\alpha + (1-\alpha)\cos\theta^{(k,\ell)} \cos\phi^{(k,\ell)}\right] \gamma^{(\ell)_c}}, \tag{2.7}$$

where, like κ in equation (2.6), η, ζ, and α are material constants, $\theta^{(k,\ell)}$ and $\phi^{(k,\ell)}$ are the angles between the slip directions and the slip-plane normals, respectively, of the kth and ℓth slip systems, $\gamma^{(\ell)_c}$ is the creep strain of the ℓth system, and the sum extends to all systems in the same grain. We further assume that ψ stays appreciably constant and may also be treated as a material constant. Parameter α is "the degree of isotropy in work hardening;" when $\alpha = 1$, this corresponds to Taylor's (1938) isotropic hardening and when $\alpha = 0$, it reduces to Prager's kinematic hardening (Weng, 1979). It also represents the internal back-stress under cyclic deformation and has proved essential to the study of cyclic creep (or more precisely plasticity) acceleration and deceleration at low temperature (Weng, 1983).

If the local stress field in the considered grain is denoted by σ_{ij}, the resolved shear stress is simply given by

$$\tau^{(k)} = \nu_{ij}^{(k)} \sigma_{ij}, \tag{2.8}$$

where $\overset{(k)}{\nu}_{ij}$ is the Schmid-factor tensor of the kth slip system, given by

$$\nu_{ij} = \frac{1}{2}(b_i n_j + b_j n_i), \tag{2.9}$$

in terms of its unit slip direction b_i and slip-plane normal n_i. In equation (2.8), and henceforth, the Einstein summation convention for a repeated index is adopted.

While the gliding process involves no diffusion, the climb process is associated with the diffusion of atoms or vacancies. Taking into account such a process, Weertman (1965) found that the force acting on a dislocation climb can be calculated from the Peach-Koehler (1950) equation, but with the stress tensor replaced by its deviatoric component. The resolved normal stress then is given by

$$\overset{(k)}{\mu}_{ij} = \overset{(k)}{\mu}_{ij}\sigma'_{ij}, \tag{2.10}$$

where σ'_{ij} is the deviatoric component of σ_{ij} ($\sigma'_{ij} = \sigma_{ij} - 1/3\ \delta_{ij}\sigma_{kk}$, δ_{ij} being the Kronecker delta), and μ_{ij} is given by

$$\mu_{ij} = b_i b_j. \tag{2.11}$$

Since both positive and negative $\psi + \sigma_n$ may lead to dislocation climb, its absolute value should be used in equations (2.6) and (2.7); the sign of creep rate $\dot{\gamma}^c$ is determined soley by τ.

Once the creep rate is calculated from the constitutive equations (2.6) and (2.7), the creep-rate tensor of the considered grain follows from all contributions of its slip systems

$$\dot{\epsilon}^c_{ij} = \sum_k \overset{(k)}{\nu}_{ij}\overset{(k)}{\dot{\gamma}}_c, \tag{2.12}$$

where $\dot{\gamma}^c = \dot{\gamma}^c_s + \dot{\gamma}^c_t$, and the summation extends to all slip systems in the grain.

In passing we note that, if the creep rate of a slip system associated with the dislocation climb from A to B is denoted by $\dot{\epsilon}^c_n$ (as opposed to $\dot{\gamma}^c$ for the shear rate), the produced creep-rate tensor of the grain would be

$$\dot{\epsilon}^c_{ij} = \sum_k \left(\overset{(k)}{\mu}_{ij} - \frac{1}{3}\delta_{ij}\right)\overset{(k)}{\dot{\epsilon}}_{n_c}, \tag{2.13}$$

assuming the kinetics of vacancy equilibrium to be fast enough so that the vacancy concentration stayed appreciably constant during dislocation climb, leading to no net volume change. Such a process was considered by Nabarro (1967), who also arrived at a power-law creep by assuming dislocation climb to be the sole source of creep strain. It is interesting to see that the creep-strain tensors generated by the glide and the climb processes of a given slip system are always mutually orthogonal, as

$$\nu_{ij}\left(\mu_{ij} - \frac{1}{3}\delta_{ij}\right) = 0, \tag{2.14}$$

for a given system. Since in reality the climb distance from A to B is very short as compared to the glide distance from B to C, the creep rate given by equation (2.13) is practically negligible.

3 The Local Stress in a Grain and the Averaging Process

Primarily due to the difference in grain orientations, the deformation of a polycrystal is highly heterogeneous. Such a heterogeneous deformation necessarily leads to a stress redistribution. As a consequence the local stress σ_{ij} needed to calculate the creep rates in equations (2.6) and (2.7) generally differs from the externally applied one $\bar{\sigma}_{ij}$; their difference largely depends on the extent of creep deformation in the considered grain and the aggregate. Such a grain-to-grain variation perhaps can be most conveniently evaluated by the self-consistent relation.

Recognizing that creep deformation can take place under a constant stress, Weng (1981) identified that the incremental creep strain is truely a "stress-free" strain in the sense of Eshelby (1957), and found that the Kröner (1961) and Budiansky-Wu (1962) relation can be used to evaluate the incremental stress variation of a constituent grain. Over a time increment dt, if the incremental creep strain of a grain is denoted by $d\epsilon^c_{ij}$, and that of the aggregate by the corresponding barred (averaging) quantity $d\bar{\epsilon}^c_{ij}$, the incremental stress variation of a grain can be written as

$$d\sigma_{ij} = d\bar{\sigma}_{ij} - 2\mu(1-\beta)(d\epsilon^c_{ij} - d\bar{\epsilon}^c_{ij}), \tag{3.1}$$

where—assuming the grain to be elastically isotropic—μ is the shear modulus and, in terms of Poisson's ratio ν, $\beta = 2(4-5\nu)/15(1-\nu)$. This equation was first used by Brown (1970) to compute the time-dependent deformation of polycrystals.

The incremental creep strain of a grain is given by

$$d\epsilon^c_{ij} = \dot{\epsilon}^c_{ij}dt, \tag{3.2}$$

where the creep rate in turn is contributed by the creep rates of all slip systems through equation (2.12). Once $d\epsilon^c_{ij}$ is determined for all grain orientations, the incremental creep strain of the polycrystalline aggregate can be calculated from their orientational average. Symbolically we write

$$d\bar{\epsilon}^c_{ij} = \left\{d\epsilon^c_{ij}\right\}, \tag{3.3}$$

for such an average. It may be noted that $d\bar{\sigma}_{ij} = \{d\sigma_{ij}\}$ results automatically from equation (3.1).

These equations, together with the constitutive equations (2.6) and (2.7) allow one to determine the development of creep strain of a polycrystalline metal. Briefly the computational procedure goes as follows. Under an applied $\bar{\sigma}_{ij}$, the initial stress σ_{ij} for all grains are equal to the applied one, and this allows one to determine τ and σ_n from equations (2.8) and (2.10) for all slip-systems in each constituent grain. These values are substituted into the constitutive equations (2.6) and (2.7) to compute $\dot{\gamma}^c$ and then by way of equations (2.12), (3.2), and (3.3), the incremental creep strains $d\epsilon^c_{ij}$ and $d\bar{\epsilon}^c_{ij}$ can be determined. Substitution of these incremental strains into the self-consistent relation (3.1) gives rise to the stress variation $d\sigma_{ij}$ during dt ($d\bar{\sigma}_{ij} = 0$ if $\bar{\sigma}_{ij}$ is kept constant). Then using the new stress $\sigma_{ij} + d\sigma_{ij}$ for each grain as the basis, the incremental strains $d\epsilon^c_{ij}$ and $d\bar{\epsilon}^c_{ij}$ may be similarly determined for the next time-increment. This process may be repeated to obtain the creep curve under $\bar{\sigma}_{ij}$ for the entire history.

It is perhaps worth mentioning that other methods to compute the local stress in the constituent grains also exist. For instance, without considering dislocation climb Hutchinson (1976) has applied Hill's (1965) self-consistent relation to determine the steady creep rate of the aggregate. Similarly, based on Lin's (1971) equivalent body force method, Lin et al. (1977) have also derived the polycrystal property from the creep data of single crystals.

4 Comparison With Experiments

It is now of interest to apply the developed theory to a specific system and to see to what extent the major features of the prediction agree with experimental observations. Such a comparison requires that the experiments be carried out at high temperature, preferably above 0.5 T_m, and that the data include both steady and transient creep, and lie within the small strain range. The latter requirement will ensure no significant grain rotation, a factor not considered in the present formulation. In order to construct a macroscopic

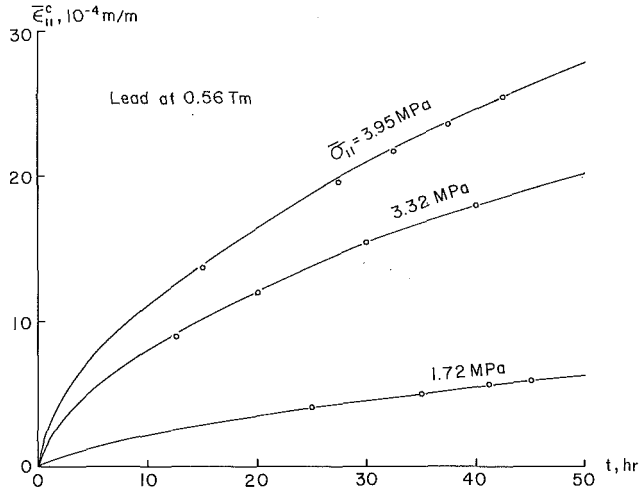


Fig. 2 Micromechanical simulations of tensile creep data (the open circles)

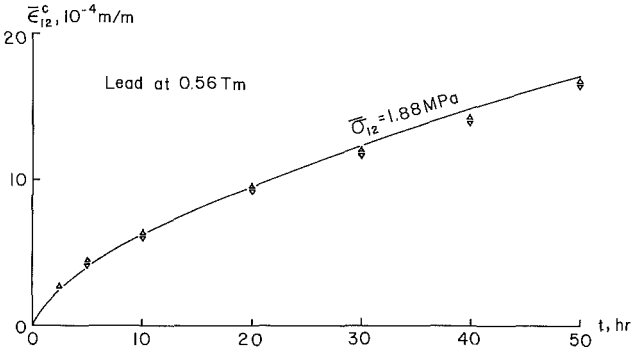


Fig. 3 Independent micromechanical prediction of shear creep, and comparison with two sets of experimental data (the triangles)

multiaxial theory, both tensile and shear results are needed. Within these conditions the writer was able to locate only one set of data, reported by Finnie (1963) on lead at 0.56 T_m (60°C).

To apply the developed micromechanical theory, first we note that lead has a face-centered-cubic crystal structure. It has four {111} slip-planes and three <110> slip-directions on each plane, resulting in twelve slip systems in each grain. The three-dimensional polycrystal model was selected to consist of 216 different grain orientations; these were obtained by rotating a basic crystal, aligned along the material axes, first about the x_1 axis then about the x_3 axis and finally about the x_2 axis. Due to the cubic symmetry, the rotation about each axis was taken to start from 0 deg and end at 75 deg, with an increment of 15 deg, leading to the required $6 \times 6 \times 6 = 216$ orientations. As judged from the generated creep-strain components, this polycrystal model was found to possess a reasonable three-dimensional isotropy. The best pair of axes for isotropy is the x_1–x_2 pair, in which the maximum deviation from the isotropic property is less than 4 percent. Consequently this plane was chosen in the numerical calculation for both tensile and shear loadings.

The elastic moduli of lead at 60°C are $\mu = 5.09$ GPa and $\nu = 0.40$. The shear modulus was calculated using the relation (Ashby, 1972)

$$-\frac{1}{\mu}\frac{d\mu}{dT} = 14.4 \times 10^{-4}/\text{K},$$

with $\mu = 5.34$ GPa at room temperature $T = 300$ K. The experimental data under three different tensile stresses and one shear stress were reproduced in Figs. 2 and 3, respectively.

There are six material constants involved—κ, η, ζ, α, ψ, and n—and these are determined from the tensile simulation of experimental data under $\bar{\sigma}_{11} = 1.72$ MPa, 3.32 MPa, and 3.95 MPa. First we note that the ability for dislocations to climb at high temperature tends to reduce the internal back-stress; together with the commonly observed network structure it appears reasonable to assume the isotropic hardening ($\alpha = 1$) in the calculation. Then, noting that primarily κ would control the steady creep rate, $\kappa + \eta/\zeta$ the initial creep rate, η/ζ^2 the initial decreasing rate of the creep rate, n the separation of creep curves at different stresses, and ψ would add the relative creep activity at lower stress, these constants were determined so that the calculated creep curves under the three tensile stresses could well simulate the experimental data. The results are: $\kappa = 1.01 \times 10^{-1}$, $\eta = 14.7$, $\zeta = 18$, $\psi = 0.2$, and $n = 2$, where stress, strain, and time are in the units of MPa, 10^{-4} m/m and hr, in turn. The time increments taken in the computations are $\Delta t = 0.1$ if $t < 2$, $\Delta t = 0.2$ if $2 \le t < 5$, $\Delta t = 0.3$ if $5 \le t < 8$, and $\Delta t = 0.5$ if $t \ge 8$. With these, we *independently* calculated the shear creep strain under $\bar{\sigma}_{12} = 1.88$ MPa. These calculated creep results were plotted along the experimental data in Figs. 2 and 3. The tensile simulations and the independent shear prediction appear to be in line with the test data.

5 A Macroscopic Multiaxial Theory

It appears that no macroscopic creep theory has ever been developed to account for the microscopic dislocation climb. Based on the properties of the micromechanical model, our objective here is to develop such a simple theory under the combined stress.

In view of the constitutive equations (2.6) and (2.7), three basic features are associated with the micromechanical theory. First, as in low-temperature plasticity, creep strains are generated solely by dislocation glide, which is governed by τ. Second, the deformation is insensitive to the hydrostatic pressure—or the first invariant of stress tensor—and involves no volume change. Finally the creep rate depends not only on the resolved shear stress and creep strain, but also on the chemical force and the resolved normal stress σ_n.

The resolved shear stress τ has been successfully modelled by the second invariant of the deviatoric stress tensor on the macro scale. Denoting this macroscopic stress by τ^*, it is given by

$$\tau^* = \sqrt{3J_2}, \tag{5.1}$$

where J_2 is the second invariant of the deviatoric stress tensor, given by $J_2 = 1/2\ \bar{\sigma}'_{ij}\bar{\sigma}'_{ij}$.

Since the resolved normal stress involves no hydrostatic pressure, its macroscopic counterpart—within the framework of the isotropic theory—would be a function of J_2 and the third invariant J_3 only. Allowing for the possibility that either J_2 or J_3 may vanish, we set this stress as a simple linear combination of the two as

$$\sigma_n^* = \frac{1}{1/\sqrt{3} + p \cdot \sqrt[3]{2}/3}(J_2^{1/2} + pJ_3^{1/3}), \tag{5.2}$$

where p is a fractional ratio and the coefficient in front of the parenthesis is introduced to normalize the combination so that $\sigma_n^* = \sigma$ under pure tension. (This latter condition is set for convenience; any other choice—say a fraction of σ—would still lead to the same constitutive form in equation (5.3), with that fraction absorbed by the parameters a, b, and c.) In terms of the three principal stresses σ_1, σ_2, and σ_3 the third deviatoric invariant is given by $J_3 = 1/27\ [(2\sigma_1 - \sigma_2 - \sigma_3)(2\sigma_2 - \sigma_3 - \sigma_1)(2\sigma_3 - \sigma_1 - \sigma_2)]$. As J_3 can sometimes become negative, only its absolute value—like σ_n—should be used (alternatively one may write it as an even function $(J_3^2)^{1/6}$). The fractional ratio p can be determined—as will be

done shortly—by comparing the tensile and shear creep curves.

In parallel to equations (2.6) and (2.7), we may introduce the macroscopic constitutive equation for the effective creep rate $\dot{\epsilon}^{*c}$ as

$$\dot{\epsilon}^{*c} = \left(a + \frac{b}{d+\epsilon^{*c}}\right) \cdot (c+\sigma_n^*)\tau^{*m}, \tag{5.3}$$

where a, b, d, and c are macroscopic constants, with c representing the chemical force, and the effective creep strain is given by

$$\epsilon^{*c} = (\sqrt{2}/3)\left[\left(\bar{\epsilon}_{11}^c - \bar{\epsilon}_{22}^c\right)^2 + \left(\bar{\epsilon}_{22}^c - \bar{\epsilon}_{33}^c\right)^2 + \left(\bar{\epsilon}_{33}^c - \bar{\epsilon}_{11}^c\right)^2 + 6\left(\bar{\epsilon}_{12}^c + \bar{\epsilon}_{23}^c + \bar{\epsilon}_{31}^c\right)\right]^{1/2}. \tag{5.4}$$

The creep strain-rate components follow from

$$\dot{\bar{\epsilon}}_{ij}^c = \frac{3}{2}\frac{\dot{\epsilon}^{*c}}{\tau^*} \cdot \bar{\sigma}_{ij}', \tag{5.5}$$

along the direction of the deviatoric stress.

The suggested macroscopic theory falls within the framework of normality structure (see for instance Rice, 1970). Assuming Φ as a creep potential, we may recast these constitutive relations into

$$\dot{\bar{\epsilon}}_{ij}^c = \lambda(J_2, J_3)\frac{\partial \Phi}{\partial \bar{\sigma}_{ij}'}, \tag{5.6}$$

with

$$\Phi = \frac{1}{m+1}\left(a + \frac{b}{d+\epsilon^{*c}}\right)\left(\sqrt{3J_2}\right)^{m+1}, \tag{5.7}$$

$$\lambda = c + \frac{1}{1/\sqrt{3} + p \cdot \sqrt[3]{2}/3}\left(J_2^{1/2} + pJ_3^{1/3}\right). \tag{5.8}$$

5.1 Determination of the Fractional Parameter p. Depending on the objective, the parameter p may be determined by applying the macroscopic theory to reproduce the creep curves of either the experimental data or the micromechanical calculations. This can be conveniently accomplished by considering two creep curves: one under pure tension and the other under pure shear (whose $J_3 = 0$).

Under pure tension one has $\sigma_n^* = \tau^* = \bar{\sigma}_{11}$, $\dot{\epsilon}^{*c} = \dot{\bar{\epsilon}}_{11}^c$, and $\epsilon^{*c} = \bar{\epsilon}_{11}^c$; under pure shear one finds $\sigma_n^* = \bar{\sigma}_{12}/(1/\sqrt{3}) + p \cdot \sqrt[3]{2}/3$), $\tau^* = \sqrt{3}\,\bar{\sigma}_{12}$, $\dot{\epsilon}^{*c} = 2/\sqrt{3}\,\dot{\bar{\epsilon}}_{12}^c$, and $\epsilon^{*c} = 2/\sqrt{3}\,\bar{\epsilon}_{12}^c$. In both cases the effective creep strain ϵ^{*c} can be obtained as a function of time **t** by integrating equation (5.3) as

$$\epsilon^{*c} - \frac{b}{a}\ln\frac{\epsilon^{*c}+d+b/a}{d+b/a} = a(c+\sigma_n^*)\tau^{*m}t. \tag{5.9}$$

When translated into the tensile and shear components under the respective loadings, this equation leads to

$$\bar{\epsilon}_{11}^c - \frac{b}{a}\ln\frac{\bar{\epsilon}_{11}^c + d + b/a}{d+b/a} = a \cdot (c+\sigma_{11}) \cdot \bar{\sigma}_{11}^m t, \tag{5.10}$$

$$\frac{2}{\sqrt{3}}\bar{\epsilon}_{12}^c - \frac{b}{a}\ln\frac{2/\sqrt{3}\cdot\bar{\epsilon}_{12}^c + d + b/a}{d+b/a} = a \cdot \left(c + \frac{\sqrt{3}\,\bar{\sigma}_{12}}{1 + p\cdot\sqrt[3]{2}/\sqrt{3}}\right) \cdot (\sqrt{3}\,\bar{\sigma}_{12})^m t. \tag{5.11}$$

The five macroscopic parameters—a, b, d, c, and m—can be determined readily from three tensile creep curves using equation (5.10). For lead, the results are: $a = 2.62 \times 10^{-3}$, $b = 9.44 \times 10^{-2}$, $d = 1.5$, $c = 0.85$, and $m = 2.1$, where stress, strain, and time are in the units of MPa, 10^{-4} m/m and hr, in turn. With these parameters, the shear creep strain at a given time **t** from either the experiment or the micromechanical theory can be used in equation (5.11) to find the value of p. We used the shear strain at $t = 50$ hr. (Fig. 3) to compute this value, and found

$$p = \begin{cases} 0.04 & \text{from experiment,} \\ -0.02 & \text{from the micromechanical theory.} \end{cases} \tag{5.12}$$

Both results suggest that the fractional contribution of J_3 to σ_n^* is relatively small and that, as a first approximation, one may set

$$p = 0, \quad \sigma_n^* = \sqrt{3\,J_2} = \tau^*. \tag{5.13}$$

Remarkably, while two distinct microstresses—σ_n and τ—appear in the microscopic constitutive equations, only the second invariant of the deviatoric stress appears in the macroscopic ones.

6 Concluding Remarks

We have, therefore, developed a micromechanical theory for the high temperature creep of metals. This theory takes dislocation climb to be the rate-controlling mechanism and dislocation glide to be responsible for the generation of creep strains. A micro constitutive equation, involving the chemical force, the resolved normal and shear stresses, and active and latent hardening, is developed for the slip systems. The ensuing creep behavior are that, like the low-temperature plasticity, it is independent of hydrostatic stress and involves no volume change, but that, dislocation climb, not dislocation glide, is rate-controlling and that both the chemical force and the resolved normal stress are also present in the rate-equation. This theory is applied to model the creep behavior of lead at 0.56 T_m under both tension and shear, and the results are found to be satisfactory.

Based on the basic features of the microchemical model, a simple macroscopic multiaxial theory has also been developed to account for the effect of dislocation climb on the macro scale. The effective normal stress—reflecting the mechanical force responsible for the dislocation climb—is found to depend primarily on the second invariant of the deviatoric stress tensor and only weakly so on the third invariant. The constitutive equations at high temperature, therefore, can be described by the chemical force, creep strain, and the second invariant of the deviatoric stress.

Acknowledgment

This work was supported by the US National Science Foundation, Solid Mechanics Program, under Grant MEA-8317887, and Solid and Geo-Mechanics Program, under Grant MSM 86-14151.

References

Ashby, M. F., 1972, "A First Report on Deformation-Mechanism Maps," *Acta Metallurgica*, Vol. 20, pp. 887–897.

Brown, G. M., 1970, "A Self-Consistent Polycrystalline Model for Creep under Combined Stress States," *Journal of the Mechanics and Physics of Solids*, Vol. 18, pp. 367–381.

Budiansky, B., and Wu, T. T., 1962, "Theoretical Prediction of Plastic Strains of Polycrystals," *Proceedings 4th U.S. National Congress of Applied Mechanics*, pp. 1175–1185.

Dorn, J. E., 1954, "Some Fundamental Experiments on High Temperature Creep," *Journal of the Mechanics and Physics of Solids*, Vol. 3, pp. 85–116.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proceedings of the Royal Society, London*, Vol. A241, pp. 376–396.

Evans, H. E., and Knowles, G., 1977, "A Model of Creep in Pure Materials," *Acta Metallurgica*, Vol. 25, pp. 963–975.

Finnie, I., 1963, "An Experimental Study of Multiaxial Creep in Tubes,"

Proceedings Joint International Conference on Creep, Institution of Mechanical Engineers, Vol. 178, pp. 221–226.

Frost, H. J., and Ashby, M. F., 1982, *Deformation-Mechanism Maps*, Pergamon Press, Oxford.

Hill, R., 1965, "Continuum Micro-Mechanics of Elastoplastic Polycrystals," *Journal of the Mechanics and Physics of Solids*, Vol. 13, pp. 89–101.

Hirth, J. P., and Lothe, J., 1982, *Theory of Dislocations*, McGraw-Hill, New York.

Hutchinson, J. W., 1976, "Bounds and Self-Consistent Estimates for Creep of Polycrystalline Materials," *Proceedings of the Royal Society, London*, Vol. A348, pp. 101–127.

Kröner, E., 1961, "Zur Plastischen Verformung des Vielkristalls," *Acta Metallurgica*, pp. 155–161.

Lin, T. H., 1971, "Physical Theory of Plasticity," *Advances in Applied Mechanics*, Yih, C. S., ed., Vol. 11, Academic Press, New York.

Lin, T. H., Yu, C. L., and Weng, G. J., 1977, "Derivation of Polycrystal Creep Properties from the Creep Data of Single Crystals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 64, pp.73–78.

Nabarro, F. R. N., 1967, "Steady-State Diffusional Creep," *Philosophical Magazine*, Vol. 16, pp. 231–237.

Peach, M., and Koehler, J. S., 1950, "The Forces Exerted on Dislocations and the Stress Fields Produced by Them," *Physical Review*, Vol. 809, pp. 436–439.

Rice, J. R., 1970, "On the Structure of Stress-Strain Relations for Time-Dependent Plastic Deformation in Metals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp.728–737.

Taylor, G. I., 1938, "Plastic Strain in Metals," *Journal of the Institute of Metals*, Vol. 62, pp. 307–324.

Weertman, J., 1955, "Theory of Steady-State Creep Based on Dislocation Climb," *Journal of Applied Physics*, Vol. 26, pp. 1213–1217.

Weertman, J., 1965, "The Peach-Koehler Equation for the Force on a Dislocation Modified for Hydrostatic Pressure," *Philosophical Magazine*, Vol. 11, pp. 1217–1223.

Weng, G. J., 1979, "Kinematic Hardening Rule in Single Crystals," *International Journal of Solids and Structures*, Vol. 15, pp. 861–870.

Weng, G. J., 1981, "Self-Consistent Determination of Time-Dependent Behavior of Metals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 41–46.

Weng, G. J., 1983, "The Influence of Fatigue Stress on the Creep Behavior of Metals," *Acta Metallurgica*, Vol. 31, pp. 207–212.

Crack Paralleling an Interface Between Dissimilar Materials

J. W. Hutchinson

M. E. Mear[1]

J. R. Rice

Division of Applied Sciences,
Harvard University,
Cambridge, MA 02138

A crack paralleling a bonded plane interface between two dissimilar isotropic elastic solids is considered. When the distance of the crack from the interface is small compared to the crack length itself and to other length scales characterizing the geometry, a simple universal relation exists between the Mode I and Mode II stress intensity factors and the complex stress intensity factor associated with the corresponding problem for the crack lying on the interface. In other words, if the influence of external loading and geometry on the interface crack is known, then this information can immediately be used to generate the stress intensity factors for the sub-interface crack. Conditions for cracks to propagate near and parallel to, but not along, an interface are derived.

1 Introduction

Bonded interfaces between dissimilar elastic materials often separate by cracking, as would be expected if the toughness of the interface is low compared to that of the abutting materials. In some instances cracking is observed to occur approximately parallel to the interface but with the crack lying entirely within one of the two materials. The aim of this paper is to analyze subinterface cracks which parallel the interface and to examine conditions under which they might be expected.

The mathematical problem which is analyzed is introduced in Fig. 1. Each material is taken to be isotropic and linearly elastic. The interface lies along the x_1 axis with material #1 lying above and #2 below. Plane strain deformations are considered. Attention will be restricted to subinterface cracks which lie below the interface at a distance h which is small compared to the length of the crack L and to all other relevant geometric length quantities in the problem. As indicated in Fig. 1, we will consider the asymptotic problem for the semi-infinite subinterface crack. The remote field in the asymptotic problem is prescribed to be the near-tip field of the interface crack problem (everywhere but in material #2 between the crack and the interface). That is, the solution to the subinterface crack problem at any point a fixed distance from the tip approaches the solution to the corresponding interface crack problem as $h \to 0^+$ with L fixed. Thus, at distances from the tip which are large compared to h and small compared to L, the near-tip field of the interface crack problem pertains. Posing the problem in this manner permits us to develop a universal relation between the Mode I and II stress intensity factors of the subinterface crack and the corresponding "complex" stress intensity factor of the interface crack. This relation is

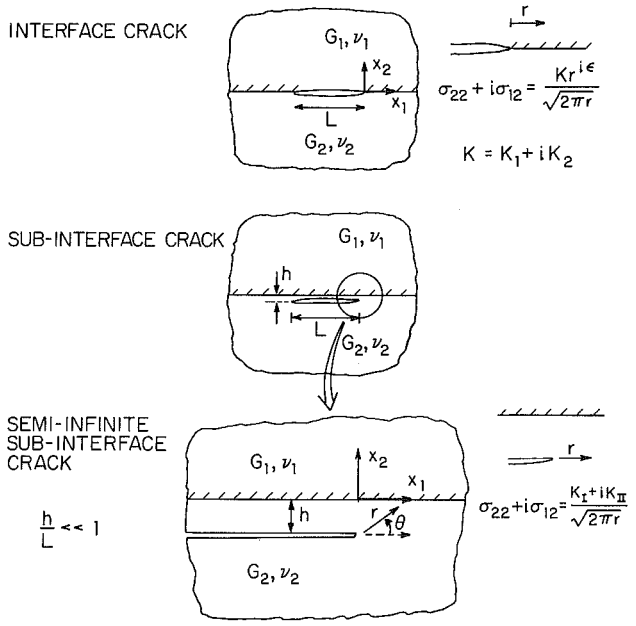


Fig. 1 Relation of asymptotic subinterface crack problem to interface crack problem

otherwise independent of loading, crack length, and external geometry.

With the universal relation in hand, we examine conditions under which propagation of a parallel sub-interface crack should be expected. When conditions do favor such cracks, the analysis predicts the separation distance from the interface.

2 Formulation and Solution

The singular near-tip field of the *interface crack problem*

[1]Currently, Department of Aerospace Engineering and Engineering Mechanics, University of Texas at Austin, Austin, Texas.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May 11, 1987.

Table 1 Values of $\phi(\alpha, \beta)$

β \ α	−.8	−.6	−.4	−.2	−.1	0	.1	.2	.4	.6	.8
−.4	−.1202	−.0801	−.0467	−.0186	−.0060	.0058	.0168	.0273	.0468	.0653	.0810
−.2	−.1814	−.1162	−.0708	−.0351	−.0197	−.0056	.0075	.0197	.0419	.0618	.0798
−.1	−.2057	−.1281	−.0764	−.0368	−.0199	−.0046	.0096	.0227	.0465	.0675	.0865
−.05	−.2167	−.1328	−.0779	−.0363	−.0187	−.0027	.0120	.0256	.0501	.0718	.0912
−.02	−.2229	−.1354	−.0785	−.0356	−.0176	−.0012	.0139	.0277	.0527	.0748	.0946
0	−.2270	−.1369	−.0787	−.0350	−.0167	0	.0153	.0293	.0547	.0770	.0970
.02	−.2309	−.1384	−.0788	−.0343	−.0156	.0013	.0168	.0311	.0567	.0793	.0995
.05	−.2366	−.1403	−.0787	−.0330	−.0138	.0035	.0193	.0339	.0601	.0830	.1035
.1	−.2456	−.1431	−.0780	−.0301	−.0101	.0079	.0243	.0393	.0663	.0900	.1110
.2	−.2620	−.1468	−.0744	−.0219	−.0003	.0191	.0367	.0528	.0815	.1065	.1287
.4	−.2902	−.1449	−.0566	.0055	.0307	.0531	.0733	.0917	.1242	.1522	.1769

(England, 1965; Erdogan, 1965; Rice and Sih, 1965) gives rise to tractions directly ahead of the tip ($\theta = 0$) given by

$$\sigma_{22} + i\sigma_{12} = K(2\pi r)^{-1/2} r^{i\epsilon} \tag{2.1}$$

where $K = K_1 + iK_2$ is the complex stress intensity factor, $i = \sqrt{(-1)}$, and

$$\epsilon = \frac{1}{2\pi} ln\left[\frac{G_1 + G_2(3-4\nu_1)}{G_2 + G_1(3-4\nu_2)}\right] \tag{2.2}$$

where G is the shear modulus and ν is Poisson's ratio. Here $K = (k_1 + ik_2)\sqrt{\pi}\cosh \pi\epsilon$ where $k_1 + ik_2$ is the complex intensity factor as originally introduced by Rice and Sih (1965). The $\sqrt{\pi}$ is standard in converting the lower case k's of that period to K's; we include the factor $\cosh \pi\epsilon$ so that the magnitude of the traction vector on the interface is given by $\sqrt{(\sigma_{22}^2 + \sigma_{12}^2)} = |K|/\sqrt{2\pi r}$, analogously to the homogeneous material case. The associated crack face displacements a distance r behind the tip are given by

$$\delta_2 + i\delta_1 = 2\,\frac{[(1-\nu_1)/G_1 + (1-\nu_2)/G_2]}{(1+2i\epsilon)\cosh \pi\epsilon} K(r/2\pi)^{1/2} r^{i\epsilon} \tag{2.3}$$

where $\delta_\alpha = u_\alpha(-r, 0^+) - u_\alpha(-r, 0^-)$. The fact that equation (2.3) predicts interpenetration in a (usually) small neighborhood of the crack tip is not relevant in the present context. Dimensional considerations dictate that K must be of the form

$$K = (\text{applied stress}) \times (\sqrt{L}\, L^{-i\epsilon}) \times f \tag{2.4}$$

where L is a length quantity such as crack length and f is a nondimensional possibly complex function of dimensionless combinations of the material moduli and the geometric parameters. Two specific examples will be given in Section 3. The energy release-rate (per unit extension along the interface per unit length of crack front) is

$$\mathcal{G} = \left[\frac{(1-\nu_1)/G_1 + (1-\nu_2)/G_2}{4\cosh^2 \pi\epsilon}\right] K\bar{K} \tag{2.5}$$

where $\bar{}$ denotes the complex conjugate.

The tractions on the line directly ahead of the *subinterface crack* tip satisfy

$$\sigma_{22} + i\sigma_{12} = (K_{\mathrm{I}} + iK_{\mathrm{II}})(2\pi r)^{-1/2} \tag{2.6}$$

where K_{I} and K_{II} are the standard Mode I and Mode II stress intensity factors. The energy release-rate is

$$\mathcal{G} = \left[\frac{1-\nu_2}{2G_2}\right](K_{\mathrm{I}}^2 + K_{\mathrm{II}}^2) \tag{2.7}$$

As discussed earlier, the remote stresses in the *semi-infinite subsurface crack problem* are required to approach (for all θ but $\theta = \pi$) the characteristic Williams singular field of the interface crack, which can be written as

$$\sigma_{\alpha\beta} = \mathrm{Re}[K(2\pi r)^{-1/2} r^{i\epsilon} \tilde{\sigma}_{\alpha\beta}(\theta)] \tag{2.8}$$

with universal (complex) angular dependence $\tilde{\sigma}_{\alpha\beta}(\theta)$ for a given material pair. The remote crack face displacements approach equation (2.3). The only length quantity in the semi-infinite sub-interface crack problem is h. From dimensional considerations and by linearity it follows that

$$K_{\mathrm{I}} + iK_{\mathrm{II}} = cKh^{i\epsilon} + d\bar{K}h^{-i\epsilon} \tag{2.9}$$

where c and d are dimensionless complex constants depending only on dimensionless combinations of the moduli of the materials. The depth of the crack below the interface must appear as the factor $h^{i\epsilon}$ to combine with $L^{-i\epsilon}$ in equation (2.4) as the dimensionless term $(h/L)^{i\epsilon}$.

By considering a unit advance of the semi-infinite crack, one concludes by an energy argument, or equivalently by application of the J integral, that the energy release-rate given by equation (2.7) must be equal to that given in equation (2.5). That is

$$K_{\mathrm{I}}^2 + K_{\mathrm{II}}^2 = q^2 K\bar{K} \tag{2.10}$$

where

$$q = \left\{\frac{1}{2\cosh^2 \pi\epsilon}\left[\frac{G_2(1-\nu_1)}{G_1(1-\nu_2)} + 1\right]\right\}^{1/2} \tag{2.11}$$

Using an argument similar to that of Thouless et al. (1987), one can show that $d = 0$, and then substitution of equation (2.9) into (2.10) gives $c\bar{c} = q^2$ and thus

$$c = qe^{i\phi} \tag{2.12}$$

so that the relation (2.9) is fully determined apart from the single dimensionless function ϕ of the elastic moduli. A further simplification is achieved when use is made of Dundurs' (1969) observation that for problems of this class the moduli dependence can be expressed in terms of just two (rather than three) special nondimensional combinations. In plane strain, Dundurs' parameters are

$$\alpha = \frac{G_1(1-\nu_2) - G_2(1-\nu_1)}{G_1(1-\nu_2) + G_2(1-\nu_1)} \tag{2.13}$$

and

$$\beta = \frac{1}{2}\,\frac{G_1(1-2\nu_2) - G_2(1-2\nu_1)}{G_1(1-\nu_2) + G_2(1-\nu_1)} \tag{2.14}$$

where the roles of 1 and 2 are switched from Dundurs' definitions. These parameters vanish for identical materials across the interface and they change sign when the materials are switched. The quantities ϵ and q can be reexpressed as

$$\epsilon = \frac{1}{2\pi} ln\left[\frac{1-\beta}{1+\beta}\right] \tag{2.15}$$

and

$$q = \left[\frac{1-\beta^2}{1+\alpha}\right]^{1/2} \tag{2.16}$$

Table 2 Some special systems:

SYSTEM ①/②	α	β	ε	φ
Al_2O_3/Cu	.51	.089	−.028	.078
Cu/Al_2O_3	−.51	−.089	.028	−.104
MgO/Au	.51	.011	−.0036	.069
Au/MgO	−.51	−.011	.0036	−.108
Si/Cu	.094	−.033	.0105	.012
Cu/Si	−.094	.033	−.0105	−.014
MgO/Ni	.14	−.015	.0049	.020
Ni/MgO	−.14	.015	−.0049	−.023
Al_2O_3/Ti	.56	.12	−.039	.089
Ti/Al_2O_3	−.56	−.12	.039	−.114
Al_2O_3/Nb	.57	.060	−.019	.081
Nb/Al_2O_3	−.57	−.060	.019	−.122

	G	ν		G	ν
Au:	0.293×10^{11}N/M^2	0.417	Cu:	0.478	0.345
Ti:	0.434	0.322	Al_2O_3:	1.792	0.207
Ni:	0.808	0.314	Nb:	0.377	0.392
MgO:	1.283	0.175	Si:	0.688	0.220

An integral equation formulation of the semi-infinite subinterface crack problem is given in the Appendix. Numerical solution of the integral equation for various combinations of α and β has been carried out, the numerical values for $\phi(\alpha, \beta)$ are presented in Table 1. As discussed in the Appendix, the accuracy of these numerical results is believed to be within a small fraction of a percent. For sufficiently small α and β the linear approximation (obtained by a numerical fit in the range of small α and β)

$$\phi = 0.1584\,\alpha + 0.0630\,\beta \tag{2.17}$$

provides an adequate estimate of ϕ. For example, with $\alpha = 0.05$ and $\beta = 0.005$ the error of this formula is only 1.3 percent while with $\alpha = 0.14$ and $\beta = -0.015$ it is 6.7 percent.

Combining equations (2.12) and (2.9) gives the basic result for the stress intensity factors of the subinterface crack in terms of the complex stress intensity factor of the corresponding interface crack for conditions when $(h/L) \ll 1$:

$$K_I + iK_{II} = qe^{i\phi}Kh^{i\epsilon} \tag{2.18}$$

Note that $K_I = K_1$ and $K_{II} = K_2$ when α and β both vanish.

3 Applications and Implications

Moduli and values of α, β, ϵ, and ϕ are presented for six representative material combinations in Table 2. The shear modulus and Poisson's ratio listed for each material are polycrystalline values derived from Simmons and Wang (1971). The values for the cubic materials are the average of the Hashin-Shtrikman bounds and the values for the noncubic materials are the average of the Voigt and Reuss bounds. The largest values of ϵ are attained by the Cu/Al_2O_3 systems.

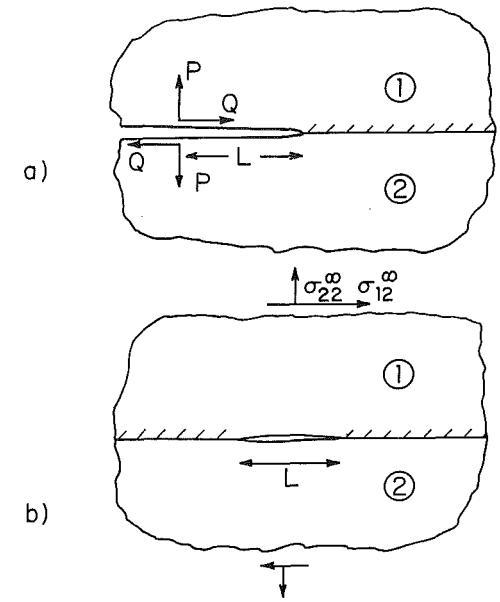


Fig. 2 Two basic interface crack problems

Solutions to a number of different interface crack problems exist in the literature. See Atkinson (1979) and Park and Earmme (1986) for recent discussions. The complex intensity factors for the two basic problems (Erdogan, 1965; Rice and Sih, 1965) shown in Fig. 2 will serve for discussion purposes here. For the semi-infinite crack along the interface between two elastic half-spaces and loaded by concentrated loads (per unit thickness) a distance L behind the tip,

$$K = (P + iQ)\,(\pi L/2)^{-1/2} L^{-i\epsilon} \cosh\, \pi\epsilon \tag{3.1}$$

In the case of a finite crack of length L on the interface between two half-spaces which are subject to remote stresses σ_{22}^{∞} and σ_{12}^{∞}, the complex intensity factor at the right tip is

$$K = (\sigma_{22}^{\infty} + i\sigma_{12}^{\infty})(1 + 2i\epsilon)(\pi L/2)^{1/2} L^{-i\epsilon} \tag{3.2}$$

(See the discussion in Rice and Sih for the behavior of the remote component σ_{11}. Those authors also gave solutions for a periodic row of collinear cracks along the interface of a solid under remote stressing, and for concentrated loads on the faces of a finite length crack.)

In all cases K can be written as

$$K = |K|\, e^{i\gamma} L^{-i\epsilon} \tag{3.3}$$

since $|L^{-i\epsilon}| = |e^{-i\epsilon \ln L}| = 1$. Then, by equation (2.18), the intensity factors of the semi-infinite subinterface crack are

$$K_I = q\,|K| \cos[\gamma + \phi + \epsilon\, \ln(h/L)] \tag{3.4}$$

and

$$K_{II} = q\,|K| \sin[\gamma + \phi + \epsilon\, \ln(h/L)] \tag{3.5}$$

Before asking what value of h/L is consistent with crack growth parallel to the interface, we first examine an interesting feature of the solution for arbitrary small values of h/L. If the collection of terms $\phi + \epsilon\, \ln(h/L)$ is small, as might easily be the case judging from the systems listed in Table 2, then the stress intensity factors are well approximated by

$$K_I + iK_{II} \cong qL^{i\epsilon}K = q\,|K|\, e^{i\gamma} \tag{3.6}$$

For example, in the case of the concentrated wedge force (3.1)

$$K_I + iK_{II} \cong q\,(P + iQ)\,(\pi L/2)^{-1/2} \tag{3.7}$$

assuming ϵ itself is small. Apart from the factor q, this is just

the classical result when no material discontinuity occurs. Similarly, the result for the finite crack (3.2) becomes

$$K_I + iK_{II} \cong q(\sigma_{22}^{\infty} + i\sigma_{12}^{\infty})(\pi L/2)^{1/2} \tag{3.8}$$

which is again the classical result multiplied by q. In other words, in these examples when ϵ and $\phi + \epsilon\, ln(h/L)$ are both small, the ratio of the Mode II to Mode I stress intensity factors is the same as in the corresponding classical problem but the energy release-rate is that of the interface crack as reflected by the factor q.

Now consider situations in fatigue, stress corrosion or under monotonic loading when the crack will tend to advance in its own plane approximately parallel to the interface. Assuming the fracture properties of material #2 are homogeneous along with its moduli, the crack will only advance in its plane if $K_{II} = 0$. If $K_{II} > 0$ it will tend to deflect downward away from the interface, while if $K_{II} < 0$ it will tend to grow upward. By equation (3.5), the condition for the crack to advance parallel to the interface in pure Mode I is

$$\sin[\gamma + \phi + \epsilon\, ln(h/L)] = 0 \tag{3.9}$$

or

$$\gamma + \phi + \epsilon\, ln(h/L) = 2\pi n;\ n = 0, \pm 1, \ldots \tag{3.10}$$

with the associated Mode I intensity

$$K_I = q\,|K| \tag{3.11}$$

Values of h/L from equations (3.10) are

$$h/L = \exp[(2\pi n - \gamma - \phi)/\epsilon];\ n = 0, \pm 1, \ldots \tag{3.12}$$

but only those values (if any) will be physically meaningful which are small compared to unity but not so small that the parts of the crack faces make contact, as will be discussed below. The crack length L increases as the crack advances and thus h cannot remain strictly constant. However, if h at the tip satisfies equations (3.12) approximately as L increases the slope dh/dL of the crack, the path will be small (and equal to the value given by equations (3.12)), with the crack thus nearly paralleling the interface when h/L is small.

As an illustration, consider the symmetric wedge loading ($Q = 0$) of the geometry in Fig. 2(a). By equations (3.1) and (3.3), $\gamma = 0$. For the material systems listed in Table 2, the largest magnitude of ϵ is 0.04, and it is readily seen that the only physically meaningful solution from equations (3.12), if any, is that associated with $n = 0$, i.e.,

$$h/L = \exp[-\phi/\epsilon] \tag{3.13}$$

Of the systems in Table 2, only Cu/Si, Si/Cu, Ni/MgO, and MgO/Ni have positive values of ϕ/ϵ and might therefore propagate a subinterface crack of the kind envisioned here for this particular geometry and loading. For Cu/Si, $h/L = 0.26$; while for Si/Cu, $h/L = 0.32$. The accuracy of these estimates may be somewhat questionable since they probably lie outside the range of h/L where the asymptotic analysis is accurate. For Ni/MgO, $h/L = 0.009$ and for MgO/Ni, $h/L = 0.017$, and these estimates should be accurate. Evidently the crack could satisfy a $K_{II} = 0$ criterion by propagating near the interface in either phase. We do not investigate here the configurational stability of those paths but expect, following Cotterell and Rice (1980), that only a path with a negative crack-parallel nonsingular stress term at the tip is stable. The conclusions for a finite crack paralleling the interface in Fig. 2(b) under remote tensile loading ($\sigma_{12}^{\infty} = 0$) are similar. Now, $\gamma \cong 2\epsilon$ and $h/L = \exp[-2 - \phi/\epsilon]$. Only the systems noted above will permit propagation of the subinterface crack parallel to the interface under the pure tensile loading. In general, however, it is important to note that the possibility of propagating a subinterface crack depends on both the material properties *and* the loading combination, so that subinterface cracks in the other systems may occur for other geometries and loadings.

The discussion and the analysis given above assume that contact between the crack faces of the subinterface crack does not occur. In applications where the near tip conditions of the subinterface crack is in pure Mode I and where h/L is not ludicrously small, it is unlikely that contact of the crack faces will be an issue. If the corresponding interface crack problem does indicate contact well away from the tip, at distances as large as h or greater, then the possibility of contact in the subinterface crack problem should certainly be checked. Solutions to (3.12) for h/L are only physically meaningful when h/L is not so small that contact will certainly occur or, what is more likely, that h is not so small that the crack lies so close to the interface that the material at the tip has properties which are affected by the existence of the interface.

Acknowledgment

This work was supported in part by the Office of Naval Research under Contract ONR-N00014-86-K-0753-Subagreement VB38639-0 from the University of California, Santa Barbara, California, and by the Division of Applied Sciences, Harvard University.

References

Atkinson, C., 1979, "Stress Singularities and Fracture Mechanics," *Applied Mechanics Reviews*, Vol. 32, pp. 123–135.

Bilby, B., and Eshelby, J. D., 1968, "Dislocations and the Theory of Fracture," *Fracture, An Advanced Treatise*, Vol. I, H. Liebowitz, ed., Academic Press, New York, pp. 99–182.

Cotterell, B., and Rice, J. R., 1980, "Slightly Curved or Kinked Cracks," *Int. J. Fracture*, Vol. 16, pp. 155–169.

Dundurs, J., 1969, "Edge-Bonded Dissimilar Orthogonal Elastic Wedges Under Normal and Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 650–652.

England, A. H., 1965, "A Crack Between Dissimilar Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 400–402.

Erdogan, F., 1965, "Stress Distribution in Bonded Dissimilar Materials with Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 403–410.

Head, A. K., 1953, "Edge Dislocations in Inhomogeneous Media," *Proceedings of the Physical Society*, London, Vol. B66, pp. 793–801.

Park, J. W., and Earmme, Y. Y., 1986, "Application of Conservation Integrals to Interface Crack Problems," *Mechanics of Materials*, Vol. 5, pp. 261–276.

Rice, J. R., and Sih, G. C., 1965, "Plane Problems of Cracks in Dissimilar Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 418–423.

Rice, J. R., 1968, "Mathematical Analysis in the Mechanics of Fracture," *Fracture, An Advanced Treatise*, Vol. II, H. Liebowitz, ed., Academic Press, New York, pp. 191–311.

Simmons, G., and Wang, H., 1971, *Single Crystal Elastic Constants and Calculated Aggregate Properties, A Handbook*, M.I.T. Press.

Thouless, M. D., Evans, A. G., Ashby, M. F., and Hutchinson, J. W., 1987, "The Edge Cracking and Spalling of Brittle Plates," *Acta Metall.*, Vol. 35, No. 6, pp. 1333–1341.

APPENDIX

Integral Equation Formulation and Solution

A dislocation formulation of the integral equation for the semi-infinite crack problem of Fig. 1 is used. The general procedure for formulating such problems is outlined by Bilby and Eshelby (1968) and Rice (1968) and the formulation and solution of a similar problem was given by Thouless et al. (1987).

The depth h of the crack below the interface will be taken to be unity here since the dependence of the solution on h is explicitly given in the body of the paper. With reference to Fig. 1, let $x = x_1$ and $y = x_2 + 1$ be coordinates centered at the crack tip with $y = 0$ parallel to the interface. Let $b_x(\xi)$ and $b_y(\xi)$ be the x and y components of an edge dislocation located on the x axis at $x = \xi$. This problem was first solved by Head (1953). At a point $(x, 0)$ the traction on a plane parallel to the x axis induced by the dislocation at $(\xi, 0)$ can be compactly derived us-

ing Muskhelishvili methods for plane strain elasticity and is given by

$$\sigma_{yy}(x)+i\sigma_{xy}(x)=2\bar{B}(\xi)(x-\xi)^{-1}+B(\xi)H_1(x-\xi)+\bar{B}(\xi)H_2(x-\xi) \quad (A.1)$$

where $i=\sqrt{(-1)}$, $\bar{}$ denotes the complex conjugate, and where

$$H_1(\zeta)=-16\delta\zeta/(4+\zeta^2)^2$$

$$H_2(\zeta)=-[(\lambda+\delta)\zeta-2i(\delta-\lambda)]/(4+\zeta^2)+8\delta/(\zeta-2i)^3$$

$$B(\xi)=G_2[b_x(\xi)+ib_y(\xi)]/[4\pi i(1-\nu_2)]$$

Here,

$$\delta=(\beta-\alpha)/(\beta+1)$$

and

$$\lambda=(\alpha+\beta)/(\beta-1)$$

where α and β are defined by equations (2.13) and (2.14).

The distribution of dislocations is chosen to give zero tractions on $y=0$ for $x<0$, i.e., $B(\xi)$ with b_x+ib_y now reinterpreted to correspond to a dislocation density must satisfy

$$\int_{-\infty}^{0}\{B(\xi)H_1(x-\xi)+\bar{B}(\xi)[2(x-\xi)^{-1}+H_2(x-\xi)]\}d\xi=0 \quad (A.2)$$

where the contribution with the $(x-\xi)^{-1}$ integrand is the Cauchy principal value integral.

The integral equation (A.2) is supplemented by the condition that the crack opening displacements far from the tip are the same as those for the interface crack. That is for large negative ξ, from equation (2.3),

$$b_x(\xi)+ib_y(\xi)=-\frac{d}{d\xi}[\delta_x(\xi)+i\delta_y(\xi)]$$

$$=\frac{[(1-\nu_1)/G_1+(1-\nu_2)/G_2]}{\cosh\,\pi\epsilon}\frac{i\bar{K}(-\xi)^{-i\epsilon}}{\sqrt{2\pi(-\xi)}} \quad (A.3)$$

and, therefore, as $\xi\to-\infty$

$$B(\xi)=(2\pi)^{-3/2}(1-\beta^2)^{1/2}(1+\alpha)^{-1}\bar{K}(-\xi)^{-1/2-i\epsilon} \quad (A.4)$$

To put the integral equation into a form suitable for numerical solution, make the change of variables

$$x=(u-1)/(u+1),\quad -1<u<1$$

$$\xi=(t-1)/(t+1),\quad -1<t<1$$

and let

$$\zeta\equiv x-\xi=2(u-t)/[(u+1)(t+1)]$$

Then, with $A(t)=B(\xi)$, equation (A.2) can be reduced to

$$\int_{-1}^{1}\bar{A}(t)(u-t)^{-1}dt+\int_{-1}^{1}\{A(t)H_1(\zeta)+\bar{A}(t)[1+t+H_2(\zeta)]\}(1+t)^{-2}dt=0 \quad (A.5)$$

The approximation for $A(t)$ was taken as

$$A(t)=(2\pi)^{-3/2}\frac{\bar{K}(1-\beta^2)^{1/2}}{2^{i\epsilon}(1+\alpha)}\frac{(1+t)^{1/2+i\epsilon}}{\sqrt{1-t}}+\frac{(1+t)}{\sqrt{1-t}}\sum_{k=1}^{N}c_kT_{k-1}(t) \quad (A.6)$$

where the c's are complex coefficients which must be obtained by the solution process and $T_j(t)$ is the Chebyshev polynomial of the first kind of degree j. The lead term in equation (A.6) gives the correct asymptotic behavior (A.4) as $\xi\to-\infty$ or, equivalently, as $t\to-1$. The stress intensity factors are given by

$$K_I+iK_{II}=(2\pi)^{3/2}\lim_{\xi\to 0-}\{(-\xi)^{1/2}\bar{B}(\xi)\}$$

or

$$K_I+iK_{II}=K(1-\beta^2)^{1/2}(1-\alpha)^{-1}+(2\pi)^{3/2}\sqrt{2}\sum_{k=1}^{N}\bar{c}_k \quad (A.7)$$

The solution procedure is essentially the same as that employed by Thouless et al. (1987). When the representation (A.6) is substituted into (A.5) the integral equation becomes an equation of the form

$$\sum_{k=1}^{N}[c_kI_1(u,k)+\bar{c}_kI_2(u,k)]=KI_3(u)+\bar{K}I_4(u) \quad (A.8)$$

where the integral expressions for the I's are readily identified. For example,

$$I_1(u,k)=\int_{-1}^{1}H_1(\zeta)T_{k-1}(t)(1+t)^{-1}(1-t)^{-1/2}dt \quad (A.9)$$

These integrals are evaluated numerically for specific values of u and k. Some further reduction of the integrals is necessary to render them in a form suitable for efficient numerical evaluation. Moreover, great care must be taken to ensure that the integration scheme provides a sufficiently accurate estimate of each integral. Accurate evaluation of these integrals is the major obstacle to accurate evaluation of the stress intensity factors.

With the real and imaginary parts of c_k for $k=1, N$ denoting the set of $2N$ unknowns, the real and imaginary parts of equation (A.8) are satisfied at N points u_i on the interval $-1<u<1$. The numerical results reported in Table 1 were computed using Gauss-Legendre points for the u_i. The solution procedure produces both K_I and K_{II}, yet from energy-release considerations the sum of the squares of the intensity factors is known (2.10). This provides an independent check on the accuracy of the numerical solution. The results reported in Table 1 were computed with $N=20$. The independent check (2.10) was satisfied to better than 0.1 percent for essentially all the (α, β) pairs reported in the Table. It is believed that the accuracy of ϕ is comparable.

On the Use of a Path-Independent Line Integral for Axisymmetric Cracks With Nonaxisymmetric Loading

An-Yu Kuo
Associate Member,
Structural Integrity Associates,
San Jose, CA 95118

A path-independent line integral J_A is derived for axisymmetric cracks under nonaxisymmetric loading conditions. The nonaxisymmetric crack problem is solved by expressing its boundary conditions as the sum of a Fourier series. Contribution on J_A from each of the Fourier terms in the nonaxisymmetric problem is shown to be decoupled from each other. Relationships between J_A and stress intensity factors are also presented for linear elastic fracture problems. Application of J_A to numerical fracture mechanics analysis is demonstrated by considering two example problems: an infinitely long circular bar with a penny-shaped internal crack at its center, and a circumferentially cracked pipe (both are under remote tension, bending, and torsion).

1 Introduction

Since the discovery of the path-independent J integral by Eshelby (1956) and Rice (1968), and later the J_K, L_K, and M integrals by Knowles and Sternberg (1972), path independent integrals have been widely used in fracture mechanics analyses. One of the most important applications of the path independent integrals is to calculate the stress intensity factor or energy release rate for a cracked solid. Since these integrals are path-independent, the stress intensity factor or energy release rate can be determined without having an accurate stress distribution near the crack tip. For instance, when the finite element method is used for a fracture mechanics analysis, even if the finite element grid near the crack tip is relatively coarse, the stress intensity factor and energy release rate can still be accurately predicted based on a path-independent integral evaluated at a contour far away from the crack tip.

While the energy release rate for cracks in a plane stress or plane strain problem can be readily determined by a path independent line integral (see, e.g., Budiansky and Rice, 1973), attempts in extending the concept to an axisymmetric problem have also been pursued in parallel. In the case of a linear elastic, axisymmetric solid under axisymmetric loading conditions, it has been shown by Astiz et al. (1977) and Bergkvist and Lan Huong (1977) that the energy release rate per unit crack advance in the radial direction per unit crack length along the circumferential direction is given by:

$$J=\frac{1}{R_c}\left\{\int_\Gamma [Wrdz-(T_r u_{r,r}+T_z u_{z,r})rds] - \int_B (W-T_\theta u_r/r)dB\right\} \quad (1)$$

where r, z, and θ are radial, axial, and circumferential coordinates, respectively, u_r and u_z are displacements, T_r, T_z, and T_θ are tractions, R_c is the r coordinate of the crack tip, W is strain energy density, Γ is a contour enclosing the crack tip (see Fig. 1), and B is the area enclosed by Γ. Note that an extra area integral has been introduced in the above integral. Equation (1) has been used by many researchers to calculate energy release rates for axisymmetric crack problems, e.g., Kumar et al. (1981). For power law materials, He and Hutchinson (1981) have eliminated the need of such an extra area integral in equation (1) by introducing

$$M=2\pi R_c^2 J=\int_A [Wx_i n_i - T_i u_{i,j} x_j - \left(\frac{2n-1}{n+1}\right) T_i u_i]\, dA \quad (2)$$

where $i, j, k = 1, 2, 3$, A is a torus-like surface ringing the edge of the axisymmetric crack, and n is the power-law constant.

This paper extends the use of the M integral introduced by He and Hutchinson (1981) to the mixed-mode fracture problems for an axisymmetric crack with nonaxisymmetric loading, e.g., a circumferentially cracked pipe under combined tension, bending, and torsion.

2 Line Integral for Axisymmetric Loading

For the convenience of later use in this paper, equation (2) is briefly rederived in this section by a method different from that used by He and Hutchinson (1981) or Astiz et al. (1977).

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 27, 1987; final revision May 4, 1987.

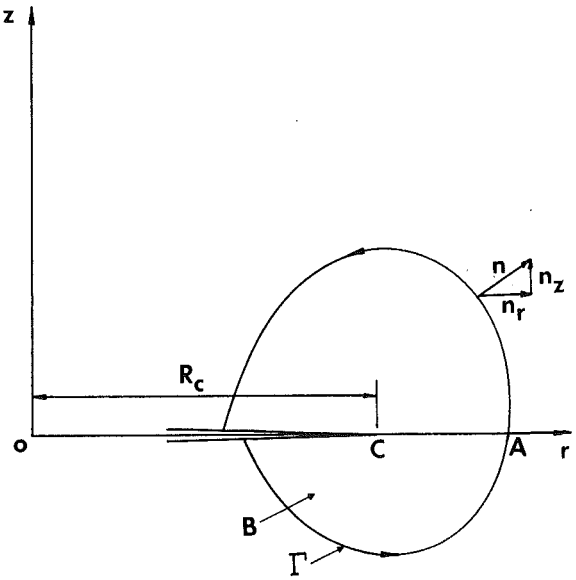


Fig. 1 An axisymmetric crack in an axisymmetric solid

Consider an elastic solid with a power law stress-strain relationship of

$$(\epsilon/\epsilon_0) = \alpha\,(\sigma/\sigma_0)^n \tag{3}$$

where ϵ and σ are uniaxial strain and stress, respectively, ϵ_0 and σ_0 are yield strain and yield stress, respectively, and α and n are material constants.

The three conservation integrals presented by Knowles and Sternberg (1972) are

$$J_k = \int_A (Wn_k - T_i u_{i,k})\, dA \tag{4}$$

$$L_k = \int_A e_{kij}\,(Wx_j n_i + T_i u_j - T_i u_{i,i} x_j)\, dA \tag{5}$$

$$M = \int_A \left[Wx_i n_i - T_i u_{i,j} x_j - \left(\frac{2n-1}{n+1}\right) T_i u_i\right] dA \tag{6}$$

where $i, j, k = 1, 2$, or 3, A is a closed surface with components n_i as its outward unit normal, e_{ijk} is an alternating tensor, W is strain energy density, x_i is the ith component of a Cartesian coordinate, and T_i and u_i are tractions and displacements, respectively. In equations (4) and (5), only J_1, J_2, and L_3 are useful for the present problem because J_3, L_1, and L_2 would result in integrals with a nonzero contribution from the crack surfaces.

Next, let's consider crack problems in an elastic, homogeneous, isotropic, axisymmetric solid under axisymmetric loading conditions. As a result of its axisymmetry, all the physical quantities, including displacements, stresses and strains, are functions of r and z only, where r, z, and θ are cylindrical coordinates with the z axis being the axis of revolution. It is further assumed that the x_3 axis of a Cartesian coordinate system coincides with the z axis of the cylindrical coordinate system and that the crack surfaces are located at the plane defined by $z = 0$ and are stress free. To apply the conservation integrals defined in equations (4–6), an integration surface A is chosen, as illustrated in Fig. 1, by rotating around the z axis an arbitrary contour, Γ, which encloses the crack tip. Thus, A is the outer surface of a torus-shaped body with Γ as the boundary of its cross section.

With the transformation $x_1 = r\cos\theta$, $x_2 = r\sin\theta$, $x_3 = z$, and similar relationships between (T_1, T_2, T_3) and (T_r, T_z), and (u_1, u_2, u_3) and (u_r, u_z), and (n_1, n_2, n_3) and (n_r, n_z), the conservation integrals defined in equations (4–6) become

$$J_1 = J_2 = 0 \tag{7}$$

$$L_3 = 0 \tag{8}$$

$$M = 2\pi \int_\Gamma \left[Wx_\alpha n_\alpha - T_\alpha u_{\alpha,\beta} x_\beta - \left(\frac{2n-1}{n+1}\right) T_\alpha u_\alpha\right] rds \tag{9}$$

where $\alpha, \beta = r$ or z. Physically, equations (7–9) imply that the energy release rates are zero for the crack to translate along the x_1 and x_2 axes or to rotate about the x_3 axis, and that the energy release rate is not zero for the crack to expand uniformly (see Budiansky and Rice, 1973).

Since the integrands in both equations (1) and (9) are proportional to $(1/R)$ as R approaches zero where R is the distance to the crack tip, it can be shown that the M integral defined by equation (9) and the J integral defined by equation (1) are related by

$$J = \frac{M}{2\pi R_c^2} \tag{10}$$

where R_c is the r coordinate of the crack tip. Therefore, the energy release rate per unit crack advance in the radial direction per unit crack front length along the circumferential direction can be calculated by

$$J_A = \frac{M}{2\pi R_c^2} = \int_\Gamma \left[Wx_\alpha n_\alpha - T_\alpha u_{\alpha,\beta} x_\beta - \left(\frac{2n-1}{n+1}\right) T_\alpha u_\alpha\right] \frac{C}{R_c^2}\, ds \tag{11}$$

A subscript A is used in equation (11) because, as will be discussed in the next section, the line integral defined by equation (11) is no longer equal to the energy release rate J when the axisymmetric crack is subject to nonaxisymmetric loadings.

3 Non Axisymmetric Loads in a Linear Elastic Solid

The J_A integral defined in equation (11) can be extended to crack problems in a linear elastic, homogeneous, isotropic, axisymmetric solid under nonaxisymmetric loading conditions. Solutions to such a non axisymmetric loading case are usually obtained by expressing the nonaxisymmetric loads and its resulting displacement fields as the sum of a Fourier series with circumferential coordinate θ as their arguments, and, thereafter, determining the Fourier coefficients with the properties of linearity and orthogonality of the harmonic functions. Such a Fourier series expansion solution technique for problems of nonaxisymmetric loads in an axisymmetric solid has been widely used in finite element analyses (see Wilson, 1965, for example).

Consider an axisymmetric solid D with outer surface A and prescribed stresses and displacements on A_σ and A_u, respectively, where A_σ and A_u are subsets of A. The nonaxisymmetric boundary conditions on A_σ and A_u can then be expanded in Fourier series as

$$T_\alpha(r,z,\theta) = \sum_{m=0}^{\infty} [T_{\alpha m}(r,z)\cos(m\theta) + \bar{T}_{\alpha m}(r,z)\sin(m\theta)] \text{ on } A_\sigma \tag{12}$$

and

$$u_\alpha^0(r,z,\theta) = \sum_{m=0}^{\infty} [u_{\alpha m}^0(r,z)\cos(m\theta) + \bar{u}_{\alpha m}^0(r,z)\sin(m\theta)] \text{ on } A_u \tag{13}$$

where $\alpha = r$, z, or θ, T_α are applied tractions on A_σ, and u_σ^0 are prescribed displacements on A_u. Thus, solutions to the problem can be written as

$$u_\alpha(r,z,\theta) = \sum_{m=0}^{\infty} [u_{\alpha m}(r,z)\cos(m\theta) + \bar{u}_{\alpha m}(r,z)\sin(m\theta)] \quad (14)$$

$$\epsilon_{\alpha\beta}(r,z,\theta) = \sum_{m=0}^{\infty} [\epsilon_{\alpha\beta m}(r,z)\cos(m\theta) + \bar{\epsilon}_{\alpha\beta m}(r,z)\sin(m\theta)] \quad (15)$$

$$\sigma_{\alpha\beta}(r,z,\theta) = \sum_{m=0}^{\infty} [\sigma_{\alpha\beta m}(r,z)\cos(m\theta) + \bar{\sigma}_{\alpha\beta m}(r,z)\sin(m\theta)] \quad (16)$$

where $\alpha, \beta = r, z$, or θ. It is easy to see that, in equations (12)–(16), the $m = 0$ terms represent the solution of uniform tension or torsion, and the $m = 1$ terms correspond to the solution of pure bending.

Substitution of equations (14)–(16) into equations (4)–(6), yields

$$J_1 = J_2 = L_3 = 0 \quad (17)$$

and

$$J_A = \frac{M}{2\pi R_c^2} = \sum_{m=0}^{\infty} [J_{Am}(u_{\alpha m},\epsilon_{\alpha\beta m},\sigma_{\alpha\beta m}) + J_{Am}(\bar{u}_{\alpha m},\bar{\epsilon}_{\alpha\beta m},\bar{\sigma}_{\alpha\beta m})] \quad (18)$$

where $\alpha, \beta = r, z, \theta$, and

$$J_{Am}(u_{\alpha m},\epsilon_{\alpha m},\sigma_{\alpha m}) = \int_\Gamma [W_m(n_r r + n_z z) - T_{\alpha m}(u_{\alpha m,r} r + u_{\alpha m,z} z + \frac{u_{\alpha m}}{2})] \frac{r}{(1+\delta_{m0})R_c^2} ds \quad (19)$$

$$W_m = (\sigma_{\alpha\beta m}\epsilon_{\alpha\beta m})/2 \quad (20)$$

$$T_{\alpha m} = \sigma_{\alpha r m} n_r + \sigma_{\alpha z m} n_z \quad (21)$$

$$\delta_{m0} = 1 \quad \text{if } m = 0, \quad \delta_{m0} = 0 \quad \text{if } m \neq 0 \quad (22)$$

From equation (18), it is seen that, for the problems of an axisymmetric solid under nonaxisymmetric loading conditions, not only displacements, stresses, and strains, but also the J_A integral can be expanded as the sum of a Fourier series, i.e., each one of the harmonic components in a Fourier series expansion of the solution is decoupled with all the other harmonics and can be solved individually.

It is worth noting that J_A is not a function of θ and is no longer equal to the energy release rate J. In fact, J_A can be interpreted as the average energy release rate over the entire crack front $0 \leq \theta < 2\pi$. Similarly, J_{Am} is the average energy relase rate over the entire crack front when the crack is under the mth mode loading. However, as will be discussed in the next section, there exists a relationship between J_A and the stress intensity factors or J.

4 Relationships to Stress Intensity Factors

In this section, we'll establish a relationship between the J_A integral and stress intensity factors in linear elastic fracture mechanics.

For an axisymmetric crack, Sih (1971) has shown that the asymptotic stress distribution near the crack tip, which is independent of the crack geometry and loading conditions, can be written in a form of

$$\sigma_{\alpha\beta} = \frac{1}{\sqrt{2\pi R}} [K_I(\theta) f_{\alpha\beta}(\theta_1) + K_{II}(\theta)\, g_{\alpha\beta}(\theta_1) + K_{III}(\theta)\, h_{\alpha\beta}(\theta_1)] \quad (23)$$

where $\alpha, \beta = r, z$, or θ, $R = [(r-R_c)^2 + z^2]^{1/2}$, $\theta_1 = \sin^{-1}(z/R)$ is the angle between the r axis and the line connecting the crack tip and the point of interest in the r-z plane, and $f_{\alpha\beta}(\theta_1)$, $g_{\alpha\beta}(\theta_1)$, and $h_{\alpha\beta}(\theta_1)$ are functions of θ_1 and the Poisson's ratio ν. Details of $f_{\alpha\beta}$, $g_{\alpha\beta}$, and $h_{\alpha\beta}$ can be found in the paper by Sih (1971). Thus, the asymptotic stress distribution near the crack tip of an axisymmetric crack can be written as

$$\sigma_{\alpha\beta}(r,z,\theta) = \frac{1}{\sqrt{2\pi R}} \sum_{m=0}^{\infty} \{[K_{Im}\cos(m\theta) + \bar{K}_{Im}\sin(m\theta)]\, f_{\alpha\beta}(\theta_1) + [K_{IIm}\cos(m\theta) + \bar{K}_{IIm}\sin(m\theta)]\, g_{\alpha\beta}(\theta_1) + [K_{IIIm}\cos(m\theta) + \bar{K}_{IIIm}\sin(m\theta)]\, h_{\alpha\beta}(\theta_1)\} \quad (24)$$

In other words, the stress intensity factors can also be expressed as the sum of a Fourier series,

$$K_q(\theta) = \sum_{m=0}^{\infty} [K_{qm}\cos(m\theta) + \bar{K}_{qm}\sin(m\theta)], q = I, II, III \quad (25)$$

Substitution of equation (24) into equation (18) yields a relationship between J_A and the stress intensity factors can be written as follows

$$J_A = \sum_{m=0}^{\infty} [J_{AM}(K_{Im},K_{IIm},K_{IIIm}) + J_{Am}(\bar{K}_{Im},\bar{K}_{IIm},\bar{K}_{IIIm})] \quad (26)$$

where

$$J_{Am} = [(1-\nu^2)(K_{Im}^2 + K_{IIm}^2) + 2(1+\nu)K_{IIIm}^2]/[E(1+\delta_{m0})] \quad (27)$$

where ν and E are Poisson's ratio and Young's modulus, respectively, and δ_{m0} is a Delta function defined in equation (22). The energy release rate $J(\theta)$ at any location of the crack front can then be calculated based on $K_q(\theta)$ and the well-known J-K relationship (similar to equation (27)) for linear elastic solids.

Therefore, as can be seen from equations (26) and (27), for a general, three-dimensional fracture problem of an axisymmetric crack, the stress intensity factors can be calculated separately for each of its Fourier components. Usually, the Fourier expansion of the boundary conditions for an axisymmetric crack can be done in such a way that the three fracture modes (modes *I*, *II*, and *III*) are decoupled from each other and the stress intensity factors can be determined by equations (25) and (26). However, if it is difficult to decouple the three fracture modes for any of the Fourier terms, e.g., the interface crack problems, an additional conservation law similar to that introduced by Yau et al. (1980) can be derived to decouple the three K's. Application of J_A to the interface crack problems will be reported elsewhere.

5 Application to Numerical Analysis

In the foregoing we have established J_A as a path-independent line integral for an axisymmetric crack under nonaxisymmetric loading conditions. To investigate the effectiveness and applicability of the J_A integral in actual engineering analyses, let's first consider an example problem of an infinitely long circular bar containing a penny-shaped internal crack at its center. As shown in Fig. 2, the radius of the penny-shaped crack, a, is one half of the bar radius, b. Three loading conditions (remote tension, bending, and torsion) were applied at the ends of the bar. A two-dimensional finite element program, which can handle nonaxisymmetric loads in an ax-

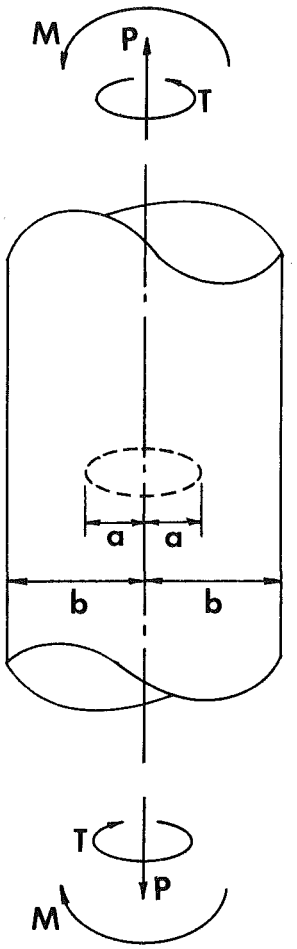


Fig. 2 An infinite bar with a penny-shaped crack

Table 1 Solutions for a penny-shaped crack in a bar

	$\frac{K_{Im}}{\sqrt{\sigma_0 a}}$	$\frac{K_{Im}}{\sqrt{\sigma_0 a}}$	$\frac{K_{IIIm}}{\sqrt{\sigma_0 a}}$
Load Case	Tension m = 0	Bending m = 1	Torsion m = 0
Path 1	1.5587	0.5207	0.4722
Path 2	1.2095	0.3448	0.3727
Path 3	1.2086	0.3665	0.3763
Path 4	1.2091	0.3733	0.3763
Path 5	1.2088	0.3737	0.3773
Path 6	1.2079	0.3723	0.3774
Path 7	1.2088	0.3746	0.3790
Path 8	1.2102	0.3762	0.3794
Reference	1.2196	0.3837	0.3797

isymmetric solid, was used to solve the above crack problem. The finite element mesh for the problem is illustrated in Fig. 3. In this finite element model, only half of the structure was modeled due to its inherent symmetry about the crack plane, and a total of 80 eight-node elements and 277 nodes were used. The two elements adjacent to the crack tip were quarter-node crack tip elements (Barsoum, 1974) and the rest of the elements were the typical eight-node isoparametric elements. The applied loads at the remote ends were

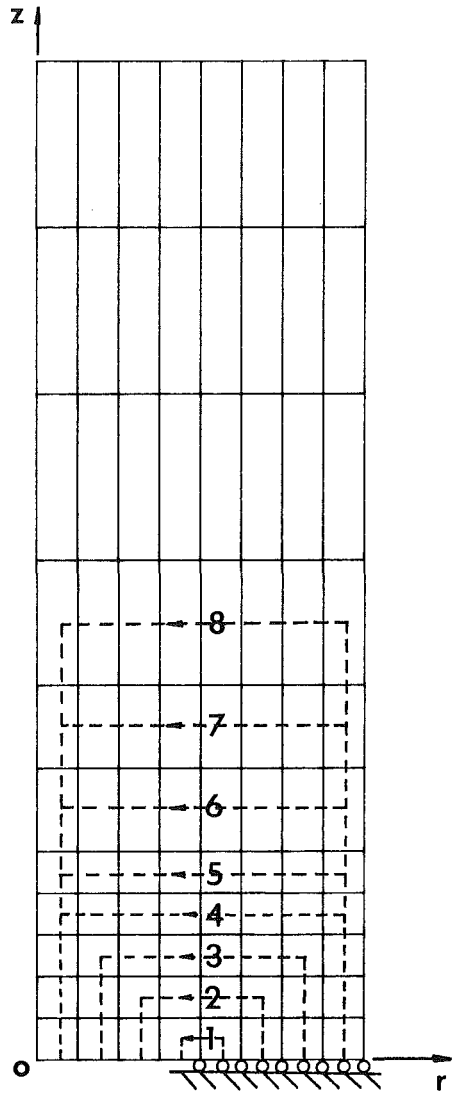


Fig. 3 Finite element mesh and integration contours

$P=\sigma_0\pi b^2$, $M=\sigma_0\pi b^4/4$, and $T=\sigma_0\pi b^4/2$

such that the maximum stresses at the remote ends of the bar were σ_0 for all three loading cases. Also shown in Fig. 3 are eight integration contours for calculating J_A.

Results of the finite element analysis and a reference solution given by Benthem and Koiter (1973) for the same problem are compared and summarized in Table 1. It is seen from this table that the stress intensity factors predicted by finite element methods and J_A are essentially path independent except the first path for the tension and torsion cases and the first three paths for the bending case. Errors at the first three paths are expectable since the finite element mesh used in that region was relatively coarse and stress distributions near the crack tip are thus not expected to be highly accurate. However, with only 80 elements in the model, the stress intensity factors calculated based on J_A from far field finite element solutions, e.g., stress intensity factors at path 8 in Fig. 3 and Table 1, are within 2 percent of the reference solution given by Bentham and Koiter (1973).

Secondly, as illustrated in Fig. 4, let's consider an infinitely long pipe with a 360 degree circumferential crack at its inside surface and with applied loads P, M, and T at infinity. For the interest of pressure vessel industry, the pipe and crack sizes are chosen to be $t/R_i=0.18$, and $a/t=1/2$. Like the first problem, it is assumed that the applied loads P, M, or T will each result

Table 2 Solutions for a circumferential crack in a pipe

	$\frac{K_{Im}}{\sqrt{\sigma_0 a}}$	$\frac{K_{Im}}{\sqrt{\sigma_0 a}}$	$\frac{K_{IIIm}}{\sqrt{\sigma_0 a}}$
Load Case	Tension m = 0	Bending m = 1	Torsion m = 0
Path 1	3.4768	3.1233	1.9193
Path 2	2.7163	2.3067	1.5558
Path 3	2.7163	2.3140	1.5542
Path 4	2.7147	2.3119	1.5549
Path 5	2.7176	2.3117	1.5498
Path 6	2.7142	2.3118	1.5686
Path 7	2.7135	2.3194	1.5769
Path 8	2.7139	2.3119	1.5772
Reference	2.7296	NA	NA

in a maximum stress of σ_0 at the uncracked section of the pipe. The same finite element mesh shown in Fig. 3 can still be used for this example problem except that a r translation has to be imposed in the coordinates of the previous finite element model. Tension solution to this problem has been given by Buchalet and Bamford (1976) and Labbens et al. (1976). Unfortunately, the bending and torsion solutions cannot be found after a literature survey. Results of the second example problem are summarized in Table 2. As shown in this table, except for the first path, stress intensity factors calculated by J_A are path-independent and are within 1 percent of the reference solutions for the tension case. Up to now, researchers and engineers often treat the combined tension and bending crack problems of a cracked pipe by taking the maximum combined stress and use that as the uniform remote tensile stress. However, most of the time, the maximum combined stress is mainly attributable to bending, and thus, as can be seen from Table 2, the approach of assuming the pipe is under remote tension would result in a conservative stress intensity factor and crack growth prediction. A more realistic calculation of the stress intensity factors based on J_A can ease such conservatism and provide a more reasonable crack growth prediction and allowable crack size for in-service inspection of cracked pipes.

6 Discussion and Conclusions

So far, application of the J_A integral has been restricted to homogeneous materials. However, path independence of J_A can be extended to the problems of an interface crack between two dissimilar, linear elastic solids with the fact that all the integrands in equation (19) are continuous across the material interface (C-A in Fig. 1). That is, the energy release rate of an axisymmetric crack between two disimilar, linear elastic, axisymmetric solids can be evaluated by integrating J_A along any contours which enclose the crack tip. Moreover, application of J_A to anisotropic materials can also be achieved by following a similar derivation procedure discussed in the above sections, but will not be elaborated on in this paper.

To sum up, the following conclusions can be drawn:

(1) For axisymmetric cracks under nonaxisymmetric loading conditions, J_A is the sum of the Fourier components J_{Am} which are decoupled from each other.

(2) J_A is the average energy release over the entire crack front.

(3) For a nonaxisymmetric fracture problem, the stress intensity factors and energy release rate J can be calculated directly from J_A.

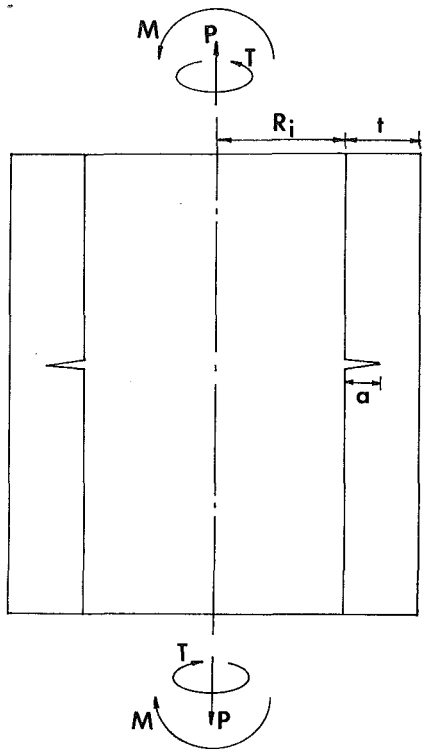


Fig. 4 An infinite pipe with a 360-degree internal circumferential crack

(4) J_A can also be used to calculate the energy release rate for interfacial cracks between two dissimilar materials.

7 References

Astiz, M. A., Elices, M., and Galrez, V. S., 1977, "On Energy Release Rates," *Fracture 1977*, Vol. 3, ICF4, Waterloo, Canada, June 19–24, pp. 395–399.

Barsoum, R. S., 1976, "Application of Quadratic Isoparametric Finite Elements in Linear Fracture Mechanics," *Internal Journal of Fracture*, Vol. 10, pp. 603–605.

Benthem, J. P. and Koiter, W. T., 1973, "Asymptotic Approximations to Crack Problems," *Method Analysis and Solutions of Crack Problems*, Sih, G. C., ed., Noordhoff International Publishing, Holland.

Berkvist, H., and Lan Huong, G.-L., 1977, "J-integral Related Quantities in Axisymmetric Cases," *International Journal of Fracture*, Vol. 13, pp. 556–558.

Buchalet, C. B., and Bamford, W. H., 1976, "Stress Intensity Factor Solution for Continuous Surface Flaws in Reactor Pressure Vessels," *Mechanics of Crack Growth*, ASTM, STP-590, pp. 385–402.

Budiansky, B., and Rice, J. R., 1973, "Conservation Laws and Energy Release Rates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, pp. 201–203.

Eshelby, J. D., 1956, "The Continuum Theory of Lattice Defects," *Solid State Physics*, Vol. 3, Seitz, F., and Turnbull, D., eds., Academic Press, New York, pp. 79–141.

He, M. Y., and Hutchinson, J. W., 1981, "The Penny-Shaped Crack and the Plane Strain Crack in an Infinite Body of Power-Law Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 830–840.

Knowles, J. K., and Sternberg, E., 1972, "On A Class of Conservation Laws in Linearized and Finite Elastostatics," *Archive for Rational Mechanics and Analysis*, Vol. 7, pp. 55–129.

Kumar, V., German, M. D., and Shih, C. F., 1981, "An Engineering Approach for Elastic-Plastic Fracture Analysis," Report NP-1931, Electric Power Research Institute, Palo Alto, CA.

Labbens, R., Pellissier-Tanon, A., and Heliot, J., 1976, "Practical Method for Calculating Stress Intensity Factors," *Mechanics of Crack Growth*, ASTM, STP-590, pp. 368–384.

Rice, J. R., 1968, "Mathematical Analysis in the Mechanics of Fracture," *Fracture—An Advanced Treatise*, Vol. II, Liebowitz, H., ed., Academic Press, New York, pp. 191–308.

Sih, G. C., 1971, "A Review of the Three Dimensional Stress Problem for Cracked Plate," *International Journal of Fracture Mechanics*, Vol. 7, pp. 39–61.

Wilson, E. L., 1965, "Structural Analysis of Axisymmetric Solids," *AIAA Journal*, Vol. 3, pp. 2269–2274.

Yau, J. F., Wang, S. S., and Corten, H. T., 1980, "A Mixed-Mode Crack Analysis of Isotropic Solids Using Conservation Laws of Elasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, No. 2, pp. 335–341.

A Finite Element Study of Stable Crack Growth Under Plane Stress Conditions: Part I—Elastic-Perfectly Plastic Solids

R. Narasimhan
Research Fellow in Applied Mechanics.

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

J. F. Hall
Assistant Professor of Civil Engineering.

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

A detailed finite element study of stable crack growth in elastic-perfectly plastic solids obeying an incremental plasticity theory and the Huber-Von Mises yield criterion is performed under plane stress, small-scale yielding conditions. A nodal release procedure is used to simulate crack extension under continuously increasing external load. It is found that the asymptotic angular extent of the active plastic zone surrounding the moving crack tip is from $\theta = 0$ deg to about $\theta = 45$ deg. Clear evidence of an elastic unloading region following the active plastic zone is found, but no secondary (plastic) reloading is numerically observed. The near-tip angular stress distribution inside the active plastic zone is in good agreement with the variation inside a centered fan, as predicted by a preliminary asymptotic analysis by Rice. It is also observed that the stress components within the plastic zone have a strong radial variation. The nature of the near-tip profile is studied in detail.

1 Introduction

A slow, stable crack extension phase is often observed (Broek, 1968; Green and Knott, 1975) in elastic-plastic materials prior to catastrophic failure during which a steady increase in applied load is required to propagate the crack. The primary reason for this is the reduced singularity in the strains that results when the crack propagates into material that has already deformed plastically.

Several investigators have contributed in providing an understanding of the mechanics and the practical implications of stable crack growth by using both analytical and numerical techniques. Problems that have received wide attention are crack extension in elastic-perfectly plastic materials under the conditions of anti-plane shear and Mode I plane strain. Chitaley and McClintock (1971) constructed an asymptotic analytical solution for steady, quasi-static crack growth under anti-plane shear conditions. Following preliminary investigations by Rice (1968, 1975), Rice et al. (1980) assembled an asymptotic solution for cracks growing in an incompressible elastic-perfectly plastic material under Mode I plane strain. The solution for this problem was also found independently by Slepyan (1974) and Gao (1980). Finally, the asymptotic analysis of Drugan et al. (1982) accounted for crack growth under Mode I plane strain in elastic-perfectly plastic materials without the restriction of elastic incompressibility.

However, by contrast not many asymptotic solutions are available for cracks growing in strain-hardening materials, primarily due to the difficulty involved in the analytical treatment of the governing equations. Amazigo and Hutchinson (1977) performed an asymptotic analysis for steady-state crack extension in a linear hardening material under antiplane shear and Mode I plane strain and plane stress. Ponte Casteñeda (1987) has recently extended their analysis to include the possibility of secondary reloading and has also treated Mode II plane strain and plane stress. Nevertheless, some questions pertaining to Mode I plane strain and plane stress, in the limit as the perfect plasticity is approached, are left unanswered by his investigation (see for example Section 4 of the present paper). Also, Gao and Hwang (1981) performed a preliminary investigation about the near-tip fields for a crack growing in a material governed by a more realistic power hardening law.

Finite element studies simulating crack growth, by using nodal release procedure, were conducted by Sorensen (1978) under antiplane shear and by Sorensen (1979) and Sham (1983) under Mode I plane strain. Dean and Hutchinson (1980) and Lam and McMeeking (1984) have used a Eulerian finite element formulation to study steady-state crack advance in the above cases. On the other hand, remarkably little work has been performed regarding crack growth under Mode I plane stress, notwithstanding its practical importance, as, for example, to thin aircraft structures. Also, a study of plane stress crack gorwth is compelling, because of the possibility of direct comparison with experiments based on the optical method of caustics, which in recent years has showed great

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 10, 1986; final revision May 21, 1987. Paper No. 87-WA/APM-20.

promise towards applications in ductile fracture (Rosakis and Freund, 1982; Zehnder et al., 1986). A preliminary analysis has been performed by Rice (1982), concerning the asymptotic nature of the stress and deformation fields near a growing crack tip in an elastic-perfectly plastic material under plane stress conditions. A complete (all-round) asymptotic solution for this problem has thus far remained elusive. Achenbach and Dunayevsky (1984) have recently investigated the variation of the plastic strain field along a ray ahead of the tip, based on the assumption of the validity of the asymptotic value for the stress field (Rice, 1982) up to the elastic-plastic boundary (see Section 4 for further discussion).

As far as numerical modelling is concerned, a steady-state Eulerian finite element study has been conducted recently by Dean (1983) for plane stress crack growth following the approach of Dean and Hutchinson (1980). However, as noted in Section 4 of this paper, the analysis of Dean (1983) is not very detailed, and certain issues pertaining to the near-tip stress and deformation fields have not been examined.

In the present investigation, a very detailed finite element study that provides great resolution near the crack tip has been carried out to model stable plane stress crack growth under continuous increase in external load, by using the nodal release procedure (Sorensen, 1979; Sham, 1983). Attention is restricted to elastic-perfectly plastic materials in the present analysis. This is a continuation of our earlier work (Narasimhan and Rosakis, 1986a), which analyzed the monotonic loading of a stationary crack under plane stress, small-scale yielding conditions.

2 Numerical Analysis

Formulation. The numerical modelling of the Mode I plane stress, small-scale yielding problem was discussed in detail by Narasimhan and Rosakis (1986a), who performed the analysis of a monotonically loaded stationary crack. In the present investigation, the results obtained by them will be used as initial conditions to simulate stable crack extension. Some of the features about the numerical analysis will be briefly outlined in this section. In the present paper $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ will represent an orthonormal frame centered at the crack tip and translating with it, while $\{\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'\}$ will be a fixed orthonormal frame situated at the position of the stationary crack tip.

The upper half of a domain R containing a crack and represented entirely by finite elements is shown in Figs. 1(*a*) and 1(*b*). The leading term in displacements of the linear elastic asymptotic solution,

$$u_\alpha = K_I\sqrt{\frac{r}{2\pi}}\hat{u}_\alpha(\theta), \tag{2.1}$$

was specified as a boundary condition on the outermost contour S of the domain.[1] The loading was applied through the Mode I stress intensity factor K_I or equivalently through the far-field value of the J integral. All plastic deformation was contained within a distance from the crack tip, which was less than 1/30 of the radius of S.

The active region of Fig. 1(*a*) has a total of 1704 four-noded quadrilateral elements and 3549 degrees of freedom. The quadrilaterals were formed from four constant strain triangles with static condensation of the internal node. Static condensation was also employed in the large region surrounding the active mesh, which always remained elastic. The cutout of Fig. 1(*a*), which is a fine mesh region near the crack tip, is shown in detail in Fig. 1(*b*). The small square elements near the crack tip have a size L, which is about 1/385 of the radius R_A of the active region and about 1/3400 of the radius of S.

[1] Throughout this paper, Greek subscripts will have range 1,2, while Latin subscripts will take values 1,2,3.

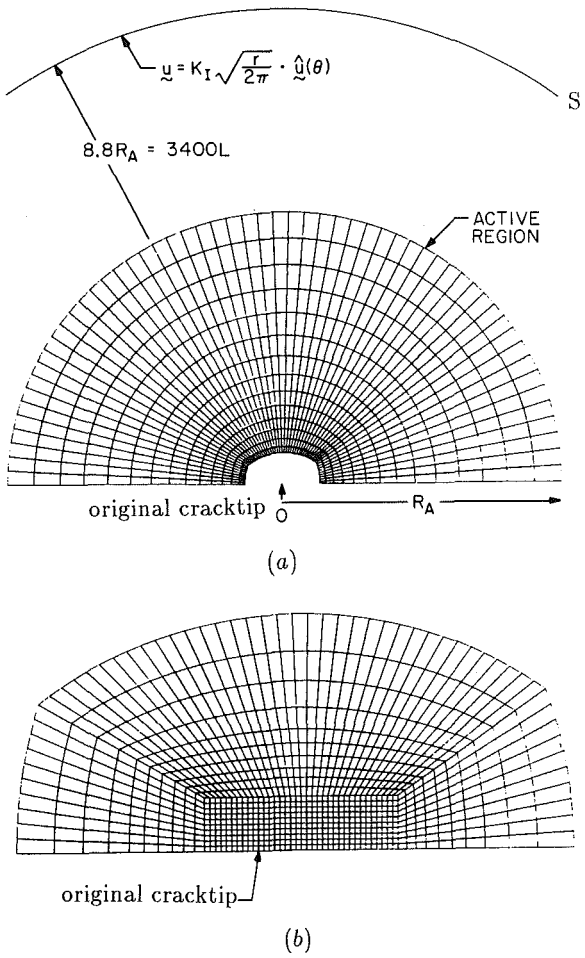


Fig. 1 Finite element mesh: (*a*) outer mesh; (*b*) fine mesh near the crack tip

Constitutive Assumptions. The material model that was considered here was that of an isotropic elastic-perfectly plastic solid. A small strain incremental plasticity theory was employed along with the Huber-Von Mises yield condition and the associated flow rule. The total strain rate tensor is assumed to be decomposed into elastic and plastic parts, and the constitutive law for material currently experiencing plastic deformation is given by,

$$\dot{\sigma}_{ij} = C^*_{ijkl}\dot{\epsilon}_{kl} = \left[C_{ijkl} - \frac{C_{ijpq}S_{pq}S_{mn}C_{mnkl}}{S_{rt}C_{rtuv}S_{uv}}\right]\dot{\epsilon}_{kl}. \tag{2.2}$$

Here C_{ijkl} is the isotropic, positive definite elasticity tensor, and S_{ij} is the deviatoric stress tensor. In the present analysis, the yield criterion and the constitutive law (2.2) were used along with the plane stress condition,

$$\sigma_{3i} \equiv 0. \tag{2.3}$$

On using equation (2.3) in equation (2.2), a constraint for $\dot{\epsilon}_{33}$ in terms of $\dot{\epsilon}_{\alpha\beta}$ can be obtained. The ratio of the Young's modulus to the yield stress in pure shear (E/τ_0) was taken as 350 and the Poisson's ratio (ν) as 0.3 in the computation.

Finite Element Scheme. A displacement based finite element method was employed and inertia effects were neglected in the analysis. The incremental finite element equilibrium equations were derived from the principle of virtual work by

linearization (Bathe, 1982). These equations were solved for each time step using the iterative Newton-Raphson method. An explicit integration procedure also known as the Tangential Predictor-Radial Return method (Schreyer et al., 1979) was employed together with subincrementation to integrate the incremental stress-strain law.

Solution Strategy. The plastic zone at the end of the stationary load history had a maximum extent, which occurred ahead of the crack tip, of about 100 times the smallest element size L. Subsequently, twenty one-element crack growth steps were simulated using the nodal release procedure, as described below. The stiff spring that was attached to the crack tip node in the x_2 direction, in order to enforce the symmetry condition ($u_2 = 0$), was removed and was replaced by the point load acting on it. This point load was subsequently relaxed to zero in twenty increments, at the end of which a traction-free element surface emerged, and the crack advanced by one element length L.

The externally applied load through the far-field J integral (or K_I) was increased simultaneously during the above nodal release procedure (as done by Sham, 1983), in order to model stable crack extension in a continuous manner. During crack growth the K field remained centered at the location of the original crack tip. This was because while the K field was specified at a radius of 3400L, the crack advanced by only 20L. For this purpose, a simple J versus crack growth (Δa) history with a constant slope was used. Following eighteen one-element crack growth steps at $T \equiv (E/\sigma_0^2) dJ/da = 5$, two crack growth steps at four different T values of 0, 5, 15, and 20 were simulated (T is the nondimensional Paris tearing modulus). The last two steps were thus carried out in order to study the effect of different rates of increase of external load on crack displacement increment. The computation was performed using a Cray XMP (Boeing Computer Services, Seattle). The total CPU time taken was about 3 CPU hours.

3 Asymptotic Fields Near Propagating Crack Tips

Plane Strain. Rice et al. (1980) assembled a near-tip solution for quasi-static crack advance under plane strain in an incompressible material ($\nu = 0.5$). This solution is essentially the Prandtl field (Rice, 1968) together with an elastic unloading sector following the centered fan. This was added to eliminate the negative plastic work that would otherwise occur at the trailing boundary of the fan. The asymptotic form of the plastic strains in the fan is given by,

$$\epsilon^p_{\alpha\beta} \sim \frac{(5-4\nu)}{2\sqrt{2}} \frac{\tau_0}{E} G_{\alpha\beta}(\theta) ln\left(\frac{\bar{R}}{r}\right), \quad r \to 0. \tag{3.1}$$

where $\bar{R}$ is an arbitrary length scale. The angular factors $G_{\alpha\beta}(\theta)$ are fully determined from an asymptotic angular integration of the plastic strain rates (Rice, 1982). It should be noted that the dominant log(r) term of (3.1) is much weaker than the $1/r$ plastic strain singularity near a monotonically loaded stationary crack tip (Rice, 1968).

Motivated by the above, Rice et al. (1980) proposed the following form for the near-tip crack opening rate during stable plane strain crack advance,

$$\dot{\delta} \sim \frac{\alpha \dot{J}}{\sigma_0} + \beta \frac{\sigma_0}{E} \dot{a} ln\left(\frac{R}{r}\right), \quad r \to 0. \tag{3.2}$$

In the above, α and β are constants and R is a length dimension, which is expected to scale with the plastic zone size under small-scale yielding conditions, so that

$$R = s\left(\frac{EJ}{\sigma_0^2}\right). \tag{3.3}$$

Here, J is the remotely applied value of the J integral which, under small scale yielding conditions, is given by,

$$\begin{aligned} J &= (1-\nu^2)\frac{K_I^2}{E} \quad \text{(plane strain)} \\ &= \frac{K_I^2}{E} \quad \text{(plane stress).} \end{aligned} \tag{3.4}$$

The constant β in (3.2) can be obtained from an all-round asymptotic solution (Rice et al., 1980; Drugan et al., 1982), whereas the constants α in (3.2) and s in equations (3.3) are undetermined from the asymptotic analysis.

The second term in (3.2) arises because of the log(r) dominant singularity in the material particle velocities. The first term in (3.2) encompasses the assumption that the higher-order terms in velocities are bounded and linear in load rate ($\dot{J}$ for small scale yielding). Also, for $\dot{a} = 0$, the right-hand side of (3.2) reduces to the correct expression for the discrete crack opening rate that is observed during the monotonic loading of a stationary crack. An asymptotic integration of (3.2) can be carried out to obtain the near-tip crack opening displacement during stable crack growth (when crack length a increases continuously with J) as follows,

$$\delta \sim \frac{\alpha r}{\sigma_0} \frac{dJ}{da} + \beta r \frac{\sigma_0}{E} ln\left(\frac{eR}{r}\right), \quad r \to 0, \tag{3.5}$$

where e is the base of the natural logarithm. As opposed to the monotonic loading of a stationary crack, equation (3.5) implies that the opening displacement at the crack tip is equal to zero during crack growth. However, as can be noticed from (3.5), the crack profile during growth exhibits a vertical tangent at the tip.

Plane Stress. The general features outlined above for plane strain apply to plane stress as well, with some modifications. No all-round asymptotic solution that satisfies all the boundary and symmetry conditions and that does not violate material stability postulates has yet been assembled for this case. However, Rice (1982) has performed a preliminary asymptotic analysis and has demonstrated that only two types of plastic sectors can exist near the crack tip. These are centered fan sectors in which radial lines are stress characteristics ($s_{rr} \equiv 0$) and constant stress sectors in which the Cartesian components of stress $\sigma_{\alpha\beta}$ are constant (not functions of angle θ). The asymptotic stress and deformation fields within the above plastic sectors and in elastic unloading sectors have been derived by Rice (1982).

The asymptotic stress field within a centered fan sector is summarized below assuming that it adjoins the $\theta = 0$ ray (similar to the stationary crack tip solution of Hutchinson, 1968):

$$\sigma_{rr} = \tau_0 \cos\theta, \ \sigma_{\theta\theta} = 2\tau_0 \cos\theta, \ \sigma_{r\theta} = \tau_0 \sin\theta, \ r \to 0 \tag{3.6}$$

Rice (1975) has demonstrated that if a centered fan adjoins the $\theta = 0$ line, then the plastic strains in front of the crack tip are given by:

$$\epsilon^p_{ij} = \frac{1}{2} \frac{\tau_0}{E} G_{ij} ln^2\left(\frac{\bar{R}}{r}\right), \ \theta = 0, \quad r \to 0, \tag{3.7}$$

where,

$$\left.\begin{aligned} G_{11} &= 0 \quad G_{12} = 0 \\ G_{22} &= 2 \quad G_{33} = -2 \end{aligned}\right\}. \tag{3.8}$$

It should be observed that the plastic strains (on the $\theta = 0$ ray) are singular as $\log^2(r)$ in plane stress, whereas in plane strain the plastic strains in the fan have a log(r) dominant singularity from (3.1). The stronger $\log^2(r)$ dominant plastic strain singularity in plane stress (which also occurs in antiplane shear) arises because the crack propagates into a centered fan region unlike in Mode I plane strain.

In plane stress, the material particle velocities have a log(r) singularity analogous to plane strain (see Rice, 1982). Hence one expects the crack opening rate during stable plane stress crack advance to have the same functional form as (3.2). Also

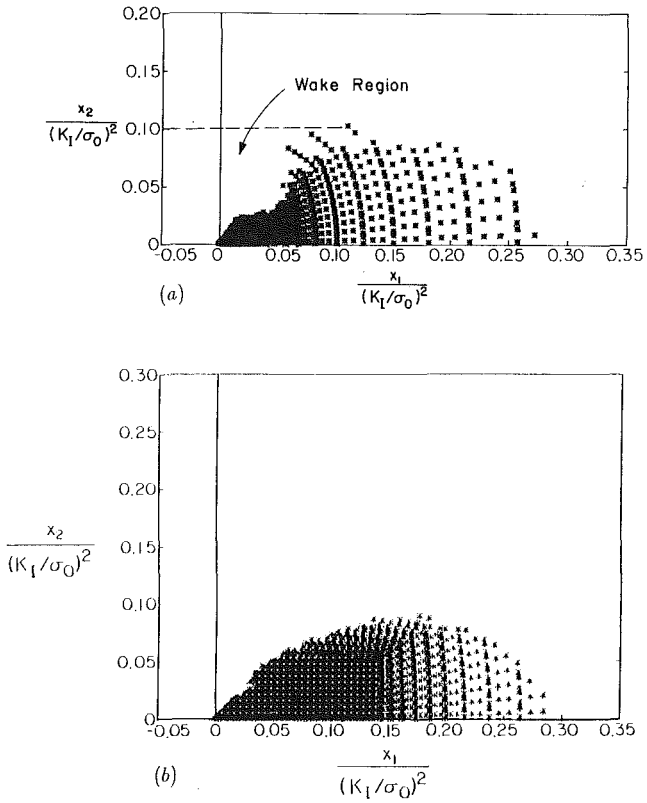


Fig. 2(a) Active plastic zone surrounding the propagating crack tip; (b) plastic zone corresponding to the stationary crack (Narasimhan and Rosakis, 1986a)

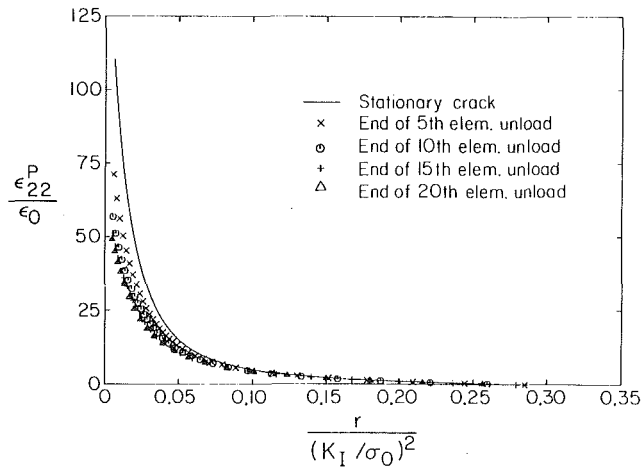


Fig. 3 Radial distribution of plastic strain ahead of the propagating crack tip for various levels of crack growth

the equations from (3.2) to (3.5) and the accompanying discussions are expected to apply for stable plane stress crack growth.

4 Results and Discussion

Plastic Zone. The active plastic zone surrounding the crack tip is shown in Fig. 2(*a*) in moving coordinates that have been made dimensionless by the parameter $(K_I/\sigma_0)^2$. For comparison purposes, the plastic zone corresponding to a stationary crack under plane stress conditions, which was obtained by Narasimhan and Rosakis (1986a), is shown in Fig. 2(*b*). In Fig. 2(*a*), the current crack tip is at the origin of the coordinate system, and this result was obtained at the end of twentieth crack growth step. A point in the figure represents an actively yielding integration station (currently on the yield surface in stress space) within an element. It can be seen from the figure that a large elastic unloading region follows the active plastic zone. The active plastic zone appears to occupy an asymptotic angular extent from $\theta = 0$ deg to about 45 deg, which will be verified later.

The elements behind the active plastic zone, which are close to the crack plane and which occupy the angular range from $\theta = 45$ to 180 deg, have unloaded elastically. These elements have previously experienced plastic yielding during the passage of the crack tip. The present numerical solution does not exhibit any secondary (plastic) reloading along the crack flank. This is in contrast to plane strain, where a secondary plastic region was found, extending behind the moving crack tip (Sham, 1983).

As can be seen from Fig. 2(*a*), the trailing boundary of the active plastic zone seems to have a kink, resulting in a shape similar to that observed in antiplane shear (Sorensen, 1978; Dean and Hutchinson, 1980). The parallel between plane stress and antiplane shear has been recognized earlier, from the presence of an intense deformation zone (centered fan) ahead of the moving tip in both cases. In the present computation, the kink in the active plastic zone developed during the first few crack growth steps and persisted with subsequent crack advance. Also, the overall features of the plastic zone did not change much after the first few crack growth steps.

The maximum radial extent of the plastic zone, which occurs directly ahead of the growing crack tip ($\theta = 0$), is $R_p \approx 0.28(K_I/\sigma_0)^2$, which is about the same as the stationary problem (Fig. 2(*b*)). Also, on comparing Figs. 2(*a*) and 2(*b*), it can be seen that the plastic zone for the propagating crack is similar in overall shape and size to that obtained for the stationary problem at points away from the crack tip. However, near the tip the two plastic zones seem to deviate in shape, primarily due to elastic unloading behind the trailing boundary during crack growth. As pointed out by Rice (1982), strong changes in plastic zone shape near the tip and a tendency to reestablish a shape similar to that for the stationary crack at points away from the tip are expected during the initial stages of crack growth. This can also be observed in the results for both antiplane shear (e.g., Sorensen, 1978) and plane strain (e.g., Sham, 1983).

Finally, it is noted that the plastic zone of Fig. 2(*a*) compares well with the steady-state result obtained by Dean (1983), from a Eulerian finite element formulation, except for the presence of the kink. However, the present finite element solution is more detailed, since it has a larger ratio of plastic zone to smallest element size of over 100 as compared to about 35 in Dean's computation. Also, unlike Dean's work, the initial phase of crack growth was simulated here under continuously increasing external load.

Radial Distribution of Plastic Strain. The radial distribution of normalized plastic strain, $\epsilon_{22}^p/\epsilon_0$, with respect to normalized distance, $r/(K_I/\sigma_0)^2$, ahead of the current crack tip is shown in Fig. 3. Results are presented for various levels of crack growth under steadily increasing value of far-field J at $T = 5$. The solid line in the figure is the plastic strain distribution ahead of a monotonically loaded stationary crack tip, which was obtained by Narasimhan and Rosakis (1986a). It can be seen that the plastic strain converges rapidly during the first few crack growth steps to an invariant distribution. For example, at a distance of $0.01(K_I/\sigma_0)^2$ ahead of the moving crack tip, the plastic strain dropped by 32 percent during the first five crack growth steps and by 17 percent, 8 percent, and 3 percent during the sixth to tenth steps, eleventh to fifteenth steps, and sixteenth to twentieth steps, respectively. Such rapid convergence was also observed in the numerical simulation of antiplane shear crack growth by Sorensen (1978).

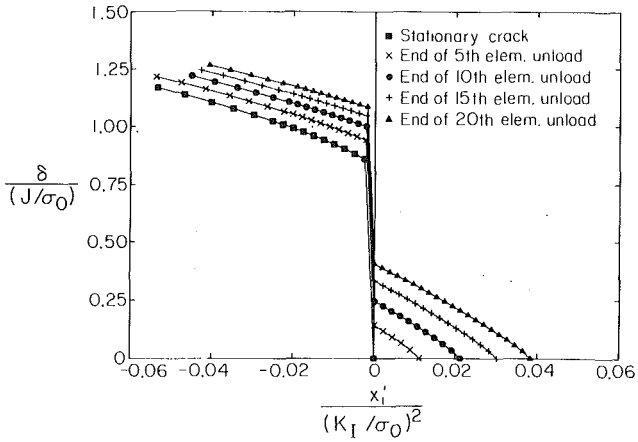


Fig. 4 Development of crack profile for various levels of crack growth

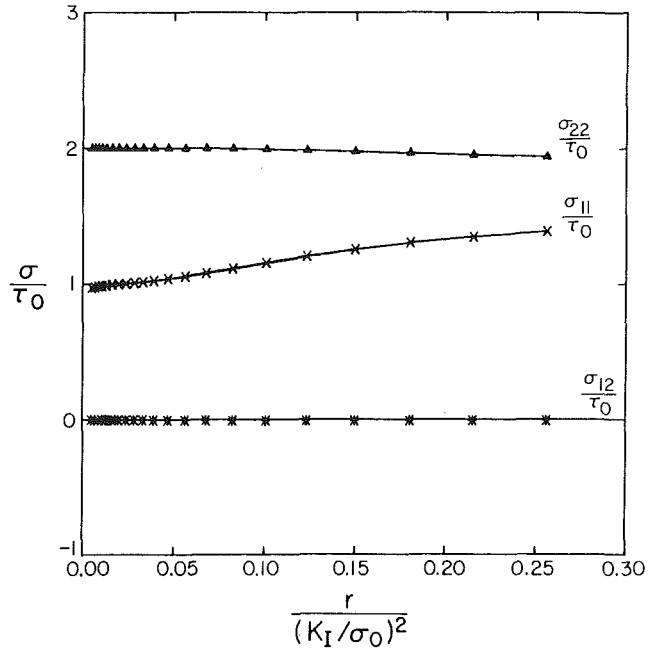


Fig. 5 Radial distribution of stresses ahead of moving crack tip

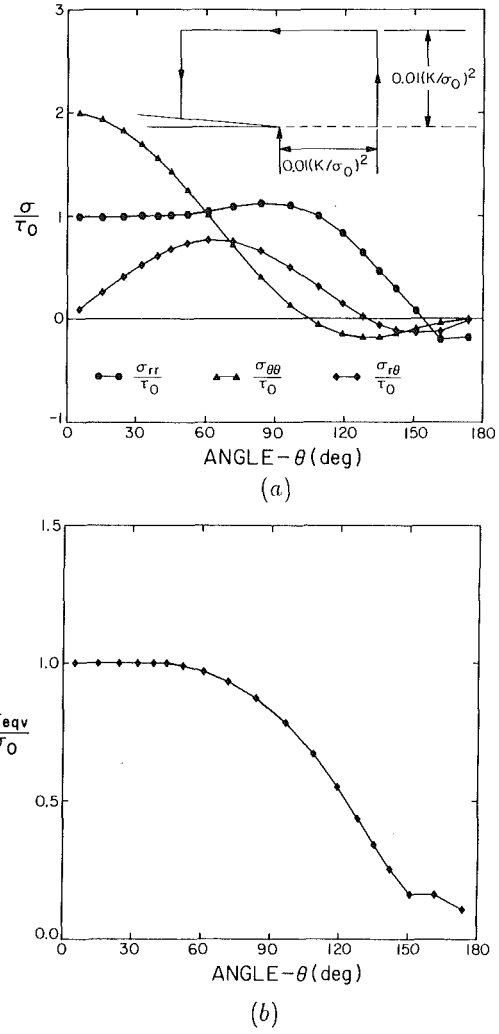


Fig. 6 Near-tip angular distribution of (a) normalized polar stress components and (b) Von Mises equivalent stress

As can be seen from Fig. 3, the plastic strain distribution ahead of the tip during growth exhibits a weaker singularity than in the stationary problem. This fact is the origin for the stable crack extension phase (McClintock and Irwin, 1965) observed in elastic-plastic materials (preceding instability), when crack growth occurs under steadily increasing external load.

Crack Profiles. The development of the crack opening profile for various levels of crack growth at $T=5$ is shown in Fig. 4 in the nondimensional form $\delta/(J/\sigma_0)$ versus $x_1'/(K_I/\sigma_0)^2$. The stationary crack profile is also plotted in the figure for comparison. As can be seen from the figure, the crack profile changes from a blunted form at the end of the stationary load history to a sharp shape during crack growth. This is because of the lessened strain concentration that results when the crack propagates into material that has already deformed plastically. In Section 5 this numerically obtained profile will be used to estimate the parameters α, β, and s in the asymptotic equation (3.5).

Radial Distribution of Stresses. The radial variation of the normalized stress components, $\sigma_{\alpha\beta}/\tau_0$, versus normalized distance ahead of the crack tip at the end of the twentieth release step is shown in Fig. 5. The centroidal values of stresses in the row of elements ahead of the tip have been used to construct this plot. The numerically obtained stresses very near the crack tip approach the asymptotic distribution given by equation (3.6) which assumes that there is a centered fan ahead of the tip. For example, from the finite element results at $r=0.01(K_I/\sigma_0)^2$, the values of σ_{11} and σ_{22} are $0.99\tau_0$ and $1.999\tau_0$, in excellent agreement with the values τ_0 and $2\tau_0$, respectively, given by (3.6).

It can be seen from Fig. 5 that the σ_{11} stress component exhibits a strong radial variation with a value of $1.40\tau_0$ at the elastic-plastic boundary. The value of σ_{11} differs from the asymptotic limit by less than 5 percent in the range $r<0.04(K_I/\sigma_0)^2$. This stress variation compares closely with that for the stationary crack (Narasimhan and Rosakis, 1986a). As noted by them, it suggests possible curving of the leading boundary of the fan at moderate distances from the tip. This will also be discussed later in connection with Fig. 7.

Near-Tip Angular Distribution of Stresses. The angular variation of the normalized polar stress components at a distance of $0.01(K_I/\sigma_0)^2$ from the moving crack tip (which is within $0.04R_p$) is shown in Fig. 6(*a*). The centroidal values of stresses in the elements lying on a rectangular contour surrounding the moving crack tip, which is shown as an inset in the figure, have been used to construct this plot. The angular variation along the above contour of the Von Mises equivalent stress, $\sigma_{eqv}=(3/2s_{ij}s_{ij})^{1/2}$ which has been normalized by σ_0, is shown in Fig. 6(*b*).

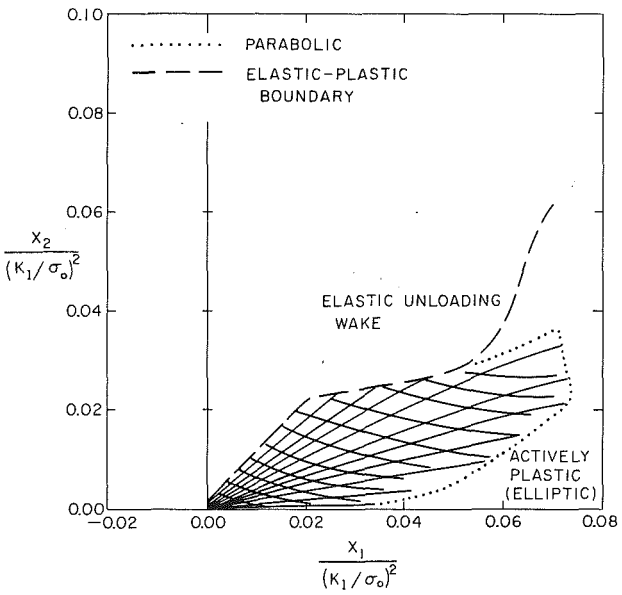


Fig. 7 Stress characteristics within active plastic zone

As can be seen from Fig. 6(*b*), σ_{eqv} becomes less than σ_0 for $\theta > 45$ deg, which suggests that the asymptotic angular extent of the active plastic zone is about 45 deg. This agrees well with visual observation of Fig. 2(*a*). However, from published results for crack advance under both antiplane shear (e.g., Dean and Hutchinson, 1980) and plane strain (e.g., Sham, 1983), where the asymptotic angular extent of active yielding was overestimated by finite element solutions, one is led to interpret the above conclusion with some caution. Also, from Fig. 6(*b*), it can be seen that σ_{eqv} does not become equal to σ_0 as θ approaches 180 deg, which implies that no secondary (plastic) reloading has been detected by this numerical solution.

The angular distribution of stresses (Fig. 6(*a*)) within the actively yielding region is in good agreement with the variation in a centered fan, as predicted by equation (3.6). For example, in the angular range 0 deg $\leq \theta <$ 45 deg, the values of $\sigma_{\theta\theta}$ and $\sigma_{r\theta}$ as given in Fig. 6(*a*) differ from that obtained using equation (3.6) by less than 1 percent and 4 percent, respectively. However, the value of σ_{rr} shown in Fig. 6(*a*) agrees with that given by equation (3.6) to within 8 percent in the angular range 0 deg $\leq \theta \leq 25$ deg and deviates substantially for 25 deg $< \theta <$ 45 deg. The reason for this discrepancy will be explained later in this section. Also, the angular stress distribution of Fig. 6(*a*) compares quite well with the finite element results of Dean (1983). However, as pointed out earlier, the present computation is considered to be more detailed than Dean's analysis.

The recent asymptotic analysis of Ponte Castañeda (1987) for steady, quasi-static crack growth in a linear hardening material is unfortunately not definitive about the asymptotic angular extent of the primary plastic zone in the limit as the perfect plasticity case is approached. He obtains a primary plastic angle of about 53.2 deg and 49 deg when the ratio E_t/E of the tangent modulus to the elastic modulus is 0.001 and 0.0001, respectively. Also the presence of a secondary reloading zone and its angular extent (which is extremely small) are not completely conclusive from his results, in the limit as E_t/E tends to be zero. The prediction of a very small reloading angle is, however, not inconsistent with the present numerical results, since such a tiny reloading zone cannot possibly be detected by a finite element scheme.

Nevertheless, the angular factors $\sigma_{ij}(\theta)$ of the dominant r^s term for the stresses given by Ponte Castañeda (1987) for $E_t/E = 0.001$ agree closely with the present numerical solution of Fig. 6, except for the σ_{rr} component, which seems to deviate in the angular range from $\theta = 25$ deg to about 100 deg. Also, the stress distribution obtained by Ponte Castañeda (1987) for $E_t/E = 0.001$ suggests yielding in compression for θ very close to 180 deg. While the present results do indicate a region near the crack flank where σ_{rr} is negative (see Fig. 6), no yielding in compression has been observed.

Stress Characteristics. The two families of stress characteristics within the active plastic zone, near the propagating crack tip, are shown in Fig. 7, using nondimensional crack tip coordinates. The dashed line in the figure is the boundary of the active plastic zone. The stress characteristics were plotted using the averaged stresses within the elements, as described by Kachanov (1974). These characteristics are lines along which the direct component of the stress deviator vanishes. The dotted line in the figure separates a region near the tip, in which the equations for the stresses are hyperbolic, from a region outside, in which they are elliptic. At each point on the dotted line, the condition for parabolicity of the governing equations for the stresses (see Kachanov, 1974) is satisfied. As can be seen from the figure, the two families of characteristics become mutually tangential to each other at every point along this dotted line, as it curves upwards from the $\theta = 0$ ray. However, it is not clear whether the elliptic region extends all the way up to the crack tip as a wedge of vanishingly small angular extent, as $r \to 0$ along the $\theta = 0$ ray, although there is some evidence to suggest this possibility. This observation might explain the difficulties encountered in constructing an all-around solution based on the asymptotic equations of Rice (1982).

Two important observations should be made from this figure. Firstly, it can be seen that a family of characteristics focusses at the crack tip in the angular range from $\theta = 0$ deg to about 25 deg, beyond which the characteristics seem to intersect the crack plane slightly behind the tip. This is probably because of the fact that, due to discretization, the crack tip is not precisely sensed in the finite element solution leading, according to the terminology of Sorensen (1978), to a "fuzzy crack tip phenomenon."

This was also observed in antiplane shear by Dean and Hutchinson (1980), who found that the active plastic zone obtained from their steady-state finite element solution extended from $\theta = 0$ deg to about 60 deg, while the charcteristics focussed at the tip only for angles less than 20 deg. For comparison, the analytical asymptotic solution of Chitaley and McClintock (1971) in antiplane shear crack growth has a centered fan region from $\theta = 0$ deg to 19.69 deg, followed by a large elastic unloading region and a tiny secondary reloading zone.

Secondly, the radial family of charcteristics in Fig. 7 bend downwards (towards the $\theta = 0$ ray) even for small distances ($r > 0.01(K_I/\sigma_0)^2$) from the tip. These two factors probably account for the strong discrepancy in the σ_{rr} stress component, between the finite element solution and the analytical asymptotic expression equation (3.6), in the angular range 25 deg $< \theta <$ 45 deg.

Finally, the strong radial variation in the stresses ahead of the crack tip (Fig. 5), combined with the observation of the change in nature of the governing equations as the distance from the crack tip is increased (Fig. 7), seems to disagree with the assumption of a constant stress field ahead of the tip made by Achenbach and Dunayevsky (1984).

5 Study of the Propagating Crack Profile

In this section, a value for the parameter β in the asymptotic crack opening rate (3.4) will be obtained by fitting the analytical asymptotic form to the numerically obtained values. The method employed is similar to that used by Sham (1983) in stable plane strain crack advance. Also, the linearity of the

higher-order term in (3.4) with respect to $\dot{J}$ will be verified from the numerical solution.

To that effect, the crack opening rate $\dot{\delta}$ is written as

$$\dot{\delta} \sim \dot{a}\frac{\sigma_0}{E}\left[f\left(\frac{a}{L}\right)+\beta ln\left(\frac{L}{r}\right)\right], \quad r\to 0, \tag{5.1}$$

where L is the smallest element size and is a convenient length scale and a is the crack length. Under small-scale yielding conditions, the function $f(a/L)$ can be shown to have the following form (see Section 3 and also Rice et al., 1980; Sham, 1983),

$$f\left(\frac{a}{L}\right)=g(T(a))+\beta ln\left(\frac{EJ(a)/\sigma_0^2}{L}\right), \tag{5.2}$$

where the quantity EJ/σ_0^2 has the dimension of length and is a measure of the plastic zone size. In the above equation, J, which is the remotely applied value of the J integral, and the nondimensional Paris tearing modulus $T=(E/\sigma_0^2)(dJ/da)$ are functions of the crack length a. If $g(T)$ is a linear function of T as was assumed in Section 3, then comparison of equation (5.2) with equations (3.2) and (3.3) gives

$$g(T)=\alpha T+\beta ln s, \tag{5.3}$$

where α and s will be taken as constants for limited amounts of crack growth.

The crack displacement increment at a fixed material point $(x_1',0)$, when the crack grows from a_1 to a_2, can be obtained by integrating (5.1) as follows (Sham, 1983),

$$\frac{E}{\sigma_0}\frac{\Delta\delta(x_1',a)}{L}\sim\Delta F+\beta\left[\frac{a_2-x_1'}{L}ln\left(\frac{eL}{a_2-x_1'}\right)-\frac{a_1-x_1'}{L}ln\left(\frac{eL}{a_1-x_1'}\right)\right]. \tag{5.4}$$

In the above equation, e is the base of the natural logarithm and

$$\left.\begin{aligned}\Delta\delta(x_1',a)&=\delta(x_1',a_2)-\delta(x_1',a_1)\\ \Delta F&=\int_{a_1/L}^{a_2/L}f(\zeta)d\zeta\end{aligned}\right\}. \tag{5.5}$$

The values of β and ΔF were obtained as the slope and axis intercept of a least-squares straight line fit to

$$\frac{E}{\sigma_0}\left(\frac{\Delta\delta(x_1',a)}{L}\right) \text{ versus } \Delta\left[\frac{a-x_1'}{L}ln\left(\frac{eL}{a-x_1'}\right)\right]$$

for successive one-element crack growth steps.

The representative straight line fits for crack growths under four different values of T of 0, 5, 15, and 20, which were simulated for the twentieth release step, is shown in Fig. 8. The first node behind the crack tip has been omitted and the data corresponding to the next five nodes have been plotted in this graph. The first node was omitted because it was observed that the crack tip element undergoes excessive rotation during the nodal release procedure. This conclusion was reached by performing a sensitivity study as described below. The average value of β based on the first six nodal points behind the crack tip was obtained as 2.1. On omitting the first node, it was found that a better straignt line fit can be made to the data corresponding to the net five nodal points (as in Fig. 8), which, however, gave a substantially lower average value of β of 1.7. The straight line fits underwent very little change on omitting the first and second nodes behind the tip, giving an average value of β of 1.67. On the basis of the above study, it is concluded that correct estimate for β, based on the crack displacement increments obtained from the finite element solution is around 1.70.

The value of ΔF obtained from the axis intercept can be taken approximately as

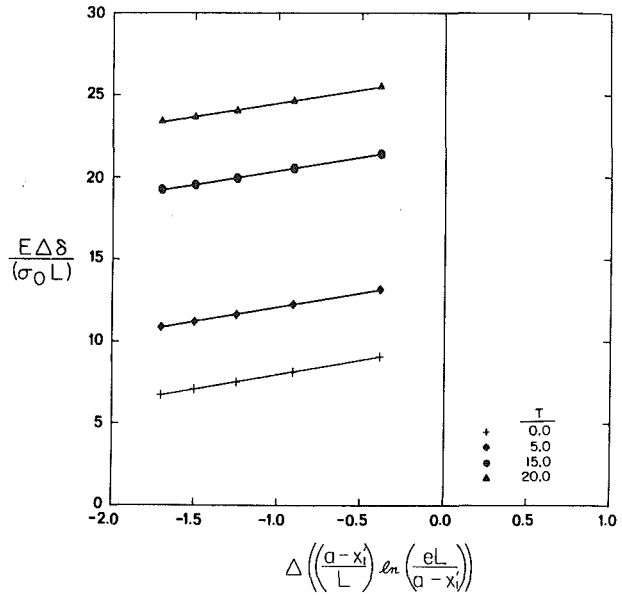


Fig. 8 Straight line fits to normalized crack displacement increments during the twentieth release step to determine β in asymptotic equation for crack opening rate

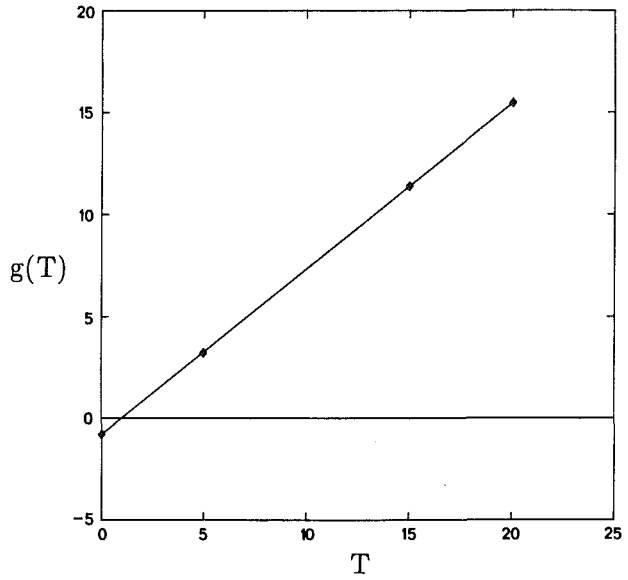


Fig. 9 Variation of higher-order term $g(T)$ in crack opening rate with respect to T. Straight line fit has been made to determine α and s.

$$\Delta F\approx\frac{(a_2-a_1)}{L}f\left(\frac{\bar{a}}{L}\right), \tag{5.6}$$

where $\bar{a}=(a_1+a_2)/2$. The value of $f(\bar{a}/L)$, which was computed from the above equation, was used along with the mean value of J during crack growth from a_1 to a_2 in equation (5.2) to determine $g(T)$. The values of $g(T)$ obtained as above for crack growth simulations under four different values of T during the twentieth release step are plotted against T in Fig. 9. It can be seen that a very accurate straight-line fit can be made to the numerically obtained points validating the assumption of linearity of $g(T)$ with respect to T made in Section 3.

On employing equation (5.3) (with β as 1.7), the values of α and s were obtained as 0.82 and 0.60 from the slope and axis intercept of the straight line fit (Fig. 9), respectively. From the analysis that included the first node behind the crack tip to determine β and ΔF (giving β as 2.1), the values of α and s were estimated as 0.82 and 0.24, respectively.

The value of α computed above is thus completely insensitive to the determination of β and is also in good agreement with the corresponding estimate from the opening displacement of the stationary crack, which was 0.85 as reported by Narasimhan and Rosakis (1986a). The value of s, on the other hand, seems to be extremely sensitive to the accuracy in determining β. This can also be observed in plane strain from the scatter in published numerical results for β and s (e.g., Sham, 1983; Dean and Hutchinson, 1980; Lam and McMeeking, 1984.

Finally, the asymptotic crack profile as given by (3.5) is plotted in nondimensional form for crack growth at $T=5.0$ in Fig. 10 with the parameters α, β and s taken as 0.82, 1.7, and 0.6, respectively. The values obtained from the finite element solution are also plotted in the figure for comparison. It is found that the predicted asymptotic crack profile is very close to the numerical solution in the range $r<0.04(K_I/\sigma_0)^2$.

A discussion of the implementation of a ductile fracture criterion will be presented in Part II of this investigation. This study will also result in the prediction of plane stress resistance curves for the case of perfect plasticity.

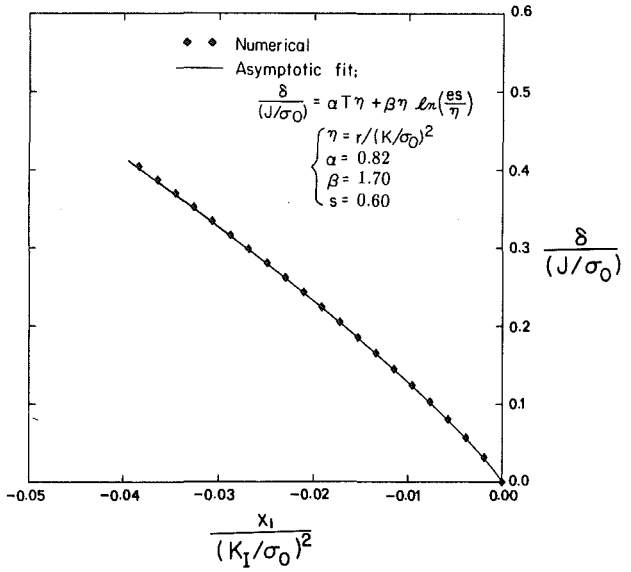


Fig. 10 Comparison of predicted asymptotic crack (solid line) with the finite element solution for stable crack growth at $T=5.0$

Acknowledgment

The authors would like to express their gratitude to Professor J. K. Knowles for his valuable advice and encouragement. This investigation was supported by the Office of Naval Research through ONR contract #N00014-85-K-0596. The computations were performed using the Supercomputer at Boeing Computer Services, Seattle. This was made possible through NSF contract #MEA-8307785. The above contracts and the facilities provided by Boeing Computer Services are gratefully acknowledged.

References

Achenbach, J. D., and Dunayevsky, V., 1984, "Crack Growth under Plane Stress Conditions in an Elastic Perfectly-Plastic Material," *Journal of Mechanics and Physics of Solids*, Vol. 32, pp. 89–100.

Amazigo, J. C., and Hutchinson, J. W., 1977, "Crack-Tip Fields in Steady Crack Growth with Linear Strain-Hardening," *Journal of Mechanics and Physics of Solids*, Vol. 25, pp. 81–97.

Bathe, K. J., 1982, *Finite Element Procedures in Engineering Analysis*, Prentice Hall, Englewood Cliffs, New Jersey.

Broek, D., 1968, "Some Considerations on Slow Crack Growth," *International Journal of Fracture*, Vol. 4, pp. 19–34.

Ponte Castañeda, P. P., 1987, "Asymptotic Fields in Steady Crack Growth with Linear Strain-Hardening," *Journal of Mechanics and Physics of Solids*, Vol. 35, pp. 227–268.

Chitaley, A. D., and McClintock, F. A., 1971, "Elastic-Plastic Mechanics of Steady Crack Growth Under Anti-Plane Shear," *Journal of Mechanics and Physics of Solids*, Vol. 19, pp. 147–163.

Dean, R. H., 1983, "Elastic-Plastic Steady Crack Growth in Plane Stress," *Elastic-Plastic Fracture: Second Symposium; Volume I—Inelastic Crack Analysis*, ASTM STP 803, American Society for Testing and Materials, pp. I-39–I-51.

Dean, R. H., and Hutchinson, J. W., 1980, "Quasi-Static Steady Crack Growth in Small-Scale Yielding," *Fracture Mechanics: Twelfth Conference*, ASTM STP 700, American Society for Testing and Materials, pp. 383–405.

Drugan, W. J., and Rice, J. R., 1984, "Restrictions on Quasi-Statically Moving Surfaces of Strong Discontinuity in Elastic-Plastic Solids," *Mechanics of Material Behavior*, Dvorak, G. J., and Shield, R. T., eds., Elsevier, 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 Mechanics and Physics of Solids*, Vol. 30, pp. 447–473.

Gao, Y. C., 1980, "Elastic-Plastic Fields at the Tip of a Crack Growing Steadily in a Perfectly Plastic Medium," *Acta Mechanica Sinicia*, in Chinese, No. 1, pp. 48–56.

Gao, Y. C., and Hwang, K. C., 1981, "Elastic-Plastic Fields in Steady Crack Growth in a Strain Hardening Material," *Proceedings of the Fifth International Conference on Fracture*, Francois, D., ed., Pergamon Press, New York, Vol. 2, pp. 669–682.

Green, G., and Knott, J. F., 1975, "On Effects of Thickness on Ductile Crack Growth in Mild Steel," *Journal of Mechanics and Physics of Solids*, Vol. 23, pp. 167–183.

Hutchinson, J. W., 1968, "Plastic Stress and Strain Fields at a Cracktip," *Journal of Mechanics and Physics of Solids*, Vol. 16, pp. 337–347.

Kachanov, L. M., 1974, *Fundamentals of the Theory of Plasticity*, Mir Publishers, Moscow, pp. 244–262.

Lam, P. S., and McMeeking, R. M., 1984, "Analysis of Steady Quasistatic Crack Growth in Plane Strain Tension in Elastic-Plastic Materials with Non-Isotropic Hardening," *Journal of Mechanics and Physics of Solids*, Vol. 32, pp. 395–414.

McClintock, F. A., and Irwin, G. R., 1965, "Plasticity Aspects of Fracture Mechanics," *Fracture Toughness Testing and its Applications*, ASTM STP 381, pp. 84–113.

Narasimhan, R., and Rosakis, A. J., 1986a, "A Finite Element Analysis of Small-Scale Yielding near a Stationary Crack Under Plane Stress," Caltech Report SM 86-21, Pasadena, CA, *Journal of Mechanics and Physics of Solids*, to appear, 1988.

Narasimhan, R., and Rosakis, A. J., 1986b, "Reexamination of Jumps across Quasistatically Propagating Surfaces Under Generalized Plane Stress in Anisotropically Hardening Elastic-Plastic Solids," Caltech Report SM 86-3, Pasadena, CA, ASME *Journal of Applied Mechanics*, 1987, Vol. 54, No. 3, pp. 519–524.

Rice, J. R., 1968, "Mathematical Analysis in Mechanics of Fracture," *Fracture: An Advanced Treatise*, Liebowitz, H., ed., Vol. 2, Academic Press, New York, pp. 191–311.

Rice, J. R., 1975, "Elastic-Plastic Models for Stable Crack Growth," *Mechanics and Mechanisms of Crack Growth*, May, M. J., ed., British Steel Corp. Physical Metallurgy Centre Publication, Sheffield, England, pp. 14–39.

Rice, J. R., 1982, "Elastic-Plastic Crack Growth," *Mechanics of Solids*, Hopkins, H. G., and Sewell, M. J., eds., Pergamon Press, Oxford, pp. 539–562.

Rice, J. R., and Sorensen, E. P., 1978, "Continuing Crack-Tip Deformation and Fracture for Plane Strain Crack Growth in Elastic-Plastic Solids," *Journal of Mechanics and Physics of Solids*, Vol. 26, pp. 163–186.

Rice, J. R., Drugan, W. J., and Sham, T. L., 1980, "Elastic-Plastic Analysis of Growing Cracks," *Fracture Mechanics: Twelfth Conference*, ASTM STP 700, American Society for Testing and Materials, pp. 189–221.

Rosakis, A. J., and Freund, L. B., 1982, "Optical Measurement of the Plastic Strain Concentration at a Cracktip in a Ductile Steel Plate," *Journal of Engineering Materials and Technology*, Transactions of ASME, Vol. 104, pp. 115–120.

Schreyer, H. L., Kulak, R. F., and Kramer, J. M., 1979, "Accurate Numerical Solutions for Elastic-Plastic Models," *ASME Journal of Pressure Vessel Technology*, Vol. 101, pp. 226–234.

Sham, T. L., 1983, "A Finite Element Study of the Asymptotic Near-Tip Fields for Mode I Plane Strain Cracks Growing Stably in Elastic-Ideally Plastic Solids," *Elastic-Plastic Fracture: Second Symposium, Volume I—Inelastic Crack Analysis*, ASTM STP 803, Shih, C. F., and Gudas, J. P., eds., American Society for Testing and Materials, Philadelphia, pp. 52–79.

Slepyan, L. I., 1974, "Growing Cracks during Plane Deformation of an Elastic-Plastic Body," *Izv. Akad. Nauk. SSSR, Mekhanika Tverdogo Tela*, Vol. 9, pp. 57–67.

Sorensen, E. P., 1978, "A Finite Element Investigation of Stable Crack Growth in Anti-Plane Shear," *International Journal of Fracture*, Vol. 14, pp. 485–500.

Sorensen, E. P., 1979, "A Numerical Investigation of Plane Strain Stable Crack Growth under Small-Scale Yielding Conditions," *Elastic-Plastic Fracture*, ASTM STP 668, American Society for Testing and Materials, pp. 151–174.

Zehnder, A. T., Rosakis, A. J., and Narasimhan, R., 1986, "Measurement of J Integral with Caustics: An Experimental and Numerical Investigation," Caltech Report SM 86-8, Pasadena, CA.

A Finite Element Study of Stable Crack Growth Under Plane Stress Conditions: Part II—Influence of Hardening

R. Narasimhan
Research Fellow in Applied Mechanics.

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

J. F. Hall
Assistant Professor of Civil Engineering.

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

A detailed finite element analysis is performed to model quasi-static crack growth under plane stress, small-scale yielding conditions in elastic-plastic materials characterized by isotropic power law hardening and the Huber-Von Mises yield surface. A nodal release procedure is used to simulate crack extension. Results pertaining to the influence of hardening on the extent of active yielding and the near-tip stress and deformation fields are presented. Clear evidence of an elastic unloading wake following the active plastic zone is found, but no secondary (plastic) reloading along the crack flank is numerically observed for any level of hardening. A ductile crack growth criterion based on the attainment of a critical crack opening displacement at a small microstructural distance behind the tip, is employed to investigate the nature of the J resistance curves under plane stress. In addition, the same criterion is employed to investigate the influence of hardening on the potential for stable crack growth under plane stress. It is found that predictions based on a perfectly plastic model may be unconservative in this respect, which is qualitatively similar to the conclusions reached in antiplane shear and Mode I plane strain.

1 Introduction

Several investigators have contributed in providing an understanding of the mechanics of stable crack growth by using both analytical and numerical techniques. Such works are reviewed in the introduction of Part I of the present investigation.

In this part a detailed finite element analysis is undertaken to model crack growth under plane stress in isotropic power hardening solids. This is a continuation of our earlier work (Narasimhan and Rosakis, 1986), which analyzed the monotonic loading of a stationary crack. Two crack growth histories (see Section 2) are simulated to study the mechanics problem of quasi-static crack extension and also to investigate the initial phase of stable growth under small-scale yielding, as would be observed in an experiment.

2 Numerical Analysis

Formulation. The numerical modeling of the Mode I plane stress, small-scale yielding problem was discussed in detail by Narasimhan and Rosakis (1986), who performed the analysis of a monontically loaded stationary crack. In the present investigation, the results obtained by them will be used as initial conditions to simulate quasi-static crack extension. The basic features of the numerical analysis have been summarized in Part I.

Constitutive Assumptions. The material model that was considered here was that of an elastic-plastic solid with an isotropic power law hardening behavior. A small strain incremental plasticity theory was employed along with the Huber-Von Mises yield condition and the associated flow rule. The Huber-Von Mises yield condition for isotropic hardending takes the form,

$$f(\sigma,\bar{\epsilon}^p) = F(\sigma) - \bar{\sigma}^2(\bar{\epsilon}^p), \tag{2.1}$$

where $F(\sigma) = 3/2\ \mathbf{S}\cdot\mathbf{S}$, and $\bar{\epsilon}^p = \int(2/3\epsilon^p_{ij}\epsilon^p_{ij})^{1/2}dt$ is the accumulated equivalent plastic strain. In the above, $\mathbf{S}$ is the deviatoric stress tensor, and $\bar{\sigma}(\bar{\epsilon}^p)$ is defined by the following power hardening rule,

$$\frac{\bar{\epsilon}^p}{\epsilon_0} = \left(\frac{\bar{\sigma}}{\sigma_0}\right)^n - \frac{\bar{\sigma}}{\sigma_0}. \tag{2.2}$$

Here σ_0 and ϵ_0 are the yield stress and strain in uniaxial tension.

The total strain rate tensor is assumed to be decomposed into elastic and plastic parts, and the constitutive law for material currently experiencing plastic deformation is given by,

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 10, 1986; final revision May 21, 1987.
Paper No. 87-WA/APM-21.

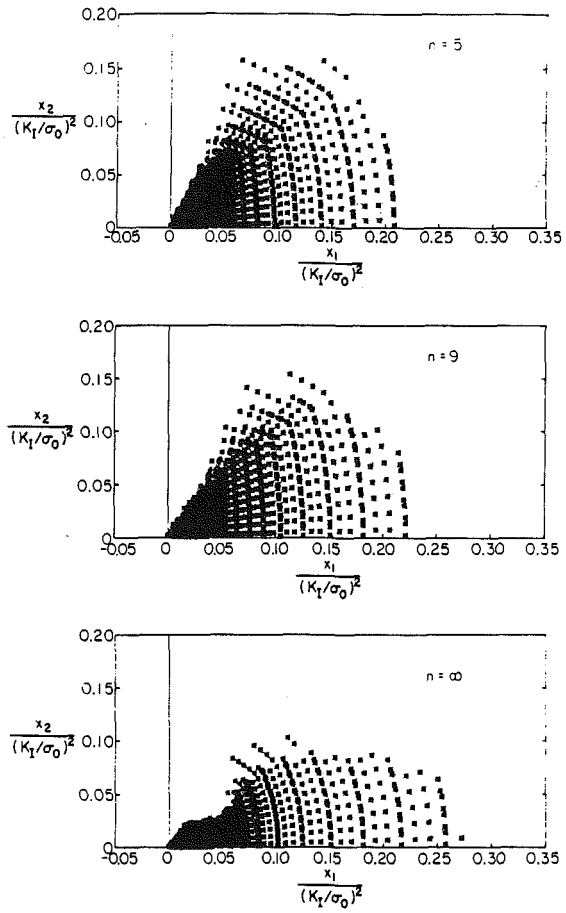


Fig. 1 Active plastic zone surrounding the propagating crack tip for various levels of hardening

$$\dot{\sigma}_{ij} = C^*_{ijkl}\dot{\epsilon}_{kl}$$

$$= \left[C_{ijkl} - \frac{C_{ijpq}S_{pq}S_{mn}C_{mnkl}}{S_{rt}C_{rtuv}S_{uv} + \frac{4}{9}\bar{\sigma}^2 H} \right] \dot{\epsilon}_{kl}. \tag{2.3}$$

Here C_{ijkl} is the isotropic, positive definite elasticity tensor and $H = d\bar{\sigma}/d\bar{\epsilon}^p$, which can be obtained from (2.2). In the present analysis, the yield criterion and the constitutive law were used along with the plane stress condition,

$$\sigma_{3i} \equiv 0. \tag{2.4}$$

On using equation (2.4) in (2.3), a constraint for $\dot{\epsilon}_{33}$ in terms of $\dot{\epsilon}_{\alpha\beta}$ may be obtained.

The computations were performed for two levels of hardening, $n = 5$ and 9. The ratio of the Young's modulus to the yield stress in pure shear (E/τ_0) was taken as 1400 and the Poisson's-ratio as 0.3 in the calculations.

Solution Strategy. In this study, two simple crack growth histories were simulated employing the nodal release procedure (see Part I and also Narasimhan et al. 1986). In the first case, the maximum plastic zone extent at the end of the stationary loading process was slightly more than 50 times the smallest element size L. Subsequently, twenty one-element crack growth steps were simulated using the nodal release procedure, holding the externally applied load fixed. This was achieved by imposing $T \equiv (E/\sigma_0^2)dJ/da = 0$ during crack growth. T is the nondimensional Paris tearing modules. The purpose of this investigation is to examine the nature of the near-tip stress and deformation fields for the mechanics problem of quasi-static crack growth without the influence of increase in applied load. Following Rice (1975), this would correspond to a hypothetical situation in which a cracked specimen is initially loaded by clamping portions of its boundary and imposing displacements, which is then followed by crack extension by saw-cutting ahead under fixed boundary displacements.

However, in an actual situation, after initiation, a crack will generally grow stably in an elastic-plastic material for an extent typically of the order of a few plastic zone sizes, during which the applied load will have to be increased to propagate the crack. A steady-state condition will then be reached, after which no further increase in applied load will be required for additional crack growth. In the second load history, stable crack extension was modelled (in a continuous manner) by simultaneously increasing the applied load during the nodal release procedure. This was accomplished by simulating fifteen one-element crack growth steps under $T = 1.5$, following the stationary loading process. The maximum extent of the plastic zone was over 100 times the smallest element length, L. Only the material with $n = 9$ was considered in this investigation.

In the following section, detailed results will be presented initially for $n = 5$ and 9 corresponding to the first load history. At the end of the section, comparison between the results for the two load histories will be made for the material with $n = 9$.

3 Results and Discussion

Active Plastic Zones. The active plastic zone surrounding the propagating crack tip after the twentieth crack growth step is shown in Fig. 1 for $n = 5$ and 9, in moving coordinates that have been made dimensionless by the self-similar parameter $(K_I/\sigma_0)^2$. The plastic zone for the stable plane stress crack growth in an elastic-perfectly plastic material is also shown for comparison.[1] The current crack tip is at the origin of the coordinate system, and a point in the figure represents an actively yielding integration station within an element.

A large elastic unloading region can be seen following the active plastic zone. No secondary (plastic) reloading along the crack flank has been observed for any level of hardening from the present numerical solution. The asymptotic angular extent of the active plastic zone, θ_p, decreases with decreasing hardening (increasing n). The values of θ_p are approximately 65 deg, 55 deg, and 45 deg for $n = 5$, 9, and ∞, respectively. The maximum radial extent of the active plastic zone, R_p, which occurs directly ahead of the crack tip, increases with decreasing hardening. The values of R_p are about $0.22(K_I/\sigma_0)^2$, $0.24(K_I/\sigma_0)^2$, and $0.28(K_I/\sigma_0)^2$ for $n = 5$, 9, and ∞, respectively.

Comparison of Fig. 1 with the plastic zone surrounding the stationary crack (Narasimhan and Rosakis, 1986) show that the active plastic zone becomes more acute (sharper) with the onset of crack growth. The results for the stationary problem show rounded plastic zones for the hardening cases, with yielding spreading beyond 90 deg near the crack tip. Strong changes in the near-tip plastic zone shape occurred during the first few crack growth steps, and then the overall features were unaltered with subsequent crack advance. The maximum radial extent of the plastic zone, R_p, given above for the propagation crack, is about the same as in the stationary problem for all levels of hardening.

A kink in the trailing boundary of the active plastic zone (Fig. 1) appears to develop for materials with low hardening and it becomes pronounced for the perfectly plastic case. The reason for this development could be related to the change in

[1] Throughout this paper, results given as $n = \infty$ will correspond to the perfectly plastic crack growth analysis of Part I.

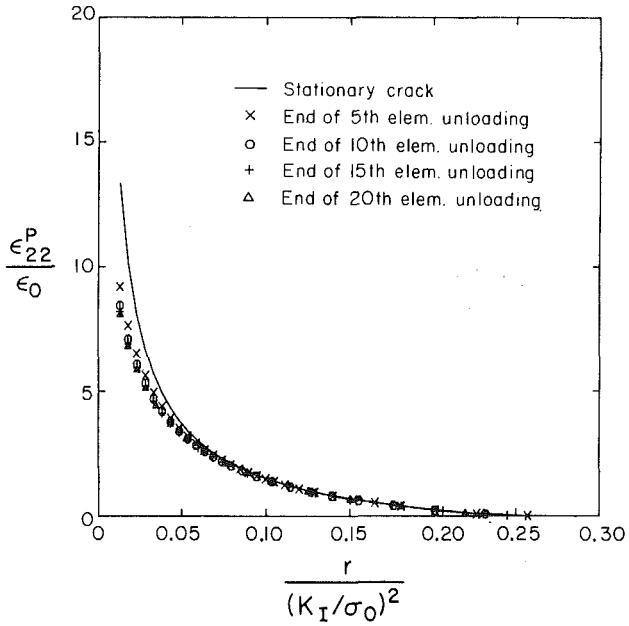


Fig. 2 Radial distribution of plastic strain ahead of the propagating crack tip for various levels of crack growth under fixed applied load for a material with $n = 9$

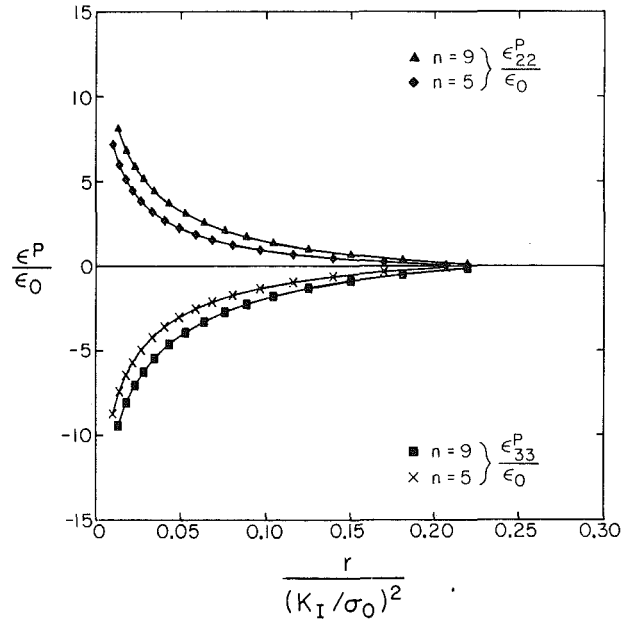


Fig. 3 Radial distribution of plastic strain ahead of the propagating crack tip at the end of the twentieth crack growth step at fixed applied load for $n = 5$ and 9

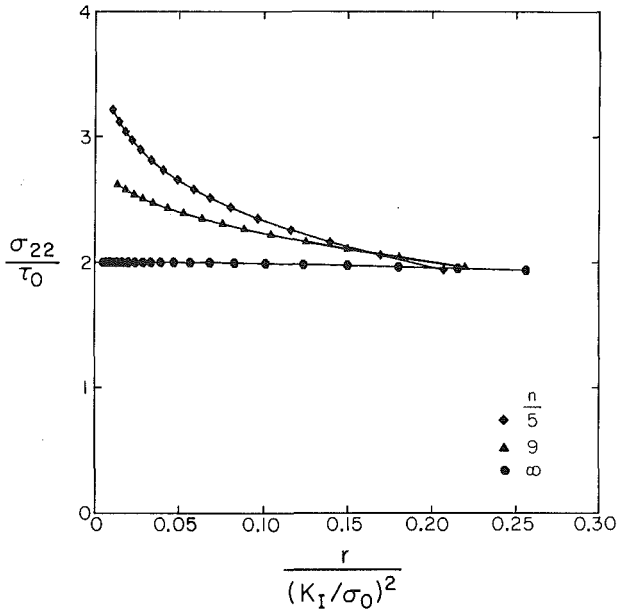


Fig. 4 Radial distribution of opening stress ahead of the moving tip

nature of the governing equations (from elliptic to hyperbolic), in the limit as the perfectly plastic case is approached. Such a behavior can also be observed from the plastic zone shapes given by Dean and Hutchinson (1980) for crack growth under antiplane shear in a linear hardening material. The similarity between the present plane stress plastic zone shapes and their antiplane shear results stems from the presence of an intense deformation zone ahead of the crack tip in both cases.

The active plastic zones of Fig. 1 and the corresponding results obtained by Dean (1983) for steady-state crack growth under plane stress in a linear hardening material have essentially the same features. However, one difference seems to be the absence of the kink in the active plastic zone for the perfectly plastic limit in Dean's solution. Nevertheless, the present analysis is more detailed, because it has a larger ratio of plastic zone to smallest element size as compared with that of Dean's computation. Also, unlike his work, the initial phase of crack growth was modelled in the present investigation.

Radial Distribution of Plastic Strains. The radial distribution of the normalized plastic strain, $\epsilon^p_{22}/\epsilon_0$, with respect to normalized distance, $r/(K_I/\sigma_0)^2$, ahead of the propagating crack tip is shown in Fig. 2 for a material with $n = 9$. Results are presented for various levels of crack growth at fixed applied load, along with the plastic strain distribution ahead of a monotonically loaded stationary crack tip, which was obtained by Narasimhan and Rosakis (1986). As can be seen from this figure, the plastic strain ahead of the moving crack tip converges rapidly during the first few crack growth steps to an invariant distribution. For example, at a distance of $r = 0.013$ $(K_I/\sigma_0)^2$ ahead of the moving tip, the plastic strain dropped by 30 percent during the first five crack growth steps and by 8 percent, 3 percent, and 1.5 percent during the sixth to tenth steps, eleventh to fifteenth steps, and sixteenth to twentieth steps, respectively.

Such rapid convergence was typical of the other hardening case ($n = 5$) as well as the perfectly plastic material. The weaker singularity in the plastic strains near the tip during crack growth, as compared with the stationary problem in Fig. 2, is due to the fact that the crack propagates into material that has already deformed plastically (Rice, 1975). The radial distribution of the plastic strains ahead of the tip at the end of the twentieth release step is shown in Fig. 3 for the two levels of hardening, $n = 5$ and 9.

Radial Distribution of Stresses. The radial distribution of the normalized opening stress, σ_{22}/τ_0, ahead of the moving crack tip is shown in Fig. 4 for $n = 5$ and 9, along with the perfect plasticity solution. As can be seen from this figure, the stress components become more strongly singular with increasing hardening. The perfect plasticity solution for σ_{22} tends to a bounded value of $1.999\tau_0$, as the crack tip is approached along the $\theta = 0$ ray, and is in excellent agreement with the preliminary asymptotic result of Rice (1982). This asymptotic limit was the same as that obtained by the numerical solution near the stationary crack tip.

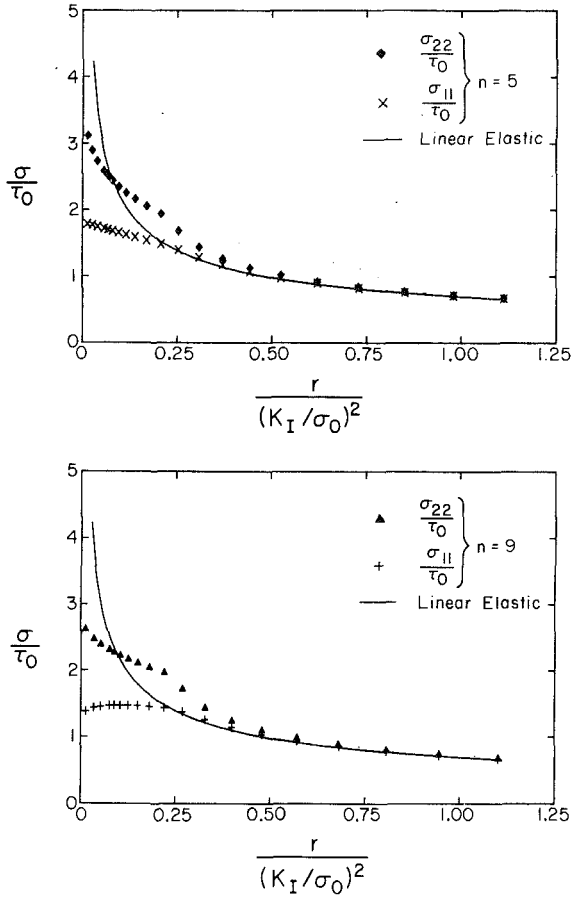


Fig. 5 Comparison of radial stress distribution ahead of the moving tip as given by the K_I field (solid line) with the finite element solution for (*a*) $n = 5$ and (*b*) $n = 9$

The stress variation for the hardening materials in Fig. 4 also differs only slightly from the stationary crack distribution, for moderate to large distances from the tip. For example, at a distance of $r = 0.018\ (K_I/\sigma_0)^2$ ahead of the tip, the ratio of the opening stress for the propagating crack to that for the stationary problem is 3.04/3.13 and 2.58/2.66 for $n = 5$ and 9, respectively. Also, as pointed out by Narasimhan and Rosakis (1986), the stress distribution (Fig. 4) appears to be relatively insensitive to the hardening level for distances from the tip exceeding about $0.15\ (K_I/\sigma_0)^2$.

In order to study the influence of the crack tip plastic zone on the stress field in the surrounding elastic region, the radial stress distribution ahead of the moving crack tip is shown on an expanded scale for $n = 5$ and 9 in Fig. 5. The singular elastic solution (K_I field) is also indicated by the solid line in the figure, for comparison. The distribution of stresses outside the plastic zone is almost identical to the corresponding result obtained for the stationary problem. The σ_{22} stress component obtained from the numerical solution differs strongly (by more than 30 percent) from that given by the K_I field at the elastic-plastic boundary ($r = R_p$). But a rapid transition in the stress distribution takes place immediately outside the plastic zone and the stresses agree closely with those of the K_I field for $r > 1.5\ R_p$.

Near-Tip Angular Distribution of Stresses. The near-tip angular distribution of the normalized polar stress components is shown in Fig. 6 for $n = 5$ and 9 along with the perfect plasticity solution. The centroidal values of stresses in the elements lying on a rectangular contour surrounding the moving crack tip, with an average radius of $0.018\ (K_I/\sigma_0)^2$ (which is within $0.08\ R_p$), have been used to make this plot. The angular variation along the above contour of the Von Mises equivalent stress, $\sigma_{eqv} = (3/2 s_{ij} s_{ij})^{1/2}$, which has been made dimensionless by σ_0, is also shown in the figure.

The assertion made earlier, that no secondary (plastic) reloading was observed (as $\theta \rightarrow 180$ deg) for any level of

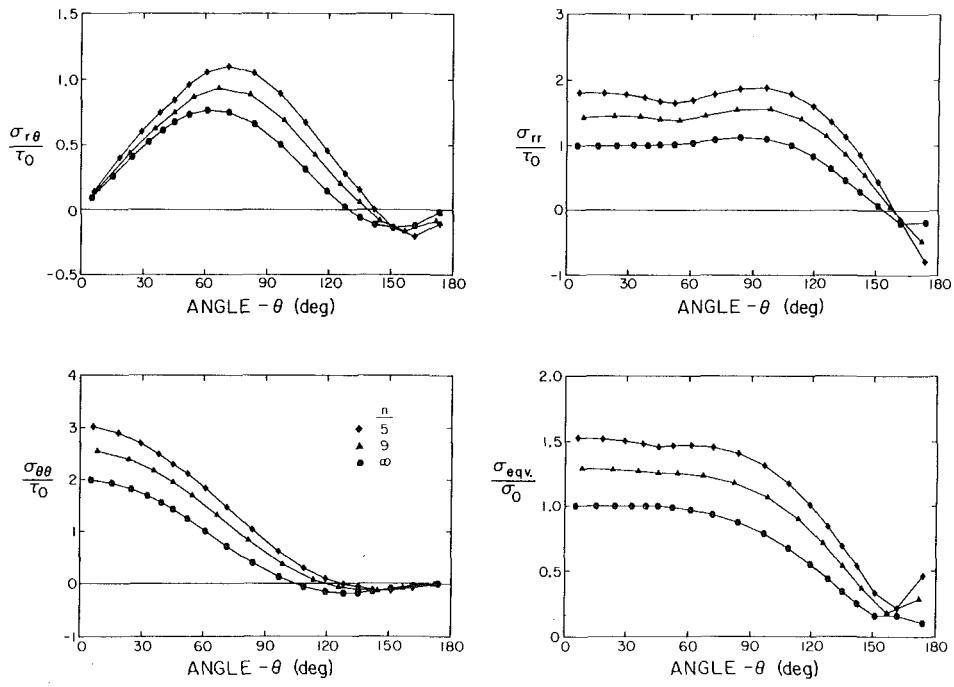


Fig. 6 Near-tip angular distribution of the normalized polar stress components and the Von Mises equivalent stress at a distance of 0.018 $(K_I/\sigma_0)^2$ from the moving tip

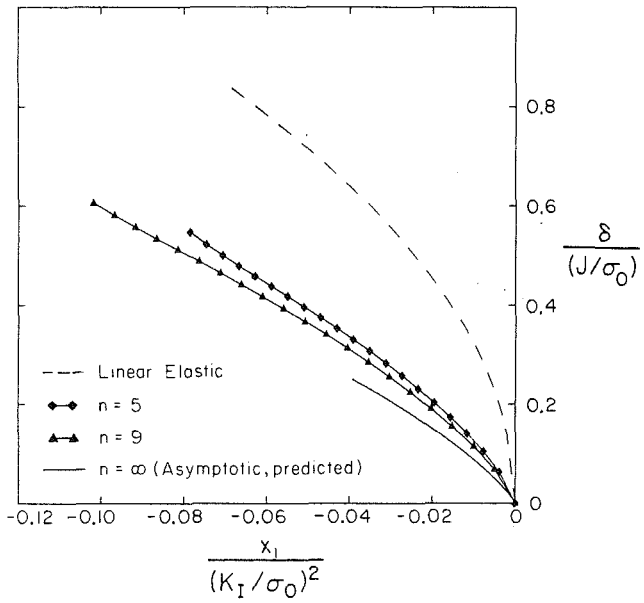


Fig. 7 Numerically obtained crack opening profiles for quasi-static crack growth under fixed applied load for $n = 5$ and 9. The dashed line is th linear elastic asymptotic solution and the solid line is the asymptotic crack displacement for steady-state crack growth in a perfectly plastic solid, as predicted in Part I.

hardening, is confirmed from this figure. Also, elastic unloading occurs for angles θ greater than about 65 deg, 55 deg, and 45 deg for $n = 5$, 9, and ∞, respectively, although it is not obvious from this figure for the hardening cases. The near-tip angular stress variation for the hardening materials appears to be qualitatively similar to the perfectly plastic case. As noted in Part I, the angular stress variation within the active plastic zone for the perfectly plastic case is in very good agreement with the distribution in a centered fan, as predicted by Rice (1982).

Crack Opening Profiles. The normalized crack opening displacement, $\delta/(J/\sigma_0)$, versus normalized distance, $x_1/(K_I/\sigma_0)^2$, along the crack flank is shown in Fig. 7 for the two hardening cases, $n = 5$ and 9, when the crack grows under fixed applied load. This profile was obtained after twenty crack growth steps and was self-similar in normalized form, in the sense that it was almost identical for different levels of crack growth. The crack opening profile for a linear elastic material is also shown by the dashed line in the figure.

The steady-state asymptotic opening profile for a crack growing in a perfectly plastic material, as predicted in Part I, is indicated by a solid line in the figure. This is given by,

$$\frac{\delta}{(J/\sigma_0)} = \beta\eta ln\left(\frac{es}{\eta}\right), \tag{3.1}$$

where

$$\eta = r/(K_I/\sigma_0)^2.$$

In the above equation e is the base of the natural logarithm. The parameters β and s which occur in equation (3.1) were estimated in Part I as 1.70 and 0.60, respectively, from a best-fit to the near-tip crack displacement increment, obtained from the numerical solution for the nonhardening case.

It can be noticed from Fig. 7 that the crack profiles vary considerably with the hardening level. This was also observed by Dean (1983) from his steady-state solution for plane stress crack growth in linear hardening solids. This also appears to be true for the crack profiles obtained under antiplane shear by Dean and Hutchinson (1980). However, the crack profiles under Mode I plane strain show comparatively less variation with the hardening level, at least near the crack tip (e.g., Dean and Hutchinson, 1980). Also, as opposed to the blunted shapes obtained for the stationary problem, the crack opening profiles during growth (Fig. 7) are sharp. This is directly traceable to the permanence of plastic deformation (Rice, 1975).

4 Ductile Crack Growth Criterion

Perfect Plasticity. Rice and Sorensen (1978) and Rice et al. (1980) proposed that a critical opening displacement, $\delta = \delta_C$, should be maintained at a small microstructural distance, r_c, behind the crack tip for continued crack growth. The near-tip crack displacement during continuous stable crack extension (see Section 3, Part I) can be written as

$$\delta = \beta\left(\frac{\sigma_0}{E}\right) r ln\left(\frac{\rho}{r}\right), \quad r \to 0, \tag{4.1}$$

where

$$\rho = Re^{(1+T\alpha/\beta)}. \tag{4.2}$$

In the above equation, $R = sEJ/\sigma_0^2$ for small-scale yielding and T is the tearing modulus. The parameters α, β, and s, which occur in equation (4.2), were estimated in Part I as 0.82, 1.70, and 0.60 respectively. The crack growth criterion stated above requires that the parameter ρ, which uniquely characterizes the near-tip crack profile, be constant for continued crack extension.

Thus, on estimating ρ from J_C and T_0, which are the values of the far-field J and the tearing modulus T at the onset of crack growth, it is possible to obtain the following differential equation for J as function of crack length a (Rice et al., 1980),

$$T = \frac{E}{\sigma_0^2}\frac{dJ(a)}{da} = T_0 - \frac{\beta}{\alpha} ln\left(\frac{J}{J_C}\right). \tag{4.3}$$

By using $J = J_C$ and $a = a_0$ as initial conditions the above equation can be integrated to give,

$$\frac{a-a_0}{(EJ_C/\sigma_0^2)} = \frac{\alpha}{\beta} e^{(\alpha T_0/\beta)}\left[E_i\left\{-\frac{\alpha T_0}{\beta}\right\} - E_i\left\{ln\left(\frac{J}{J_C}\right) - \frac{\alpha T_0}{\beta}\right\}\right], \tag{4.4}$$

where $E_i\{\cdot\}$ is the exponential integral function. It is interesting to note that the mathematical structure of equations (4.3) to (4.4) are similar to the ones deduced by Wnuk (1974) by means of his "final stretch" crack growth criterion, based on a plane-stress Dugdale, line plastic zone model.

A family of plane stress resistance curves generated from equation (4.4) corresponding to several values of T_0 with α and β taken as 0.82 and 1.70, respectively, is shown in Fig. 8. The abscissa of the figure is the extent of crack growth, made dimensionless by the quantity $0.3EJ_C/\sigma_0^2$, which is approximately equal to the maximum plastic zone extent at initiation. The flat portion of the curves corresponds to steady state crack growth when no further increase in externally applied J is required to propagate the crack. Setting $dJ/da = 0$ in (4.3) gives J corresponding to steady-state as

$$J_{SS} = J_C e^{\alpha T_0/\beta}. \tag{4.5}$$

Comparison of Fig. 8 with the corresponding plot for plane strain given by Rice et al. (1980) show that the amount of stable crack extension in plane stress if far more extensive than in plane strain. This is because the ratio α/β in plane stress as computed in the present investigation, is 0.82/1.70, which is about 4.4 times larger than the corresponding ratio of 0.6/5.46 in plane strain (Rice, 1982; Sham, 1983). Thus, for $T_0 = 5$, the ratio J_{SS}/J_C calculated from (4.5) is 11.2 and 1.73 for plane stress and plane strain, respectively.

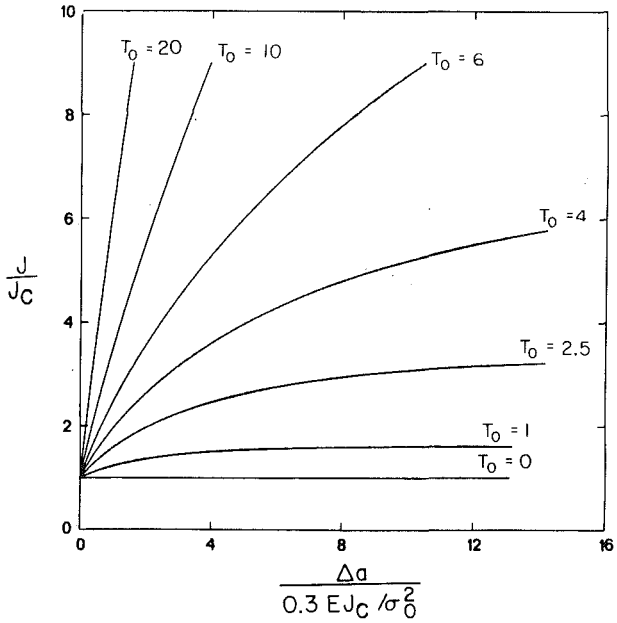


Fig. 8 Predicted normalized plane stress J resistance curves. The abscissa is the amount of crack growth normalized by a quantity which is approximately equal to the maximum plastic zone extent at initiation.

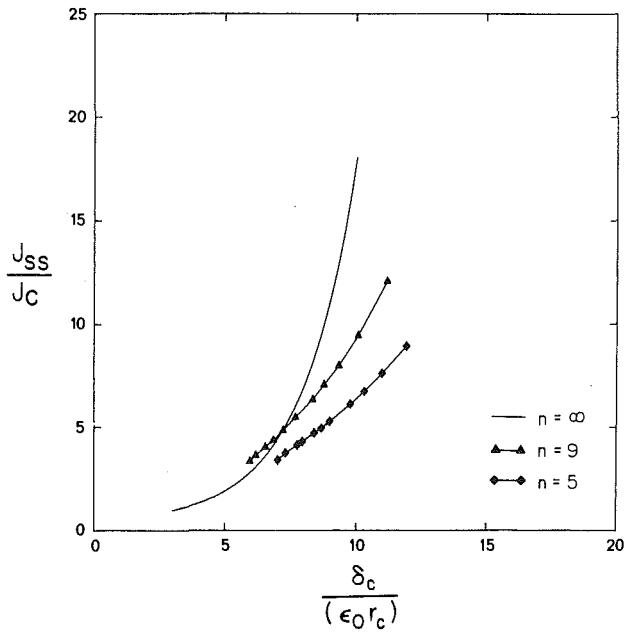


Fig. 10 Influence of hardening on J_{SS}/J_C in Mode I plane stress, as predicted by the critical displacement criterion, for continued crack growth. The solid line is the perfect plasticity result of Part I.

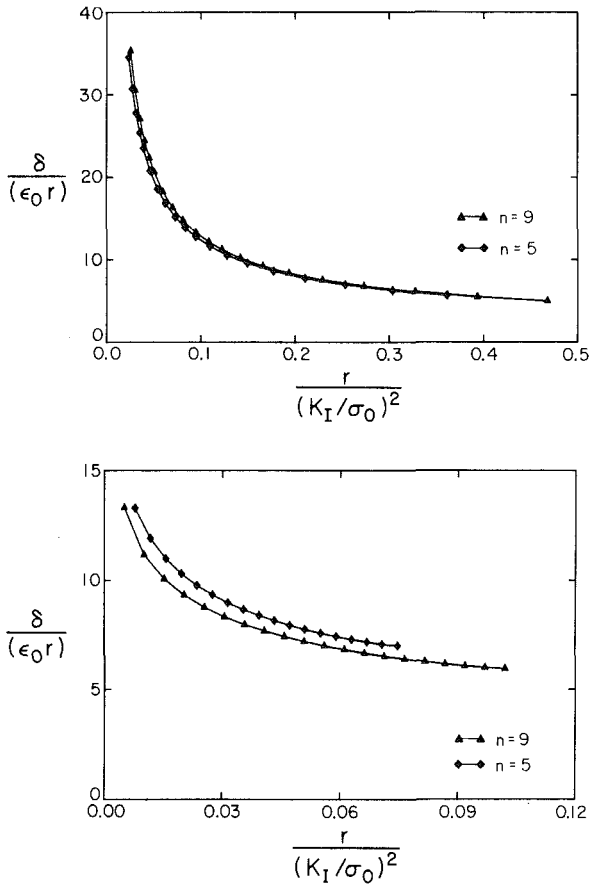


Fig. 9 Variation of $\delta/(\epsilon_0 r)$ with normalized distance along the crack flank for (*a*) the stationary problem (Narasimhan and Rosakis, 1986) and (*b*) quasi-static crack growth under fixed applied load

Hardening Solid. The above crack growth criterion can be used for both initiation and continuation at crack growth (Dean and Hutchinson, 1980) to examine the potential for stable growth from the microstructural viewpoint. To this end, the crack proflies shown in Fig. 7 for crack growth under fixed applied load were taken as steady-state profiles and were used to generate a plot of $\delta/(\epsilon_0 r)$ versus $r/(K_{SS}/\sigma_0)^2$. This is shown in Fig. 9(*b*) for the two cases of hardening, $n = 5$ and 9. The opening displacement for the stationary crack given by Narasimhan and Rosakis (1986) was used similarly to obtain the variation of $\delta/(\epsilon_0 r)$ versus $r/(K_C/\sigma_0)^2$ as shown in Fig. 9(*a*).

For a given value of the microscale parameter $\lambda_m = \delta_C/(\epsilon_0 r_c)$, the value of $r_c/(K_{SS}/\sigma_0)^2$ can be obtained from the abscissa of Fig. 9(*b*) corresponding to steady-state crack growth. The value of $r_c/(K_{IC}/\sigma_0)^2$ may be obtained similarly from Fig. 9(*a*) for initiation of crack growth. These two quantities can be used to compute the ratio of $J_{SS}/J_C = (K_{SS}/K_C)^2$, corresponding to the chosen value of the microscale parameter λ_m. The variation of J_{SS}/J_C versus $\delta_C/(\epsilon_0 r_c)$, calculated as indicated above for $n = 5$ and 9, is shown in Fig. 10. On comparing Figs. 9(*a*) and 9(*b*), it can be seen that the influence of hardening on the relationship between J_{SS}/J_C and $\delta_C/(\epsilon_0 r_c)$ arises mainly due to the results in Fig. 9(*b*), corresponding to steady-state crack growth. The effect of hardening on the variation of $\delta/(\epsilon_0 r)$ with respect to $r/(K_C/\sigma_0)^2$ at initiation is not so significant, as can be seen from Fig. 9(*a*).

For comparison purposes, the variation of J_{SS}/J_C with respect to $\delta_C/(\epsilon_0 r_c)$ for the elastic-perfectly plastic material, is also shown in Fig. 10 by the solid line. It can be shown from equations (4.1)–(4.3), along with the fact that $\delta_C = \alpha J_C/\sigma_0$ for initiation, that this relation is given by,

$$\frac{J_{SS}}{J_C} = \frac{\alpha}{s\lambda_m} e^{(\lambda_m/\beta - 1)}, \tag{4.6}$$

where $\lambda_m = \delta_C/(\epsilon_0 r_c)$. It can be seen from Fig. 10 that in the range $\mu_m > 8.0$, the ratio J_{SS}/J_C may increase significantly with a decrease in hardening. For example, corresponding to a value of $\lambda_m = 9.5$, the ratio J_{SS}/J_C is 5.8, 8.3, and 14.1 for $n = 5$, 9, and ∞, respectively. Thus, the potential for stable crack growth may be grossly overestimated by a calculation based on the perfect plasticity idealization, when the material actually possesses some hardening. Hence, predictions about the extent of stable crack growth based on the perfectly plastic

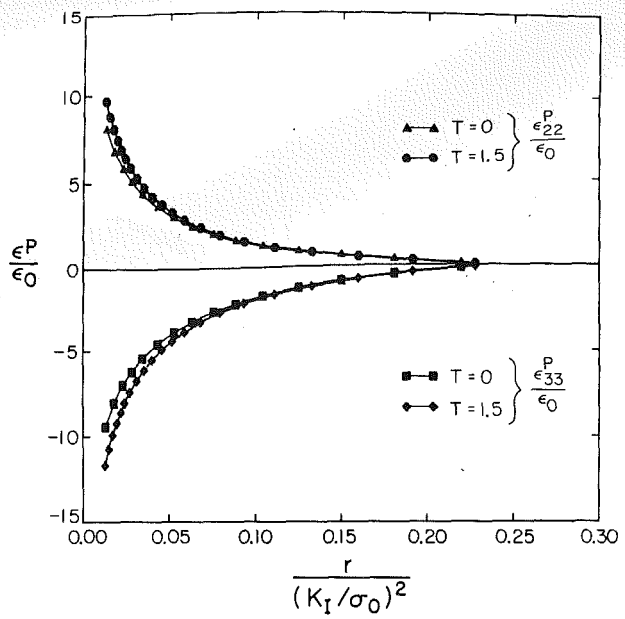


Fig. 11 Comparison of the radial distribution of plastic strains ahead of the tip for the two crack growth histories that were simulated for n = 9

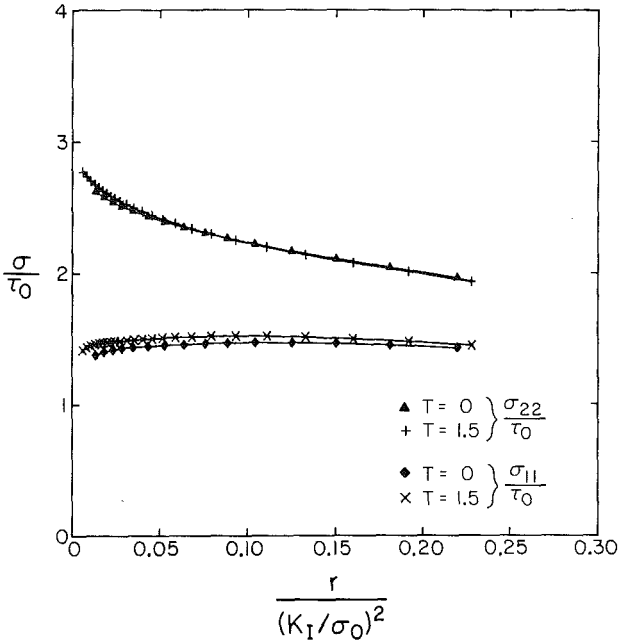


Fig. 12 Radial stress distribution ahead of the tip for n = 9 for the two crack growth histories

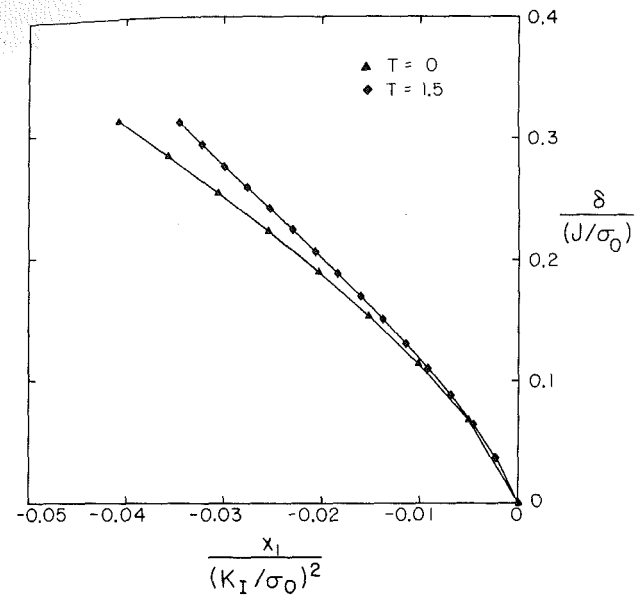


Fig. 13 Effect of increase in applied load at T = 1.5 on the near-tip crack displacement for n = 9

model may be unconservative for a hardening material when the microscale parameter exceeds a value of about 8. A qualitatively similar conclusion was reached in antiplane shear and Mode I plane strain as well, by Dean and Hutchinson (1980).

In the light of the above observation, one is compelled to examine the effects of kinematic hardening and corner formation on the yield surface, which may occur during the nonproportional loading experienced by a material particle near the crack tip. It is not clear to what extent these factors will affect the potential for stable crack growth under plane stress conditions. Dean and Hutchinson (1980) found that the influence of corner formation was not as significant as strain hardening from their numerical results for antiplane shear crack growth. However, Lam and McMeeking (1984) observed that both corner formation and kinematic hardening further reduced the potential for stable crack growth in Mode I plane strain. Thus, in this sense, even the results based on a smooth yield surface with isotropic hardening may be unconservative. It is suggested that such effects should be investigated in Mode I plane stress.

5 Comparison of Results for the Two Crack Growth Histories

In order to study the influence of increase in applied load, as would be observed in an experiment during the initial phase of stable crack extension, a crack growth history at a constant value of T = 1.5 was also simulated in this work. Only the material with n = 9 was considered in this investigation.

The active plastic zones obtained for this crack growth history compared very closely with that shown in Fig. 1, both in shape and size. During the first few crack growth steps, the active plastic zone assumed the sharpened shape of Fig. 1, which did not change with subsequent crack advance. The values of θ_p and R_p were about 55 deg and 0.24 $(K_I/\sigma_0)^2$ as reported earlier, based on the first crack growth history (at fixed applied load).

The plastic strains ahead of the moving crack tip exhibited a tendency to converge rapidly to an invariant distribution during the first few crack growth steps as in the earlier analysis (Fig. 2). The normalized plastic strains ahead of the tip at the end of the fifteenth crack growth step under T = 1.5 is shown in Fig. 11 and is compared with the result given in Fig. 3 for crack growth at T = 0. As expected, the plastic strains for T = 1.5 are slightly higher due to the influence of increase in applied load with crack growth.

The radial distribution of stresses ahead of the propagating crack tip for the two histories is shown in Fig. 12 in the nondimensional form, $\sigma_{\alpha\beta}/\tau_0$ versus $r/(K_I/\sigma_0)^2$. The effect of the increase in applied load on the stress field seems to be less significant than that on the deformation field. Also, the near-tip angular stress distribution for the two histories were almost identical. Finally, the nondimensional crack opening displacement, $\delta/(J/\sigma_0)$, as a function of position on the crack flank, $x_1/(K_I/\sigma_0)^2$, is shown in Fig. 13 for T = 0 and 1.5. Due to the increase in applied load, the crack opening displacement for T = 1.5 is higher than that for T = 0.

Acknowledgment

The authors would like to express their gratitude to Professor J. W. Knowles for his valuable advice and encouragement. This investigation was supported by the Office of Naval Research through ONR contract #N00014-85-K-0596. The computations were performed using the Supercomputer at Boeing Computer Services, Seattle. This was made possible through NSF contract #MEA-8307785. The above contracts and the facilities provided by Boeing Computer Services are gratefully acknowledged.

References

Dean, R. H., 1983, "Elastic-Plastic Steady Crack Growth in Plane Stress," *Elastic-Plastic Fracture: Second Symposium, Volume I-Inelastic Crack Analysis*, ASTM STP 803, pp. I-39–I-51.

Dean, R. H., and Hutchinson, J. W., 1980, "Quasi-Static Steady Crack Growth in Small-Scale Yielding," *Fracture Mechanics: Twelfth Conference*, ASTM STP 700, pp. 385–400.

Lam, P. S. and McMeeking, R. M., 1984, "Analysis of Steady Quasistatic Crack Growth in Plane Strain Tension in Elastic-Plastic Materials with Non-Isotropic Hardening," *Journal of Mechanics and Physics of Solids*, Vol. 32, pp. 395–414.

Narasimhan, R., and Rosakis, A. J., 1986, "A Finite Element Analysis of Small-Scale Yielding near a Stationary Crack under Plane Stress," Caltech Report SM 86-21, Pasadena, CA, *Journal of Mechanics and Physics of Solids*, to appear, 1988.

Narasimhan, R., Rosakis, A. J., and Hall, J. F., 1986, "A Finite Element Study of Stable Crack Growth under Plane Stress Conditions in Elastic-Perfectly Plastic Solids," Caltech Report SM 86-22, Pasadena, CA.

Rice, J. R., 1975, "Elastic-Plastic Models for Stable Crack Growth," *Mechanics and Mechanisms of Crack Growth*, May, N. J., ed., British Steel Corp. Physical Metallurgy Centre Publication, Sheffield, England, pp. 14–39.

Rice, J. R., 1982, "Elastic-Plastic Crack Growth," *Mechanics of Solids*, Hopkins, H. G., and Sewell, M. J., eds., Pergamon Press, Oxford, pp. 539–562.

Rice, J. R., and Sorensen, E. P., 1978, "Continuing Crack-Tip Deformation and Fracture for Plane Strain Crack Growth in Elastic-Plastic Solids," *Journal of Mechanics and Physics of Solids*, Vol. 26, pp. 163–186.

Rice, J. R., Drugan, W. J., and Sham, T. L., 1980, "Elastic-Plastic Analysis of Growing Cracks," *Fracture Mechanics: Twelfth Conference*, ASTM STP 700, pp. 189–221.

Sham, T. L., 1983, "A Finite Element Study of the Asymptotic Near-Tip Fields for Mode I Plane Strain Cracks Growing Stably in Elastic-Ideally Plastic Solids," *Elastic-Plastic Fracture: Second Symposium, Volume I—Inelastic Crack Analysis*, ASTM STP 803, pp. 52–79.

Wnuk, M. P., 1974, "Quasi-static Extension of a Tensile Crack Contained in a Viscoelastic-Plastic Solid," ASME JOURNAL OF APPLIED MECHANICS, Transactions of ASME, Ser. E, Vol. 41, pp. 231–242.

Transient Thermoelastic Fields in a Transversely Isotropic Infinite Solid With a Penny-Shaped Crack

N. Noda
Professor,
Department of Mechanical Engineering,
Shizuoka University,
Hamamatsu, Shizuoka 432, Japan

F. Ashida
Research Associate,
Department of Mechanical Engineering,
Tsuyama College of Technology,
Tsuyama, Okayama 708, Japan

The present paper deals with a transient thermoelastic problem for an axisymmetric transversely isotropic infinite solid with a penny-shaped crack. A finite difference formulation based on the time variable alone is proposed to solve a three-dimensional transient heat conduction equation in an orthotropic medium. Using this formulation, the heat conduction equation reduces to a differential equation with respect to the spatial variables. This formulation is applied to attack the transient thermoelastic problem for an axisymmetric transversely isotropic infinite solid containing a penny-shaped crack subjected to heat absorption and heat exchange through the crack surface. Thus, the thermal stress field is analyzed by means of the transversely isotropic potential function method.

1 Introduction

In recent years, a considerable effort has been devoted to calculations of thermal stresses associated with various types of cracks. Most of these calculations, however, deal with isotropic media. Thermoelastic crack problems in anisotropic media, on the other hand, have been treated only in a limited number of papers. The plane-thermoelastic problems in orthotropic media with a crack have been considered by Atkinson and Clements (1977), Ghosh (1977), Clements and Tauchert (1979), Sumi (1981, 1982), and Clements (1983). Maiti and Misra (1976), Murata and Atsumi (1977), and Tsai (1983a, 1983b) have analyzed thermal stresses in a transversely isotropic medium containing a penny-shaped crack. But all of these are steady thermoelastic problems. The transient thermoelastic problems in anisotropic media with a crack have not been considered so far. Even if the analytical solution in the Laplace transform domain may be obtained, the inverse Laplace transform must be carried out numerically because of difficulty in evaluating the analytical inverse Laplace transform. Such a procedure may require intricate numerical calculations.

In this paper, a finite difference method with respect to the time variable only is introduced to analyze the transient thermoelastic problem for an axisymmetric transversely isotropic medium containing a penny-shaped crack. In the first stage of the analysis, we shall apply this finite difference method to a mixed boundary value problem for three-dimensional transient heat conduction in an orthotropic solid. Next, we shall discuss the problem of transient thermal stress for an axisymmetric transversely isotropic infinite solid containing a penny-shaped crack subjected to heat absorption and heat exchange through the crack surface. In this method, the temperature distribution can be determined by solving a dual-integral equation. A subsequently thermal stress field is analyzed by means of the transversely isotropic potential function method (Takeuti and Noda, 1978).

Numerical examples of the temperature, axial displacement, and the stress intensity factor are illustrated for a transversely isotropic graphite and are compared with those derived under isotropic conditions. The effects of anisotropies of the transversely isotropic material constants on the stress intensity factor are shown in figures. The numerical results due to this procedure, in the case of constant heat flux through the crack surface, closely agree with the exact solutions.

2 General Discussion on Temperature Field in an Orthotropic Solid

Consider a heat conduction problem in an orthotropic body with heat generation. The governing equation for the temperature T at time t and on the position $\bar{P}(x_1, x_2, x_3)$ is

$$\Delta_1 T = T_{,t}/\kappa_1 - W(\bar{P},t)/\lambda_1 \tag{1}$$

where $\Delta_1 = \frac{\partial^2}{\partial x_1^2} + \frac{\lambda_2}{\lambda_1}\frac{\partial^2}{\partial x_2^2} + \frac{\lambda_3}{\lambda_1}\frac{\partial^2}{\partial x_3^2}$, $\kappa_1 = \lambda_1/\bar{c}_p\gamma$ is thermal diffusivity, $W(\bar{P}, t)$ is the heat generation per unit volume and unit time, γ is the density, $\bar{c}_p$ is the specific heat, and λ_i is the thermal conductivity in the x_i direction. Partial differentiation is indicated with a comma followed by the variables.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, 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 15, 1985; final revision April 1, 1987.

Paper No. 88-WA/APM-2.

The initial and boundary conditions are

$$T = \Psi(\bar{P}) \quad \text{at} \quad t = 0 \tag{2}$$

and

$$-T_{,n} + q(\bar{P},t)/\lambda_n = h_n\{T - \Theta(\bar{P},t)\}$$
$$\text{on the boundary surface} \tag{3}$$

where $\Psi(\bar{P})$ is the initial temperature, $\Theta(\bar{P}, t)$ is the temperature of the surrounding medium, $q(\bar{P}, t)$ is the heat generation on the boundary surface, $h_n = \bar{\alpha}_n/\lambda_n$ is the coefficient of relative surface heat transfer (i.e., ratio of convective coefficient to conductivity), λ_n is the thermal conductivity, and n is the normal direction on the boundary surface.

Usually, it is not easy to obtain the analytical expression for equation (1) under the initial and boundary conditions given by equations (2) and (3). In what follows, we shall introduce a finite difference formulation in terms of a single independent variable, namely, the time variable. Equation (1) along with equation (2) reduces to

$$\Delta_1 T_j = (1-\delta_{1j})(T_j - T_{j-1})/\kappa_1 \Delta t_j + \delta_{1j}\{T_j - \Psi(\bar{P})\}/\kappa_1 \Delta t_j - W(\bar{P},t_j)/\lambda_1 \tag{4}$$

where δ_{ij} is Kronecker's delta, while T_j is the temperature at time t_j with

$$t_j = \sum_{k=1}^{j} \Delta t_k, \text{ and } \Delta t_j = t_j - t_{j-1}.$$

Let us consider the case in which all time differences Δt_j (j = 1, 2, . .) have different values. (Note that the solution of equation (4) depends on the values of the time differences Δt_j.)

The general solution of equation (4) may be expressed by

$$T_j = \sum_{k=1}^{j} w_{jk} T_{kc} + T_{jp} \tag{5}$$

where

$$w_{jk} = \delta_{jk} + (1-\delta_{jk}) \prod_{m=k+1}^{j} \Delta t_k/(\Delta t_k - \Delta t_m) \tag{6}$$

and T_{jc} and T_{jp} are complemental and particular solutions of equation (4), respectively. These solutions are governed by

$$\Delta_1 T_{jc} - T_{jc}/\kappa_1 \Delta t_j = 0 \tag{7}$$

$$\Delta_1 T_{jp} - T_{jp}/\kappa_1 \Delta t_j = -\delta_{1j}\Psi(\bar{P})/\kappa_1 \Delta t_j - (1-\delta_{1j}) T_{(j-1)p}/\kappa_1 \Delta t_j - W(\bar{P},t_j)/\lambda_1 \tag{8}$$

The boundary condition (3) reduces to

$$T_{jc,n} + h_n T_{jc} = D_j(\bar{P})$$
$$\text{on the boundary surface} \tag{9}$$

where

$$D_j(\bar{P}) = -(T_{jp,n} + h_n T_{jp}) + h_n \Theta(\bar{P},t_j) + q(\bar{P},t_j)/\lambda_n - (1-\delta_{1j}) \sum_{k=1}^{j-1} w_{jk} D_k(\bar{P}) \tag{10}$$

Thus, we can obtain the temperature T_j at time t_j by solving the governing equations (7) and (8) under the boundary condition (9).

3 Thermal Stresses in a Transversely Isotropic Body With a Penny-Shaped Crack

3.1 Temperature Field. Let us consider a heat conduction problem for an axisymmetric transversely isotropic infinite body containing a penny-shaped crack with radius a, using the cylindrical coordinate system (r, z). It is assumed that the body is initially at the same uniform temperature ($T = 0$) and that the crack is to be opened by the agency of heat absorption.

The axisymmetric transient heat conduction equation in transverse isotropy without heat generation can be expressed by

$$\Delta_2 T + \lambda^2 T_{,zz} = T_{,t}/\kappa \tag{11}$$

where

$$\Delta_2 = \partial^2/\partial r^2 + r^{-1}\partial/\partial r, \quad \kappa = \lambda_r/\bar{c}_p\gamma, \quad \lambda^2 = \lambda_z/\lambda_r$$

The initial and boundary conditions can be expressed by

$$T = 0 \qquad \text{at } t = 0 \tag{12}$$

$$T_{,z} = [hT - \{-Q_0 f(r) g(t)/\lambda_z\}]H(a-r) \qquad \text{on } z = 0 \tag{13}$$

where $H(r)$ is Heaviside's unit step function, while $-Q_0$ is the heat absorption rate which is assumed to be constant.

Equations (11)–(13) reduce to the following equations by use of the finite difference formulation with respect to the time variable:

$$\Delta_2 T_j + \lambda^2 T_{j,zz} = \{T_j - (1-\delta_{j1}) T_{j-1}\}/\kappa \Delta t_j \tag{14}$$

$$\left.\begin{aligned} T_{j,z} - hT_j &= Q_0 f(r) g(t)/\lambda_z, && \text{for } 0 < r < a \\ T_{j,z} &= 0, && \text{for } a < r \end{aligned}\right\} \text{on } z = 0 \tag{15}$$

The general solution of equation (14) may be expressed by

$$T_j = \sum_{k=1}^{j} w_{jk} T_{kc} \tag{16}$$

in which w_{jk} is given by equation (6), and T_{kc} is

$$T_{kc} = \int_0^\infty A_k p J_0(pr) \exp(-s_k z/\lambda)\, dp \tag{17}$$

where

$$s_k = \sqrt{p^2 + (\kappa \Delta t_k)^{-1}}$$

Substitution of equation (16) into equations (15) yields to a dual-integral equation:

$$\left.\begin{aligned} &\int_0^\infty p A_j (s_j + \lambda h) J_0(pr)\, dp = -\lambda D_j f(r), && \text{for } 0 \leqq r < a \\ &\int_0^\infty p s_j A_j J_0(pr)\, dp = 0, && \text{for } a < r \end{aligned}\right\} \tag{18}$$

where

$$D_j = Q_0 g(t_j)/\lambda_z - (1-\delta_{1j}) \sum_{k=1}^{j-1} w_{jk} D_k$$

This dual-integral equation reduces to an infinite system of simultaneous algebraic equations for the unknown coefficients $\bar{A}_{jn}$:

$$\sum_{n=0}^{\infty} \bar{A}_{jn} \int_0^\infty (1 + \lambda h/s_j) Z_{np} J_m^2(pa/2)\, dp = -\lambda D_j f_m, \quad (m = 0, 1, 2, \ldots) \tag{19}$$

where

$$A_j = \sum_{n=0}^{\infty} \bar{A}_{jn} Z_{np}/p s_j$$

$$Z_{np} = J_{2n+3/2}(pa)/\sqrt{p},$$

$$f(r) = f_0 + 2\sum_{m=1}^{\infty} f_m \cos(m\theta),$$

$$f_m = \int_0^{\pi} f\{a\sqrt{2(1-\cos\theta)}/2\}\cos(m\theta)\,d\theta/\pi$$

Then, the temperature T_j is given by

$$T_j = \sum_{k=1}^{j} w_{jk} \sum_{n=0}^{\infty} \bar{A}_{kn} \int_0^{\infty} s_k^{-1} Z_{np} J_0(pr)\exp(-s_k z/\lambda)\,dp \tag{20}$$

In a similar fashion, the steady state temperature T_∞ may be expressed as

$$T_\infty = \sum_{n=0}^{\infty} \bar{A}_{\infty n} \int_0^{\infty} p^{-1} Z_{np} J_0(pr)\exp(-pz/\lambda)\,dp \tag{21}$$

where the coefficients $\bar{A}_{\infty n}$ are the solutions of infinite system of simultaneous algebraic equations:

$$\sum_{n=0}^{\infty} \bar{A}_{\infty n} \int_0^{\infty} (1+\lambda h/p) Z_{np} J_m^2(pa/2)\,dp = -\lambda Q_0 g(\infty) f_m/\lambda_z, \quad (m=0,1,2,\ldots) \tag{22}$$

3.2 Thermal Stresses. Let us consider a three-dimensional transient thermal stress problem for an axisymmetric transversely isotropic infinite solid containing a penny-shaped crack. The stress-strain relations are

$$\left.\begin{aligned} \sigma_{rr} &= c_{11}\epsilon_{rr} + c_{12}\epsilon_{\theta\theta} + c_{13}\epsilon_{zz} - \beta_1 T \\ \sigma_{\theta\theta} &= c_{12}\epsilon_{rr} + c_{11}\epsilon_{\theta\theta} + c_{13}\epsilon_{zz} - \beta_1 T \\ \sigma_{zz} &= c_{13}\epsilon_{rr} + c_{13}\epsilon_{\theta\theta} + c_{33}\epsilon_{zz} - \beta_3 T \\ \sigma_{rz} &= c_{44}\epsilon_{rz} \end{aligned}\right\} \tag{23}$$

where c_{ij} and β_i are material constants of transverse isotropy.

Now, we shall introduce the potential functions Ω, χ, ϕ_1 and ϕ_2 defined by

$$\left.\begin{aligned} u_r &= (\phi_1+\phi_2)_{,r} + \Omega_{,r} \\ u_z &= (k_1\phi_1 + k_2\phi_2)_{,z} + \chi_{,z} \end{aligned}\right\} \tag{24}$$

As we substitute equations (24) into the corresponding displacement equations, we find that the potential functions must satisfy the following equations (Takeuti and Noda, 1978):

$$\left.\begin{aligned} c_{11}\Delta_2\Omega + c_{44}\Omega_{,zz} + (c_{13}+c_{44})\chi_{,zz} &= \beta_1 T \\ (c_{13}+c_{44})\Delta_2\Omega + c_{44}\Delta_2\chi + c_{33}\chi_{,zz} &= \beta_3 T \end{aligned}\right\} \tag{25}$$

$$\Delta_2\phi_i + \mu_i\phi_{i,zz} = 0, \quad (i=1,2) \tag{26}$$

where

$$c_{11}c_{44}\mu^2 + (2c_{13}c_{44} + c_{13}^2 - c_{11}c_{33})\mu + c_{33}c_{44} = 0$$

k_1 and k_2 are

$$k_i = (c_{11}\mu_i - c_{44})/(c_{13}+c_{44}), \quad (i=1,2)$$

Substituting equations (24) into equations (23), the thermal stress components can be represented by the potential functions as

$$\left.\begin{aligned} \sigma_{rr} &= \sum_{i=1}^{2}\{(c_{11}-c_{12})\phi_{i,rr} - (c_{12}\mu_i - c_{13}k_i)\phi_{i,zz}\} + c_{11}\Omega_{,rr} + c_{12}\Omega_{,r}/r + c_{13}\chi_{,zz} - \beta_1 T \\ \sigma_{\theta\theta} &= -\sum_{i=1}^{2}\{(c_{11}-c_{12})\phi_{i,rr} + (c_{11}\mu_i - c_{13}k_i)\phi_{i,zz}\} + c_{12}\Omega_{,rr} + c_{11}\Omega_{,r}/r + c_{13}\chi_{,zz} - \beta_1 T \\ \sigma_{zz} &= -\sum_{i=1}^{2}(c_{13}\mu_i - c_{33}k_i)\phi_{i,zz} + c_{13}\Delta_2\Omega + c_{33}\chi_{,zz} - \beta_3 T \\ \sigma_{rz} &= c_{44}\left\{\sum_{i=1}^{2}(1+k_i)\phi_i + \Omega + \chi\right\}_{,rz} \end{aligned}\right\} \tag{27}$$

Now, we get admissible solutions for Ω, χ, ϕ_1 and ϕ_2 as follows:

$$\left.\begin{aligned} \Omega_j &= \sum_{k=1}^{j} w_{jk} \sum_{n=0}^{\infty} \bar{A}_{kn} \int_0^{\infty} p^{-2} s_k^{-1} F_k Z_{np} \times J_0(pr)\exp(-s_k z/\lambda)\,dp \\ \chi_j &= \sum_{k=1}^{j} w_{jk} \sum_{n=0}^{\infty} \bar{A}_{kn} \int_0^{\infty} p^{-2} s_k^{-1} G_k Z_{np} \times J_0(pr)\exp(-s_k z/\lambda)\,dp \end{aligned}\right\} \tag{28}$$

$$\phi_{ij} = \int_0^{\infty} p^{-2} H_{ij} J_0(pr)\exp(-pz/\sqrt{\mu_i})\,dp, \quad (i=1,2) \tag{29}$$

where F_k, G_k, H_{1j}, and H_{2j} are unknown coefficients.

Substituting equations (28) into equations (25), F_k and G_k are determined as follows:

$$\left.\begin{aligned} F_k &= \{\beta_1(c_{33}s_k^2 - c_{44}p^2\lambda^2) - \beta_3(c_{13}+c_{44})s_k^2\}\cdot\overline{FG} \\ G_k &= \{\beta_1(c_{13}+c_{44})p^2\lambda^2 - \beta_3(c_{11}p^2\lambda^2 - c_{44}s_k^2)\}\cdot\overline{FG} \end{aligned}\right\} \tag{30}$$

where

$$\overline{FG} = p^2\lambda^2/\{(c_{13}+c_{44})^2 p^2 s_k^2\lambda^2 - (c_{11}p^2\lambda^2 - c_{44}s_k^2)(c_{33}s_k^2 - c_{44}p^2\lambda^2)\}$$

The substitution of equations (28) and (29) into equations (24) and (27) yields the displacement and thermal stress components as

$$u_{zj} = -\int_0^{\infty} p^{-1} J_0(pr)\left\{\sum_{i=1}^{2} H_{ij}k_i\exp(-pz/\sqrt{\mu_i})/\sqrt{\mu_i} + \sum_{k=1}^{j} w_{jk}\sum_{n=0}^{\infty}\bar{A}_{kn}p^{-1}G_k Z_{np}\exp(-s_k z/\lambda)/\lambda\right\}dp \tag{31}$$

$$\sigma_{zzj} = -\int_0^{\infty} J_0(pr)\left\{\sum_{i=1}^{2} H_{ij}(c_{13} - c_{33}k_i/\mu_i)\exp(-pz/\sqrt{\mu_i})\right.$$

$$+\sum_{k=1}^{j} w_{jk}\sum_{n=0}^{\infty}\bar{A}_{kn}(F_k c_{13}-G_k c_{33}s_k^2/p^2\lambda^2+\beta_3)Z_{np}$$

$$\times\exp(-s_k z/\lambda)/s_k\Big\}dp \tag{32}$$

$$\sigma_{rzj}=c_{44}\int_0^{\infty}J_1(pr)\Big\{\sum_{i=1}^{2}H_{ij}(1+k_i)\exp(-pz/\sqrt{\mu_i})/\sqrt{\mu_i}$$

$$+\sum_{k=1}^{j} w_{jk}\sum_{n=0}^{\infty}\bar{A}_{kn}(F_k+G_k)Z_{np}$$

$$\times\exp(-s_k z/\lambda)/p\lambda\Big\}dp \tag{33}$$

The boundary conditions of the displacement and thermal stresses on the crack surface may be given by

$$\sigma_{rzj}=0 \qquad \text{on } z=0 \tag{34}$$

$$\left.\begin{aligned}&\sigma_{zzj}=0, \quad \text{for } 0\leqq r<a\\ &u_{zj}=0, \quad \text{for } a<r\end{aligned}\right\}\text{on } z=0 \tag{35}$$

The substitution of equation (33) into equation (34) yields

$$H_{2j}=-\sqrt{\mu_2}\Big\{(1+k_1)H_{1j}/\sqrt{\mu_1}$$

$$-\sum_{k=1}^{j} w_{jk}\sum_{n=0}^{\infty}\bar{A}_{kn}(F_k+G_k)Z_{np}/p\lambda\Big\}/(1+k_2) \tag{36}$$

Substituting equations (31) and (32) into equations (35), the following dual-integral equation is obtained:

$$\left.\begin{aligned}&\int_0^{\infty}p^2P_jJ_0(pr)dp=\int_0^{\infty}U_{jp}J_0(pr)dp, \quad \text{for } 0\leqq r<a\\ &\int_0^{\infty}pP_jJ_0(pr)dp=0, \qquad\qquad \text{for } a<r\end{aligned}\right\} \tag{37}$$

where

$$P_j=\{(c_{13}-c_{33}k_1/\mu_1)\sqrt{\mu_1}(1+k_2)$$

$$-(c_{13}-c_{33}k_2/\mu_2)\sqrt{\mu_2}(1+k_1)\}\Big\{(k_1-k_2)H_{1j}/\sqrt{\mu_1}$$

$$-\sum_{k=1}^{j} w_{jk}\sum_{n=0}^{\infty}\bar{A}_{kn}(F_k k_2-G_k)Z_{np}/p\lambda\Big\}$$

$$/p^2(1+k_2)(k_1-k_2) \tag{38}$$

$$U_{jp}=-\sum_{k=1}^{j} w_{jk}\sum_{n=0}^{\infty}\bar{A}_{kn}\bar{q}_{kn}(s_k,p) \tag{39}$$

and

$$\bar{q}_{kn}(s_k,p)=[[F_k[c_{13}+s_k\{k_2\sqrt{\mu_1}(c_{13}-c_{33}k_1/\mu_1)$$

$$-k_1\sqrt{\mu_2}(c_{13}-c_{33}k_2/\mu_2)\}/p\lambda(k_1-k_2)]-G_k[c_{33}s_k^2/p^2\lambda^2$$

$$+s_k\{\sqrt{\mu_1}(c_{13}-c_{33}k_1/\mu_1)-\sqrt{\mu_2}(c_{13}-c_{33}k_2/\mu_2)\}$$

$$/p\lambda(k_1-k_2)]+\beta_3]]Z_{np}/s_k \tag{40}$$

The solution of this dual-integral equation is given by (Sneddon, 1966)

$$P_j=2\int_0^{a}\sin(p\eta)\int_0^{\eta}\delta\{\eta^2-\delta^2\}^{-1/2}$$

$$\int_0^{\infty}U_{j\xi}J_0(\xi\delta)d\xi d\delta d\eta/\pi p \tag{41}$$

The stress intensity factor K-Mode I is defined by

$$K_{1j}=\lim_{r\to a}\sqrt{2\pi(r-a)}[\sigma_{zzj}]_{z=0}$$

$$=2\int_0^{\infty}p^{-1}\sin(pa)U_{jp}dp/\sqrt{\pi a} \tag{42}$$

Replacing U_{jp} in equation (39) by the following equation,

$$U_{\infty p}=-\sum_{n=0}^{\infty}A_{\infty n}\bar{q}_{\infty n}(p,p) \tag{43}$$

we can obtain the stress intensity factor $K_{1\infty}$ for the steady case.

3.3 Exact Solution for $h=0$. If the relative heat transfer coefficient h is equal to zero, the exact solution of temperature can be obtained by use of the Laplace and Hankel transforms. The stress intensity factor may be expressed by means of the same transversely isotropic potential function method as follows:

$$T=\int_0^{\infty}J_0(\delta r)\int_0^{\infty}\bar{T}_{\delta\omega}\cos(\omega z/\lambda)d\omega d\delta \tag{44}$$

Table 1 Comparison of $(\widehat{T})_{\rho=\zeta=0}$ by this method with exact one for $B_i = 0$ and Case I of heat absorption.

$\left(\sum_{n=0}^{\infty}=\sum_{n=0}^{N}\text{, E.S.: Exact Solution}\right)$

N \ t'	0.01	0.1	1	10	100	∞
5	- 0.107	- 0.367	- 0.770	- 0.973	- 1.042	- 1.073
10	- 0.121	- 0.371	- 0.770	- 0.971	- 1.039	- 1.070
15	- 0.118	- 0.371	- 0.770	- 0.971	- 1.039	- 1.070
E.S.	- 0.120	- 0.377	- 0.779	- 0.974	- 1.039	- 1.069

Table 2 Comparison of $(\widehat{u_\zeta})_{\rho=\zeta=0}$ by this method with exact one for $B_i = 0$ and Case I of heat absorption.

$\left(\sum_{n=0}^{\infty}=\sum_{n=0}^{N}\text{, E.S.: Exact Solution}\right)$

N \ t'	0.01	0.1	1	10	100	∞
5	0.018	0.137	0.423	0.598	0.659	0.689
10	0.018	0.134	0.419	0.590	0.651	0.681
15	0.018	0.134	0.419	0.591	0.651	0.681
E.S.	0.018	0.137	0.425	0.595	0.653	0.681

Table 3 Comparison of $\widehat{K_I}$ by this method with exact one for $B_i = 0$ and Case I of heat absorption.

$\left(\sum_{n=0}^{\infty}=\sum_{n=0}^{N}\text{, E.S.: Exact Solution}\right)$

N \ t'	0.01	0.1	1	10	100	∞
5	0.021	0.101	0.308	0.460	0.512	0.541
10	0.021	0.100	0.305	0.455	0.508	0.535
15	0.021	0.100	0.305	0.455	0.508	0.535
E.S.	0.021	0.100	0.305	0.455	0.508	0.535

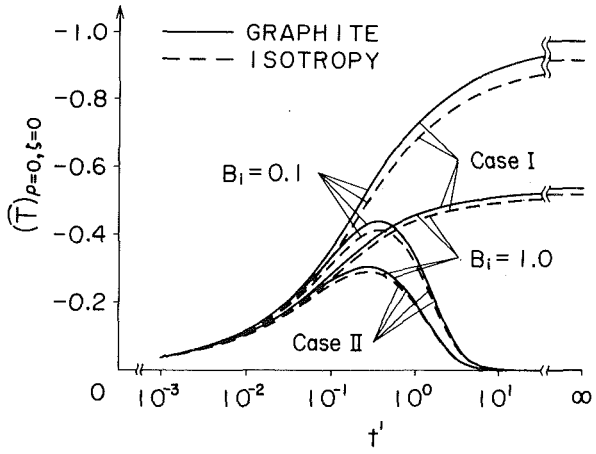


Fig. 1 Comparisons of $(\bar{T})_{\rho=\zeta=0}$ for graphite with that for isotropic conditions

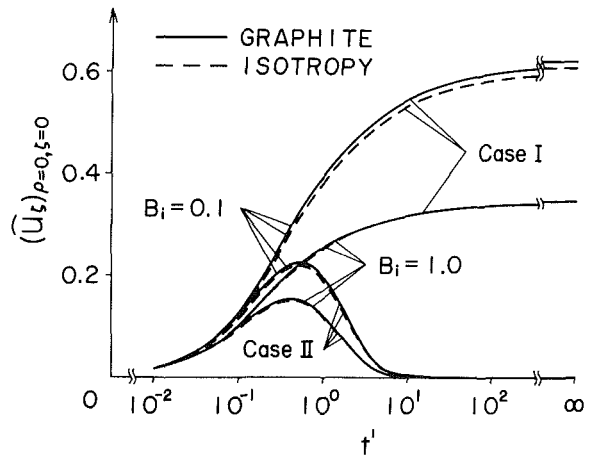


Fig. 2 Comparisons of $(\bar{u}_\zeta)_{\rho=\zeta=0}$ for graphite with that for isotropic conditions

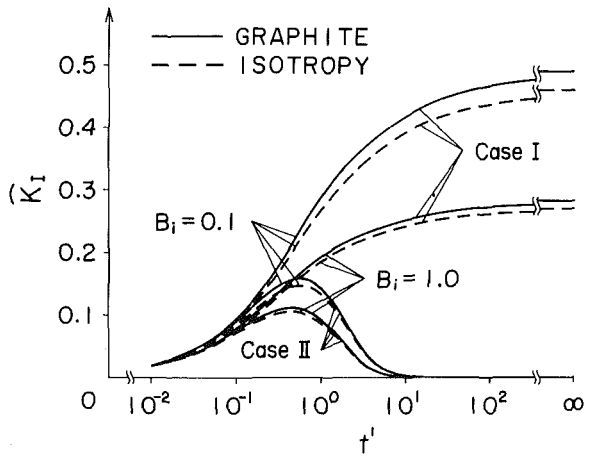


Fig. 3 Comparisons of $\bar{K}_I$ for graphite with that for isotropic conditions

$$K_I = -2\int_0^\infty \delta^{-1}\sin(\delta a)\int_0^\infty (c_{13}\delta^2 R_{\delta\omega} + c_{33}\omega^2 S_{\delta\omega}/\lambda^2 + \beta_3 \bar{T}_{\delta\omega})\,d\omega\,d\delta/\sqrt{\pi a} \tag{45}$$

where

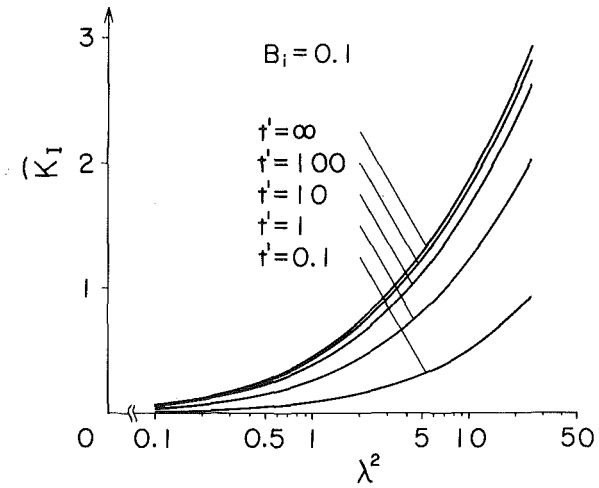


Fig. 4 Effects of λ^2 on $\bar{K}_I$ for $B_i = 0.1$ for Case I of heat absorption (M-I)

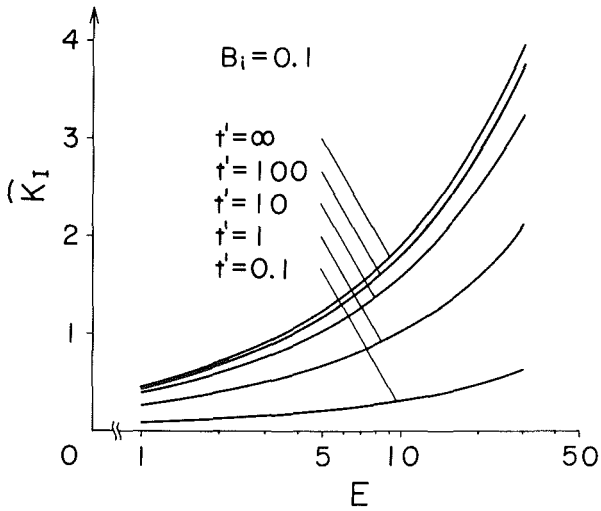


Fig. 5 Effects of E on $\bar{K}_I$ for $B_i = 0.1$ for Case I of heat absorption (M-II)

$$\left.\begin{aligned}
\bar{T}_{\delta\omega} &= -2\kappa Q_0 \delta \bar{f}\int_0^t g(t-\tau) \\
&\quad \times \exp\{-\kappa(\delta^2+\omega^2)\tau\}\,d\tau/\pi \\
R_{\delta\omega} &= \{-\beta_1(c_{33}\omega^2 + c_{44}\delta^2\lambda^2) \\
&\quad + \beta_3(c_{13}+c_{44})\omega^2\}\cdot\overline{RS} \\
S_{\delta\omega} &= \{\beta_1(c_{13}+c_{44})\delta^2\lambda^2 \\
&\quad - \beta_3(c_{11}\delta^2\lambda^2 + c_{44}\omega^2)\}\cdot\overline{RS}
\end{aligned}\right\} \tag{46}$$

and

$$\bar{f} = \lambda\int_0^a r f(r) J_0(\delta r)\,dr/\lambda_z$$

$$\overline{RS} = \lambda^2 \bar{T}_{\delta\omega}/\{(c_{11}\delta^2\lambda^2 + c_{44}\omega^2)(c_{33}\omega^2 + c_{44}\delta^2\lambda^2) - (c_{13}+c_{44})^2\delta^2\omega^2\lambda^2\}$$

4 Numerical Results and Discussion

For convenience of numerical calculations, the following dimensionless quantities are introduced:

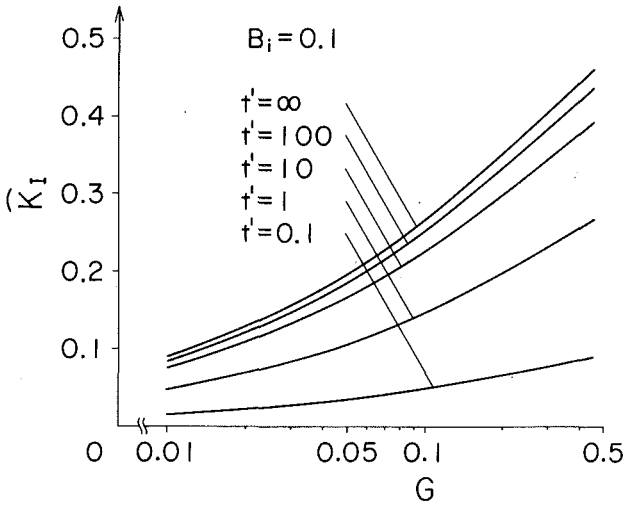


Fig. 6 Effects of G on $\widehat{K_I}$ for $B_i = 0.1$ for Case I of heat absorption (M-III)

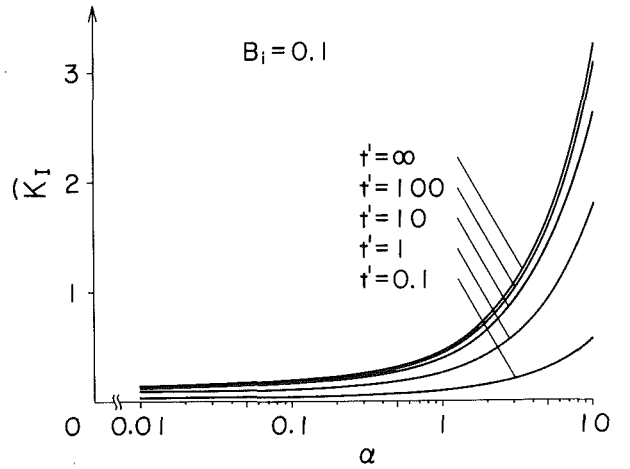


Fig. 7 Effects of α on $\widehat{K_I}$ for $B_i = 0.1$ for Case I of heat absorption (M-IV)

$$\rho = r/a, \quad \zeta = z/a, \quad t' = \kappa t/a^2,$$
$$\widehat{T} = \lambda_z T/aQ_0, \quad \widehat{u_\zeta} = u_z \lambda_z/\alpha_r a^2 Q_0,$$
$$\widehat{K_I} = K_I \lambda_z/\alpha_r E_r a^{3/2} Q_0, \quad B_i = ah,$$
$$E = E_z/E_r, \quad G = G_{rz}/E_r, \quad \alpha = \alpha_z/\alpha_r$$

where E_r, E_z, and G_{rz} are moduli of elasticity, α_r and α_z are coefficients of linear thermal expansion.

Numerical calculations were carried out for the transversely isotropic infinite solid subjected to the following two types of heat absorption through the crack surface.

Case I Heat absorption rate is constant, namely,

$$f(\rho)g(t') = H(1-\rho).$$

Case II Heat absorption rate depends on time, namely,

$$f(\rho)g(t') = H(1-\rho)\exp(-t').$$

The time variable is expressed by

$$t'_j = 10^{(j-41)/10}, \quad (j = 1, 2, 3, \ldots)$$

The material constants are taken as

$$\lambda_r = 1.172 \text{ W/(mK)}, \quad \lambda_z = 1.340 \text{ W/(mK)},$$
$$E_r = 10.4 \text{ GPa}, \quad E_z = 11.8 \text{ GPa}, \quad G_{rz} = 4.14 \text{ GPa},$$
$$\alpha_r = 3.9 \times 10^{-6} \text{ K}^{-1}, \quad \alpha_z = 3.5 \times 10^{-6} \text{ K}^{-1},$$
$$\nu_{r\theta} = \nu_{rz} = 0.11$$

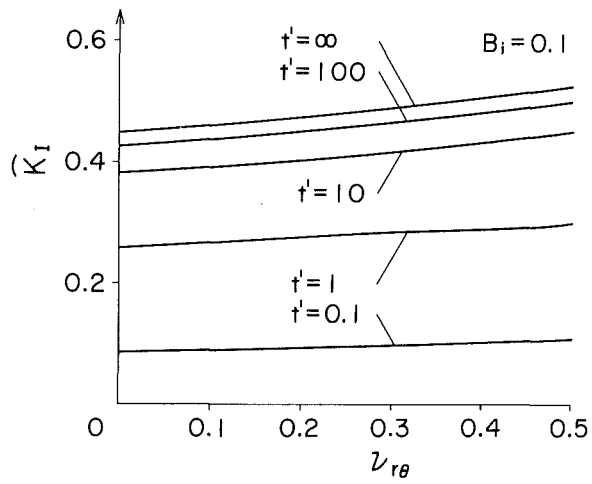


Fig. 8 Effects of $\nu_{r\theta}$ for $\widehat{K_I}$ for $B_i = 0.1$ for Case I of heat absorption (M-V)

Table 4 Various materials of transverse isotropy

Material	λ^2	E	G	α	$\nu_{r\theta}$
M-I	0.1~25	1	0.4504	1	0.11
M-II	1	1~30	0.4504	1	0.11
M-III	1	1	0.01~0.4504	1	0.11
M-IV	1	1	0.4504	0.01~10	0.11
M-V	1	1	0.4504	1	0~0.5

for a graphite and

$$\lambda^2 = E = \alpha = 1, \quad G = 0.4504, \quad \nu_{r\theta} = \nu_{rz} = 0.11$$

for isotropic conditions, where $\nu_{r\theta}$ and ν_{rz} are Poisson's ratios.

It is necessary to examine the propriety of this finite difference method and the convergence of the infinite series solution for the dual-integral equation. The temperature, axial displacement, and stress intensity factor of a graphite were calculated for various upper limit numbers of the infinite series and compared with the exact solutions for $B_i = 0$ and Case I of heat absorption. The calculated results are shown in Tables 1 to 3. It is clear from these tables that the infinite series solutions converge for $N = 10$, at a reasonable rate. The value of $\widehat{K_I}$ from the finite difference solution agrees quite well with the exact one. However, for $\widehat{T}$ and $\widehat{u_\zeta}$, some discrepancy is appreciable between the finite difference solution and the exact solution. (It should be noticed that numerical integrations in the temperature and axial displacement are more formidable than that in the stress intensity factor.)

Figures 1 to 3 compare the results of temperature, axial displacement and stress intensity factor of a graphite with those derived under isotropic conditions for Case I and II. These figures show that $\widehat{T}$, $\widehat{u_\zeta}$, and $\widehat{K_I}$ in a graphite are somewhat larger than those derived under isotropic conditions, and that $\widehat{T}$, $\widehat{u_\zeta}$, and $\widehat{K_I}$ for $B_i = 1.0$ are smaller than those for $B_i = 0.1$. The maximum values of temperature, axial displacement, and stress intensity factor for Case II appear when Fourier's number is between 0.1 and 1.0.

Figures 4 to 8 illustrate the effects of anisotropies of the transversely isotropic material constants on the stress intensity factor for $B_i = 0.1$ and Case I of heat absorption (when only one among the material constants such as λ^2, E, G, α, and $\nu_{r\theta}$ indicates various anisotropies, while the other material constants are kept equal to those of isotropic conditions as shown in Table 4). These figures show that the stress intensity factor is strongly influenced by the values of λ^2, E, and α, and that

the effects of G and $\nu_{r\theta}$ on the stress intensity factor are small in comparison to the effects of λ^2, E, and α.

The present numerical calculations were made with double precision accuracy by NEC ACOS-1000 in the Computation Center at Osaka University.

References

Atkinson, C., and Clements, D. L., 1977, "On Some Crack Problems in Anisotropic Thermoelasticity," *International Journal of Solids and Structures*, Vol. 13, pp. 855–864.

Clements, D. L., 1983, "A Thermoelastic Problem for a Crack Between Dissimilar Anisotropic Media," *International Journal of Solids and Structures*, Vol. 19, pp. 121–130.

Clements, D. L., and Tauchert, T. R., 1979, "A Thermoelastic Crack Problem for an Anisotropic Slab," *J. Austr. Math. Soc.*, Ser. B, Vol. 21, pp. 243–255.

Ghosh, S., 1977, "A Note on Two-Dimensional Thermo-Elastic Crack Problem for an Aeolotropic Solid," *Indian Journal of Mechanical Mathematics*, Vol. 15, pp. 6–15.

Maiti, S. N., and Misra, J. C., 1976, "Thermal Stresses in a Finite Aeolotropic Cylinder Due to Thermal Crack," *Rev. Roum. Sci. Tech.*, Vol. 21, pp. 339–351.

Murata, K., and Atsumi, A., 1977, "Thermal Stresses in a Transversely Isotropic Cylinder Containing a Penny-Shaped Crack," *Lett. Appl. Eng. Sci.*, Vol. 5, pp. 173–185.

Sneddon, I. N., 1966, *Mixed Boundary Value Problems in Potential Theory*, North-Holland, p. 76.

Sumi, N., 1981, "Thermal Stresses in an Orthotropic Rectangular Plate With a Rigid Ribbonlike Inclusion," *Nuclear Engineering and Design*, Vol. 66, pp. 405–411.

Sumi, N., 1982, "The Thermoelastic Problem for an Orthotropic Rectangular Plate With an Oblique Crack," *Theoretical and Applied Mechanics*, Vol. 31, pp. 135–143.

Takeuti, Y., and Noda, N., 1978, "A General Treatise on the Three-Dimensional Thermoelasticity of Curvilinear Aeolotropic Solids," *Journal of Thermal Stresses*, Vol. 1, pp. 25–39.

Tsai, Y. M., 1983a, "Thermal Stress in a Transversely Isotropic Medium Containing a Penny-Shaped Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 24–28.

Tsai, Y. M., 1983b, "Transversely Isotropic Thermoelastic Problem of Uniform Heat Flow Disturbed by a Penny-Shaped Crack," *Journal of Thermal Stresses*, Vol. 6, pp. 379–389.

Circumferential Crack at the Fixed End of a Cylinder in Flexure

J. A. Alabi
Mechanical Engineering Department,
University of Science and Technology,
Port Harcourt, Nigeria

The problem of a circumferential through-crack at the fixed end of a cylinder in flexure is considered herein. An expression for the energy-release rate, in closed form, is obtained according to linear fracture mechanics and thin shell theory. The expression consists of two terms, one of which is the energy-release due to shear and the other is the energy-release due to bending. The former is relatively negligible except for very short cylinders. A formula for the flexibility factor is also derived.

Introduction

The problem of circumferential through-cracks in pipes and cylindrical shells has been investigated by a number of authors (Folias, 1967; Erdogan and Ratwani, 1970; Duncan-Fama and Sanders, 1972; Nicholson et al., 1983; Sanders, 1982; Alabi, 1984; Alabi and Sanders, 1985; Barsoum et al., 1979; Tada et al., 1980; Smith, 1984). In most of these investigations the cylindrical shell has been subjected to simple tension or pure bending. However, a cracked pipe or cylinder may be loaded by a shear force, and the question of how the varying bending moment produced compares with a similar pipe or cylinder subjected to constant bending has practical importance. The analyses of Tada et al. (1980) and Smith (1984) have shed some light on this problem. The present work is concerned with the problem of a circumferential through-crack at the fixed end of a long circular cylinder loaded by a shear force.

The shell under consideration contains a circumferential crack at one end which is otherwise held rigid. A shearing force is applied at the other end. The investigation is along the lines of the analyses carried out in Sanders (1982) and Alabi and Sanders (1985). As in Sanders (1982) and Alabi and Sanders (1985), the solution is applicable to cylindrical shells containing a sufficiently long crack and is valid everywhere except close to the crack tips. The difficulty with the solution near the crack tips is overcome by the use of path-independent integrals to evaluate the energy-release rate. The results show that the energy-release rate is composed of two terms, one of which is the contribution due to shearing and the other is due to bending. The shear term is insignificant except for very short cylinders. Thus for all but very short cylinders, the energy-release rate is effectively proportional to the square of the cylinder length. This may be contrasted with cylinders subjected to pure bending where, as has been shown (Alabi and Sanders, 1985), the energy-release rate is independent of the length of the cylinder provided the latter is sufficiently long. This implies that a shearing force can have a much more severe effect on a circumferential crack in long cylinders than pure bending because of the large flexural stresses produced.

Equations and Boundary Conditions

The complex displacements and stress functions in a circular cylindrical shell are given by (Sanders, 1983)

$$\begin{aligned} u &= \epsilon^{-2}\ddot{\Phi}' - \nu\phi' & \chi_z &= -i\epsilon^{-2}\ddot{\Phi}' - i\nu\phi' \\ v &= \epsilon^{-2}\ddot{\Phi} - (2+\nu)\dot{\phi} & \chi_\theta &= i\epsilon^{-2}\ddot{\Phi} + i(2-\nu)\dot{\phi} \\ w &= -\ddot{\Phi} + i(1+2i\epsilon^2)\phi & \psi &= i\ddot{\Phi} + (1+2i\epsilon^2)\phi \end{aligned} \tag{1}$$

where $(\,)' \equiv \partial/\partial z\,(\,)$ and $(\,)^{\cdot} \equiv \partial/\partial\theta\,(\,)$. z and θ are dimensionless coordinates specified by the condition that axial and circumferential distances are, respectively, given by Rz and $R\theta$ where R is the cylinder radius. The physical displacements and stress functions are obtained from the real parts of equations (1). The characteristic functions Φ and ϕ both satisfy

$$\nabla^4\Omega + \ddot{\Omega} - i\epsilon^{-2}(1+2i\epsilon^2)\Omega'' = 0 \tag{2}$$

and are related to each other by $\Phi'' = \epsilon^2\phi$. The dimensionless variables in equations (1) are related to the dimensional quantities $\bar{u}$, $\bar{v}$, etc., by

$$\begin{aligned} (\bar{u},\bar{v}) &= \frac{P}{Eh}(u,v) & \bar{w} &= \frac{P}{\epsilon^2 Eh}w \\ (\bar{\chi}_z,\bar{\chi}_\theta) &= \epsilon^2 PR(\chi_z,\chi_\theta) & \bar{\psi} &= PR\psi \end{aligned} \tag{3}$$

Similarly, the dimensionless membrane and bending stress measures $N_{\alpha\beta}$ and $M_{\alpha\beta}$ are related to the dimensional forms by

$$\bar{N}_{\alpha\beta} = \frac{P}{R}N_{\alpha\beta} \quad \bar{M}_{\alpha\beta} = \epsilon^2 PM_{\alpha\beta} \tag{3a}$$

Here, P is the magnitude of the shear force applied at the "free" end of the cylinder, h is the shell thickness, E is Young's modulus, ν is Poisson's ratio, and ϵ is a small parameter given by

$$\epsilon^2 = (h/R)[12(1-\nu^2)]^{-1/2} \tag{4}$$

In terms of the characteristic function ϕ, the membrane and

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May 23, 1986; final revision March 3, 1987.

Paper No. 87-WA/APM-17.

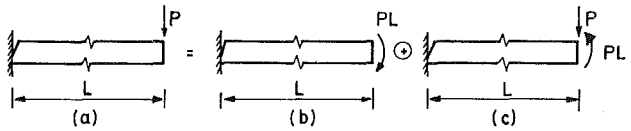


Fig. 1 Cylinder of length L loaded by a shearing force

bending stress resultants are given by (Sanders, 1983) expressions such as $N_z = \ddot{\phi}$, $M_z = -i\phi'' - i\nu\ddot{\phi}$, etc.

On the stress-free crack, boundary conditions on the stress functions are (Alabi and Sanders, 1985)

$$\chi_z = \chi_\theta = \psi = \psi' = 0 \quad \text{on} \quad z = 0, |\theta| < \alpha \tag{5}$$

Off the crack where the cylinder is fixed, we have

$$u = v = w = w' = 0 \quad \text{on} \quad z = 0, \alpha < \theta < 2\pi - \alpha \tag{6}$$

equations (1), (5), and (6) lead to the boundary conditions in terms of the characteristic functions Φ and ϕ. On the crack

$$\begin{aligned} &\Re\{(1+i\epsilon^2\nu)\phi_c\} = 0 \\ &\Re\{i\ddot{\Phi}_c + i\epsilon^2(2-\nu)\phi_c\} = 0 \\ &\Re\{\phi_c' + i\epsilon^2(2-\nu)\phi_c'\} = 0 \\ &\Re\{i\ddot{\Phi}_c' + i\epsilon^2\nu\phi_c'\} = 0 \end{aligned} \tag{7}$$

and off the crack

$$\begin{aligned} &\Re\{(i+\epsilon^2\nu)\phi_c\} = 0 \\ &\Re\{\ddot{\Phi}_c + \epsilon^2(2+\nu)\phi_c\} = 0 \\ &\Re\{i\phi_c' - \epsilon^2(2+\nu)\phi_c'\} = 0 \\ &\Re\{\ddot{\Phi}_c' - \epsilon^2\nu\phi_c'\} = 0. \end{aligned} \tag{8}$$

The symbol $\Re\{\ \}$ denotes the real part of the expression in brackets. The subscript c means that the expressions are in terms of the complete characteristic functions.

Solution

The analysis is much simpler if we consider a composite of problems (Fig. 1) in which two different types of loading are applied. The solution to the first problem in which the cracked cylinder is subjected to pure bending is known (Alabi and Sanders, 1985). What follows is the solution to the second problem in which the cracked cylinder is loaded by a shearing force and a bending moment of equal but opposite sign to that in the first problem.

First, we formulate a reduced boundary value problem in such a way that the reduced characteristic function ϕ and $\ddot{\Phi}$ vanish at infinity. The stresses very far away from the fixed end of the cylinder are required to be those corresponding to shear and bending. The solution for combined shear and bending is given by

$$\begin{aligned} &\pi\phi = -(1+2i\epsilon^2)z\cos\theta \\ &\pi\ddot{\Phi} = \frac{1}{6}\epsilon^2 z^3\cos\theta - \frac{1}{2}iz\theta\sin\theta - i(1-i\epsilon^2\nu)z\cos\theta. \end{aligned} \tag{9}$$

In addition, elementary solutions corresponding to rigid body motions (for which stresses vanish) and null effect (for which displacements vanish) are allowed to be present at infinity provided they have the proper symmetry. This list of solutions, null effect, axial translation, lateral translation, and a rotation are, respectively,

$$\begin{aligned} &\phi = i\epsilon^{-1}a, \ddot{\Phi} = -\epsilon^{-1}a(1+2i\epsilon^2); \phi = ibz, \ddot{\Phi} = -bz(1+2i\epsilon^2) \\ &\phi = 0, \ddot{\Phi} = \epsilon^{-1}c\cos\theta; \phi = 0, \ddot{\Phi} = dz\cos\theta \end{aligned} \tag{10}$$

Now put $\phi_c = \phi_E + \phi$, $\ddot{\Phi}_c = \ddot{\Phi}_E + \ddot{\Phi}$ where ϕ_E, $\ddot{\Phi}_E$ refer to the list of elementary solutions (9) and (10). Accordingly, the boundary conditions (7) and (8) reduce to

$$\begin{aligned} &\Re\{(1+i\epsilon^2\nu)\phi\} = \epsilon^{-1}a_I + \epsilon\nu a_R \\ &\Re\{i\ddot{\Phi} + i\epsilon^2(2-\nu)\phi\} = -\epsilon^{-1}a_I - \epsilon\nu a_R + \epsilon^{-1}c_I\cos\theta \\ &\Re\{\phi' + i\epsilon^2(2-\nu)\phi'\} = b_I + \epsilon^2(2-\nu)b_R + \frac{1}{\pi}\cos\theta \\ &\Re\{i\ddot{\Phi}' + i\epsilon^2\nu\phi'\} = -b_I - \epsilon^2(2-\nu)b_R + d_I\cos\theta + \\ &\qquad -\frac{1}{2\pi}\theta\sin\theta - \frac{1}{\pi}\cos\theta \end{aligned} \tag{11}$$

to be satisfied on the crack and

$$\begin{aligned} &\Re\{(i+\epsilon^2\nu)\phi\} = \epsilon^{-1}a_R + \epsilon\nu a_I \\ &\Re\{\ddot{\Phi} + \epsilon^2(2+\nu)\phi\} = \epsilon^{-1}a_R + \epsilon\nu a_I - \epsilon^{-1}c_R\cos\theta \\ &\Re\{i\phi' - \epsilon^2(2+\nu)\phi'\} = b_R - \epsilon^2(2+\nu)b_I + \epsilon^2\nu\cos\theta \\ &\Re\{\ddot{\Phi}' - \epsilon^2\nu\phi'\} = b_R - \epsilon^2(2+\nu)b_I - d_R\cos\theta \end{aligned} \tag{12}$$

off the crack. The subscripts R and I refer to the real and imaginary parts of these constants.

It is evident that for sufficiently long cracks, the reduced characteristic functions do not vary rapidly in the circumferential directions, i.e., ϕ and $\dot{\phi}$ are the same order, except very close to the crack tips. In such a situation, a solution that is accurate everywhere except in the immediate neighborhood of the crack tips exists. The solution is compounded of semi-membrane solutions governed by

$$(\ddot{\Omega} + \Omega)^{\cdot\cdot} - i\epsilon^{-2}\Omega'' = 0 \tag{13}$$

and edge-effect solutions governed by

$$\Omega'' - i\epsilon^{-2}\Omega = 0 \tag{14}$$

More precisely, semi-membrane and edge-effect solutions are of the form $\Omega(\xi, \theta, \epsilon)$ where $\xi = \epsilon z$ or $\epsilon^{-1}z$, as the case may be and the leading term in the asymptotic expansion of ϕ in powers of ϵ^2 satisfies (13) or (14). A result involving the semi-membrane solutions needed later is the following:

$$(\ddot{\Phi} + \Phi)^{\cdot\cdot} - i\phi + 2\epsilon^2(\ddot{\phi} + \phi) = 0(\epsilon^4). \tag{15}$$

Assume a solution in the form $\phi + \epsilon^2\bar{\phi}$ where ϕ is semi-membrane and $\bar{\phi}$ is edge-effect. $\bar{\Phi}$ is order ϵ^4 too small and may be neglected. The boundary conditions (11), (12), equation (15), and the relations

$$\phi' = 0(\epsilon\phi), \quad \bar{\phi}' = 0(\epsilon^{-1}\bar{\phi})$$

lead to

$$\begin{aligned} \Re\{\bar{\phi}\} &= \Re\{i\nu\ddot{\bar{\phi}}\} \\ \Re\{\bar{\phi}'\} &= 0 \end{aligned} \qquad \text{on } z = 0, |\theta| < \alpha \tag{16}$$

and

$$\begin{aligned} \Re\{i\bar{\phi}\} &= \Re\{\nu\ddot{\bar{\phi}}\} \\ \Re\{i\bar{\phi}'\} &= \nu\cos\theta \end{aligned} \qquad \text{on } z = 0 \text{ and } \alpha < \theta < 2\pi - \alpha \tag{17}$$

Denote the edge-effect solution by $\bar{\phi} = i^{1/2}G(\theta)e^{-i^{1/2}\epsilon^{-1}z}$. The conditions (16) and (17) are satisfied if:

G is real and $\Re\{i^{1/2}G\} = \Re\{i\nu\ddot{\bar{\phi}}(0,\theta\}$ on $|\theta| < \alpha$

and

$\Re\{G\} = \nu\cos\theta$ and $\Re\{i^{3/2}G\} = \Re\{\nu\ddot{\bar{\phi}}(0,\theta)\}$ on $\alpha < \theta < 2\pi - \alpha$.

The edge-effect is thus determined by the boundary values of the semi-membrane solution.

Equation (13) governing semi-membrane solutions requires only two boundary conditions on $z = 0$. Only a first-approximation solution is sought, and in view of equation (15), the first and third of boundary conditions (11) are not independent of the second and fourth, respectively, and similarly the first and third of boundary conditions (12) are not independent of the second and fourth, respectively. The problem has now been reduced to finding a (first-approximation) semi-membrane solution $\ddot{\Phi}$ for which

$$\Re\{i\ddot{\Phi}\} = -\epsilon^{-1}a_I + \epsilon^{-1}c_I\cos\theta \tag{18}$$

$$\Re\{i\ddot{\Phi}'\} = -b_I + d_I\cos\theta - \frac{1}{2\pi}\theta\sin\theta - \frac{1}{\pi}\cos\theta \tag{19}$$

at $z=0$ on the crack, and

$$\Re\{\ddot{\Phi}\} = \epsilon^{-1}a_R - \epsilon^{-1}c_R\cos\theta \tag{20}$$

$$\Re\{\ddot{\Phi}'\} = b_R - d_R\cos\theta \tag{21}$$

at $z=0$ off the crack. From the relation $\Phi'' = \epsilon^2\phi$ and equation (13) governing semi-membrane solutions follows

$$\phi = -i(\ddot{\Phi} + \Phi)'' \tag{22}$$

from which ϕ may be obtained once $\ddot{\Phi}$ is known.

Any solution to equation (13), even in θ, which vanishes at $z=\infty$, must satisfy the conditions

$$\int_0^\pi \ddot{\Phi}(0,\theta)d\theta = \int_0^\pi \ddot{\Phi}(0,\theta)\cos\theta d\theta = 0 \tag{23}$$

and it can be shown that such solutions are closely approximated by solutions to the substitute equation (Sanders, 1983)

$$\ddot{\Phi}' = -i^{3/2}\epsilon\left(\ddot{\Phi} + \frac{1}{2}\Phi\right)'' \tag{24}$$

The usefulness of this relation lies in the fact that the boundary values of $\ddot{\Phi}'$ can be expressed in terms of the boundary values of $\ddot{\Phi}$ and its second derivative with respect to θ. This allows the reduction of the boundary conditions to a system of ordinary differential equations. Now put $\ddot{\Phi}(0,\theta) = F(\theta)$ and use equation (24) to get

$$\ddot{\Phi}'(0,\theta) = -i^{3/2}\epsilon\left(\ddot{F} + \frac{1}{2}F\right) \tag{25}$$

Because of equation (25), equations (18)–(21) may be rewritten as

$$\epsilon F_I = a_I - c_I\cos\theta \tag{26}$$

$$\Re\{i^{1/2}\epsilon(\ddot{F} + \frac{1}{2}F)\} = -b_I + d_I\cos\theta - \frac{1}{2\pi}\theta\sin\theta + \\ - \frac{1}{\pi}\cos\theta \tag{27}$$

to be satisfied at $z=0$ on the crack and

$$\epsilon F_R = a_R - c_R\cos\theta \tag{28}$$

$$\Re\{i^{3/2}\epsilon(\ddot{F} + \frac{1}{2}F)\} = -b_R + d_R\cos\theta \tag{29}$$

to be satisfied at $z=0$ off the crack.

Equations (26)–(29) are a system of ordinary differential equations for the real and imaginary parts of F. The solution, even about $\theta=0$ and $\theta=\pi$, contains ten undetermined constants (in real terms). These are determined by the conditions (23) applied to F, and certain continuity conditions demanded by physical considerations. These latter conditions are that the real displacements u, v, w, and real stress functions χ_z, χ_θ, ψ must be continuous at the crack tips. In terms of F, these require that F, $\dot{F}$, $\ddot{F}$ be continuous and (23) to be satisfied by F. There are thus ten constants to be determined by ten conditions. The constants a and c turn out to be real. The solution is

$$\epsilon F = -2\sqrt{2}(b_I + d_I\cos\theta) + \frac{\sqrt{2}}{\pi}\theta\sin\theta + \frac{6\sqrt{2}}{\pi}\cos\theta + \\ + A\cos\frac{\theta}{\sqrt{2}} \text{ on the crack} \tag{30}$$

and

$$\epsilon F = -\sqrt{-2i}(a - c\cos\theta) + 2\sqrt{2}i(b_R + d_R\cos\theta) \\ \text{off the crack} \tag{31}$$

where

$$\left.\begin{aligned}
\pi A &= \frac{2\alpha - \sin 2\alpha}{\cos\alpha\sin\frac{\alpha}{\sqrt{2}} - \frac{1}{\sqrt{2}}\sin\alpha\cos\frac{\alpha}{\sqrt{2}}} \\
\pi^2 a &= 2\sqrt{2}(\sin\alpha - \alpha\cos\alpha) + \pi A\left(\frac{\alpha}{2}\cos\frac{\alpha}{\sqrt{2}} - \frac{1}{\sqrt{2}}\sin\frac{\alpha}{\sqrt{2}}\right) \\
2\sqrt{2}b_R &= -a \\
4\pi b_I &= \cos\alpha + \sqrt{2}\pi a + \frac{\pi A}{\sqrt{2}}\cos\frac{\alpha}{\sqrt{2}} \\
\sqrt{2}\pi^2 c\sin\alpha &= A\Gamma - 2\alpha^2\cos\alpha - 3\alpha\sin\alpha + \\
&\quad + 5\cos\alpha\sin^2\alpha \\
2\sqrt{2}d_R &= c \\
4\pi d_I\cos\alpha &= 2\alpha\sin\alpha + 8\cos\alpha + \frac{\pi A}{\sqrt{2}}\cos\frac{\alpha}{\sqrt{2}} - \sqrt{2}\pi c\cos\alpha \\
\Gamma &= -\frac{3}{2}\sin^2\alpha\cos\frac{\alpha}{\sqrt{2}} + \frac{1}{2\sqrt{2}}\sin\frac{\alpha}{\sqrt{2}}(\alpha + 5\cos\alpha\sin\alpha)
\end{aligned}\right\} \tag{32}$$

If required, the solution for $\ddot{\Phi}(z,\theta)$, and by equation (22) for $\phi(z,\theta)$ can be constructed by Fourier series or by means of an integral representation (Sanders, 1980).

Energy-Release Rate

The energy-release rate may be obtained by adding the energy-release rates in the composite of problems mentioned earlier. If the energy-release rate $\bar{I}$ (with respect to α) is made dimensionless according to the relation $\bar{I} = (P^2/Eh)I$, then for the cylinder loaded as shown in Fig. 1, I is given by

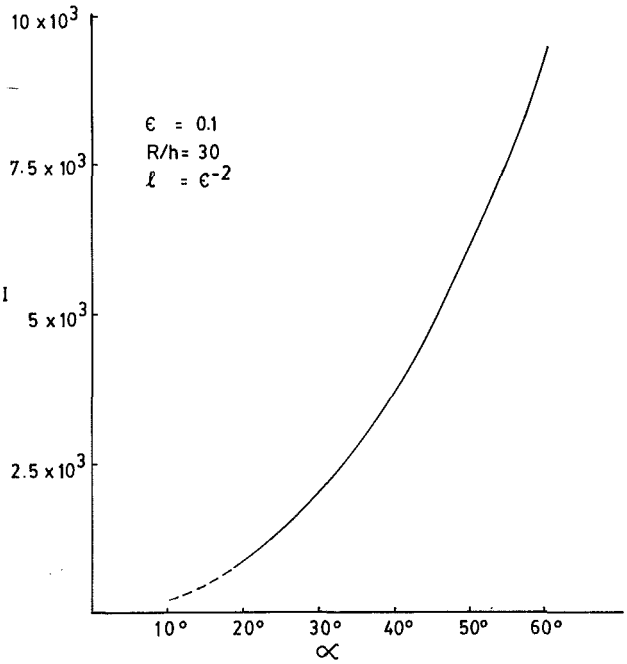


Fig. 2 Dimensionless energy-release rate versus crack angle

$$\pi I=\frac{1}{2}\int_{-\alpha}^{\alpha}\cos\theta\frac{\partial u_c}{\partial\alpha}(0,\theta)d\theta$$

$$+\frac{1}{2}\int_{-\alpha}^{\alpha}\sin\theta\frac{\partial v_c}{\partial\alpha}(0,\theta)d\theta \tag{33}$$

The formula (33) is a special case of the path-independent integrals derived by Nicholson et al. (1983) for the calculation of I in cracked shallow shells. It can be shown that formula (33) is applicable to the complete cylinder as well as to the shallow cylinder. For the complete cylinder, formula (33) can be shown to give the following formula

$$\pi I=\frac{1}{2}\frac{\partial}{\partial\alpha}\int_{-\pi}^{\pi}\cos\theta u_c(z,\theta,\alpha)d\theta$$

$$+\frac{1}{2}\frac{\partial}{\partial\alpha}\int_{-\pi}^{\pi}\sin\theta v_c(z,\theta,\alpha)d\theta \tag{34}$$

In formula (34), u_c, v_c may be replaced by u_E, v_E, respectively, because $\int_{-\pi}^{\pi} u\cos\theta d\theta=\int_{-\pi}^{\pi} v\sin\theta d\theta=0$. Both formulae are equivalent but the second is more easily evaluated to give, on account of equations (10) and (32) and equations (10) and (32) of Alabi and Sanders (1985),

$$I=\frac{\epsilon^{-1}}{\sqrt{2}\pi^2}\sin^2\alpha(\epsilon^{-2}\lambda^2+\ell^2\mu^2) \tag{35}$$

where

$$\left.\begin{aligned}\lambda^2&=8\left(1+\frac{\alpha\operatorname{cosec}^2\alpha-\cot\alpha}{2\left(2\cot\alpha+\sqrt{2}\cot\frac{\alpha}{\sqrt{2}}\right)}\right)^2\\ \mu^2&=\frac{1}{8}\lambda^2+\left(1+\frac{\alpha\operatorname{cosec}^2\alpha-\cot\alpha}{2\left(2\cot\alpha+\sqrt{2}\cot\frac{\pi-\alpha}{\sqrt{2}}\right)}\right)^2\\ \text{and}\quad \ell&=\frac{L}{R}\end{aligned}\right\} \tag{36}$$

It should be noted that formula (35) corresponds to the total energy release-rate for the configuration of Fig. 1(a) which is a superposition of the energy-release rates in Fig. 1(b) and Fig. 1(c). The second term in formula (35) may be obtained from equations (37) and (38) of Alabi and Sanders (1985) with $T=0$ and $M=PL$.

Numerical evaluation of formulae (35) and (36) shows that the first term in formula (35) is negligible except for moderate values of $\ell(\ell=0(1))$. This result is similar to that for the flexure of solid beams in which shear effects are known to be insignificant except for beams of comparable length and cross-sectional dimensions. Formula (35) shows that I depends on the length of the cylinder or pipe unlike I for a similar cylinder or pipe subjected to pure bending.

Flexibility Factor

The flexibility factor f is defined as the ratio of the rotation of the cracked cylinder (or pipe) at any cross section to that of an uncracked cylinder (or pipe) of the same length. f is always greater than unity because the crack gives rise to an additional rotation. For a cylinder of length ℓ, f is given by

$$f=1+\frac{6\epsilon^{-2}(d_R^0+\epsilon d_R^1)}{\ell^2+3\ell-6(2+\nu)} \tag{37}$$

where

$$\left.\begin{aligned}4\pi d_R^0&=9\alpha-7\cos\alpha\sin\alpha-2\alpha^2\cot\alpha+\\ &\quad+\frac{2(\alpha\operatorname{cosec}\alpha-\cos\alpha)^2}{\cot\alpha-\frac{1}{\sqrt{2}}\cot\frac{\alpha}{\sqrt{2}}}\\ \text{and}\quad 4\sqrt{2}\pi d_R^1&=4\pi d_R^0-\frac{(\alpha\operatorname{cosec}\alpha-\cos\alpha)^2}{\cot\alpha-\frac{1}{\sqrt{2}}\cot\frac{\alpha}{\sqrt{2}}}\\ &\quad+\frac{(\alpha\operatorname{cosec}\alpha-\cos\alpha)^2}{\cot\alpha+\frac{1}{\sqrt{2}}\cot\frac{\pi-\alpha}{\sqrt{2}}}\end{aligned}\right\} \tag{38}$$

Formula (37) is the flexibility factor for the cylinder in flexure as in Fig. 1(a). Numerical evaluation of formula (37) shows that the flexibility factor is, effectively, unity for the long cylinders under investigation in this paper.

Concluding Remarks

As already noted, the expression for the energy-release rate is composed of two terms, one of which is the energy-release due to bending and the other is due to shear. The latter is negligible for all but very short cylinders, i.e., $\ell=0(1)$. This finding is similar to the known result for flexure of solid beams in which shear effects are insignificant except for beams whose cross-sectional dimensions are comparable with their length.

For $\ell=0(1)$, the expressions (35) and (36) for the energy release rate would not be very accurate due to end effects. More precisely, as has been discussed elsewhere (Sanders, 1982), semi-membrane solutions decay very slowly in the axial direction and in order to avoid end effects, the distance between the crack and any boundary or load must not be less than about $2R\sqrt{R/h}$, i.e., $\ell=0(\epsilon^{-1})$. Furthermore, for $\ell=0(1)$, the particular way in which the shear force P is applied at the end of the cylinder would (if it is different from that assumed in the present analysis) lead to further deviation from the results produced here. However, both corrections are not expected to be too "large" and expressions (35) and (36) for the energy release rate should at least serve as a guide in arriving at a more accurate result for the energy-release rate in such a situation.

The results for the dimensionless energy-release rate for $\epsilon=0.1$ and $\ell=\epsilon^{-2}$ are shown in Fig. 2. Since the first term in expression (35) is small in comparison with the second for all but relatively small ℓ, the results for I for other values of ϵ and ℓ are readily obtained from the ones given here. For angles α less than about twenty degrees, a curve from a shallow shell analysis (if available) would apply (Sanders, 1982). Because of the large flexural stresses produced at the fixed end of long cylinders, plastic effects cannot be totally ignored. For such long cylinders the so-called small scale yielding condition no longer holds and this further limits the validity of the results produced herein.

References

Alabi, J. A., 1984, "Circumferential Cracks in Cylindrical Shells," Ph.D. Thesis, Division of Applied Sciences, Harvard University.

Alabi, J. A., and Sanders, Jr., J. L., 1985, "Circumferential Crack at the Fixed End of a Pipe," *Engineering Fracture Mechanics*, Vol. 22, pp. 609–616.

Barsoum, R. S., Loomis, R. W., and Stewart, B. D., 1979, "Analysis of Through-Cracks in Cylindrical Shells by the Quarter-Point Elements," *International Journal of Fracture*, Vol. 15, pp. 259–280.

Duncan-Fama, M. E., and Sanders, Jr., J. L., 1972, "A Circumferential Crack in a Cylindrical Shell Under Tension," *International Journal of Fracture Mechanics*, Vol. 8, pp. 15–20.

Erdogan, F., and Ratwani, M., 1970, "Fatigue and Fracture of Cylindrical Shells Containing a Circumferential Crack," *International Journal of Fracture Mechanics*, Vol. 6, pp. 379–392.

Folias, E. S., 1967, "A Circumferential Crack in a Pressurized Cylindrical Shell," *International Journal of Fracture Mechanics*, Vol. 3, pp. 1–11.

Nicholson, J. W., Weidman, S. T., and Simmonds, J. G., 1983, "Sanders' Energy-Release Rate Integral for a Circumferentially Cracked Cylindrical Shell," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 373–378.

Sanders, Jr., J. L., 1980, "The Cylindrical Shell Loaded by a Concentrated Normal Force," *Mechanics Today*, Vol. 5, Reissner Volume, Pergamon Press, Oxford, England, pp. 427–438.

Sanders, Jr., J. L., 1982, "Circumferential Through-Cracks in Cylindrical Shells Under Tension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 103–107.

Sanders, Jr., J. L., 1983, "Analysis of Circular Cylindrical Shells," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 1165–1170.

Smith, E., 1984, "The Effect of Non-Uniform Bending on the Stability of Circumferential Growth of a Through-Wall Crack in a Pipe," *International Journal of Engineering Sciences*, Vol. 22, pp. 127–133.

Tada, H., Paris, P. C., and Gamble, R. M., 1980, "A Stability Analysis of Circumferential Cracks in Reactor Piping Systems," *Proceedings of the 12th National Symposium on Fracture Mechanics*, ASTM STP 700, pp. 296–313.

Treatment of Body Forces in 2D Elastostatic BEM Using Particular Integrals

D. A. Pape
Graduate Student.

P. K. Banerjee
Professor.

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

A new set of direct and indirect boundary element formulations for two-dimensional elastostatics have been developed. These new formulations differ from currently popular formulations in the treatment of body forces. The method of particular integrals is used resulting in a formulation which requires neither volume nor surface integration to solve the most general body force problem. This formulation is implemented in both direct and indirect boundary element methods using quadratic isoparametric elements. The efficiency and accuracy of this formulation for these two methods are compared for a range of problems. Finally, a multi-region problem with complicated geometry is run in order to show the complete generality of the particular integral method.

Introduction

In recent years, the boundary element method has emerged as a powerful technique for solving elastostatic problems. Because only surface discretization is necessary, boundary element methods (BEM) can have significant advantages over other methods requiring full domain discretization. Such advantages include savings in data preparation and data reduction as well as decreased computing time due to the reduced dimensionality of the problem. This advantage is partially lost if, as in early formulations, volume integration is required to solve the generalized body force problem.

In order to eliminate the volume discretization, Cruse et al. (1977) and Rizzo and Shippy (1977) have utilized the field equations for the body force potential and the divergence theorem to convert the volume integrals involved in the treatment of conservative body forces into an equivalent surface integral. Although this obviously eliminates any necessity for a volume discretization, additional surface integrations are now required. This volume integral conversion method has become very popular amongst many BEM workers (Danson, 1982; Kamiya and Sawaki, 1985; Nardini and Brebbia, 1982) in spite of this drawback.

In this paper the development of a very general approach to the treatment of body forces in direct and indirect BEM analyses is presented. The method is based on the well-known concept of developing the solution of an inhomogeneous differential equation by means of a complementary solution and a particular integral. This completely general approach, when applied to boundary element formulations, does not require any volume or additional surface integration to solve the general body force problem.

The second objective of this paper is to compare two formulations of the BEM, namely the direct boundary element method (DBEM) and the indirect boundary element method (IBEM). Computer programs based on the DBEM and IBEM are developed and tested in order to expose the strengths and weaknesses of both formulations. In both numerical implementations, three noded isoparametric elements allowing quadratic variation of tractions and displacements are used.

Lastly, a complicated multi-region problem is considered in order to show the complete generality of the particular integral method for solving elastostatic body force problems.

Boundary Element Formulations Using Particular Integrals

Particular Integrals. Although the boundary element solution of elastostatic problems using the method of particular integrals has been tentatively discussed by Watson (1979) as well as Banerjee and Butterfield (1981), they have never been fully developed. Jaswon and Maiti (1968) appear to be the first to apply this concept in a boundary element analysis by solving some simple plate bending problems. The method of particular integrals is well known in the solution of inhomogeneous differential equations. For example, the governing differential equations for the deformation of a homogeneous isotropic body with any system of body forces are given by

$$(\lambda+\mu)\frac{\partial^2 u_j}{\partial x_i \partial x_j}+\mu\frac{\partial^2 u_i}{\partial x_j \partial x_j}+\psi_i=0 \qquad (1)$$

where λ and μ are Lamés constants, ψ_i are the body forces, i,

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August 15, 1986; final revision, April 2, 1987.

Paper No. 87-WA/APM-28.

$j=1, 2$ for two-dimensional problems, u_i is the displacement vector referred to a Cartesian coordinate system x_i, and the summation convention is implied. Equation (1) is a linear inhomogeneous differential equation of the form

$$L(u)+\psi_i=0 \tag{2}$$

where $L(u)$ is a self-adjoint homogeneous differential operator such as the first two terms of equation (1) and ψ_i is the known inhomogeneous part as in equation (1). Since this is a linear equation, its solution can be represented as the sum of a complementary solution u_i^c satisfying

$$L(u^c)=0 \tag{3a}$$

and a particular integral u_i^p satisfying

$$L(u^p)+\psi_i=0 \tag{3b}$$

The total solution is then simply

$$u_i=u_i^c+u_i^p \tag{4}$$

The following observations are important as far as the particular integrals are concerned:

(*i*) They can often be obtained by inspection of the inhomogeneous differential equation or by the method of undetermined coefficients.

(*ii*) There are no unique particular integrals, since any polynomial satisfying (3*b*) or any linear combination of particular integrals is a valid particular integral.

For gravitational and centrifugal loading one can obtain particular integrals from known solutions as illustrated below.

Particular Integrals for Gravity Loading. If we consider the case of gravitational acceleration directed along the x_2 axis, the body force components are given as

$$\psi_1=0 \tag{5a}$$

$$\psi_2=-\rho g \tag{5b}$$

where ρ is the mass density and g is the acceleration due to gravity. One set of particular integrals for equation (1) corresponding to this set of body forces are given, for instance, by Sokolnikoff (1956):

$$u_1^p=\frac{-\rho g}{4\mu(\lambda+\mu)}\lambda x_1 x_2 \tag{6a}$$

$$u_2^p=\frac{\rho g}{8\mu(\lambda+\mu)}[(\lambda+2\mu)x_2^2+\lambda x_1^2] \tag{6b}$$

which gives the following stress components:

$$\sigma_{11}^p=0 \tag{6c}$$

$$\sigma_{12}^p=0 \tag{6d}$$

$$\sigma_{22}^p=\rho g x_2 \tag{6e}$$

And the associated tractions are given by

$$t_i^p=\sigma_{ij}^p n_j \tag{7}$$

or:

$$t_1^p=0 \tag{8a}$$

$$t_2^p=\rho g x_2 n_2 \tag{8b}$$

Particular Integrals for Centrifugal Loading. For stresses due to the rotation of a body about a fixed axis located at the origin of a Cartesian coordinate system, the body forces are given by

$$\psi_r = \rho\omega^2 r$$

$$\psi_\theta = 0$$

where ω is the angular velocity, r is the radial distance to the point under consideration, and ρ is again the mass density. Converting to Cartesian coordinates, we get

$$\psi_1=\psi_r\cos\theta-\psi_\theta\sin\theta=\rho\omega^2 x_1 \tag{9a}$$

$$\psi_2=\psi_r\sin\theta+\psi_\theta\cos\theta=\rho\omega^2 x_2 \tag{9b}$$

or, in indicial notation,

$$\psi_i=\rho\omega^2 x_i \tag{10}$$

The particular integrals for this loading are again given by Sokolnikoff (1956) as:

$$u_i^p=\frac{-\rho\omega^2}{8(\lambda+2\mu)}(x_n x_n)x_i \tag{11}$$

The stress components are given by:

$$\sigma_{ij}^p=\frac{-\rho\omega^2}{4(\lambda+2\mu)}[(2\lambda+\mu)x_n x_n\delta_{ij}+2\mu x_i x_j] \tag{12}$$

where δ_{ij} is the Kronecker delta and the summation convention is implied. Using equation (7), we find the tractions as

$$t_i^p=\frac{-\rho\omega^2}{4(\lambda+2\mu)}[(2\lambda+\mu)x_n x_n n_i+2\mu x_i(x_\ell n_\ell)] \tag{13}$$

The final solution satisfying all of the real boundary conditions can be found by simply combining the complementary and particular solution

$$u_i=u_i^c+u_i^p \tag{14a}$$

$$t_i=t_i^c+t_i^p \tag{14b}$$

$$\sigma_{ij}=\sigma_{ij}^c+\sigma_{ij}^p \tag{14c}$$

Thus, it is only necessary to construct boundary integral formulations for the complementary solution and superimpose the particular integral. The particular integrals given above are not dependent on the problem geometry. Since they are completely general, they are independent of the boundary conditions of the problem and need only to be evaluated at each boundary point, with the required integrations the same as for a homogeneous problem. The details of how this is implemented for the DBEM and IBEM are given in the following sections.

Direct Formulation

This section provides a brief summary of the mathematical foundations of the DBEM. The interested reader may find a more detailed description in, for instance, Banerjee and Butterfield (1981). By using the reciprocal work theorem on the Kelvin singular solution and the complementary solution, one has for the displacement at an interior point ξ due to any admissible combination of tractions (t_i) and displacements (u_i) over the surface (s):

$$u_j^c(\xi)=\int_S [G_{ij}(x,\xi)t_i^c(x) - F_{ij}(x,\xi)u_i^c(x)]ds(x) \tag{15}$$

The kernels $G_{ij}(x,\xi)$ and $F_{ij}(x,\xi)$, where x is the load point and ξ is the field point, are given explicitly in Appendix A. We have used the superscript c to emphasize that this is the complementary solution, i.e., there are no body forces present.

By taking the load point to the boundary of the body, we obtain a constraint equation relating the complementary parts of the known and unknown displacements and tractions on the surface of the body:

$$C_{ij}u_i^c(\xi)=\int_S [G_{ij}(x,\xi)t_i^c(x) - F_{ij}(x,\xi)u_i^c(x)]ds(x) \tag{16}$$

where $C_{ij}=1/2\ \delta_{ij}$ for a point on a smooth boundary.

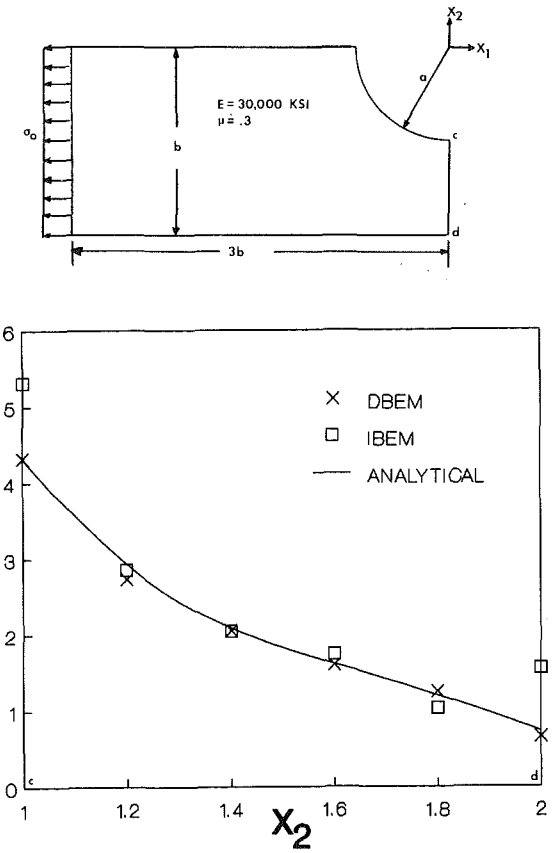


Fig. 1 Plate with circular hole

Using the strain-displacement and stress-strain relations, the stress state at any point within the body may be found as

$$\sigma_{ij}^c(\xi) = \int_S [T_{kij}(x, \xi) t_k^c(x) - E_{kij}(x, \xi) u_k^c(x)] ds(x) \tag{17}$$

The kernels T_{kij} and E_{kij} are listed in Appendix A. The stresses at a surface point can be obtained from the boundary tractions and displacements without any integration as was originally proposed by Cruse (1973).

The usual procedure for solving equation (16) is to discretize the boundary into a number of boundary elements. In doing so, we may write this equation in matrix form as

$$[G]\{t^c\} - [F]\{u^c\} = \{0\} \tag{18}$$

Recalling equations (14*a*) and (14*b*), which give the real solution as the sum of the complementary and particular solutions, we have

$$[G]\{t - t^p\} - [F]\{u - u^p\} = \{0\} \tag{19}$$

Placing all the known quantities on the right-hand side we get

$$[A]\{X\} = \{b\} + [G]\{t^p\} - [F]\{u^p\} \tag{20}$$

where b is a vector obtained from the real known tractions and displacements, A is a coefficient matrix, and X is a vector containing the unknown tractions and displacements. This equation can be solved for the only unknown, $\{X\}$. If stresses are required, we can use equation (17) discretized as

$$\{\sigma_{ij}\}^c = [T]\{t^c\} - [E]\{u^c\} \tag{21}$$

and we can use equation (14*c*) to get

$$\{\sigma_{ij}\} = [T]\{t^c\} - [E]\{u^c\} + \{\sigma_{ij}\}^p \tag{22}$$

where σ_{ij}^p is given by equation (6) for the gravitational or equation (12) for the centrifugal load problem.

For a multiregion problem, equation (20) can be formed for each subregion of the system and expressed as

$$[A]^r\{X\}^r = \{b\}^r$$

where r represents the region under consideration. These are then assembled to form the final system equations by satisfying equilibrium and compatibility at the common interfaces as described, for instance, in Banerjee and Butterfield (1981).

Indirect Formulation

In the indirect formulation, the displacement and traction at an interior point x are given in terms of a fictitious surface traction ϕ as:

$$u_i^c(x) = \int_S G_{ij}(x, \xi)\phi_j(\xi) ds(\xi) + C_i \tag{23}$$

$$t_i^c(x) = \int_S F_{ij}(x, \xi)\phi_j(\xi) ds(\xi) \tag{24}$$

where the C_i are unknown rigid body displacements which arise from the arbitrary nature of r_o in equation ($A.2$).

Again we use the superscript c to denote that this is the complementary solution. By bringing the field point x to the boundary, we obtain the following equations relating the tractions and displacements

$$u_i^c(x_o) = \int_S G_{ij}(x_o, \xi)\phi_j(\xi) ds(\xi) \tag{25}$$

$$t_i^c(x_o) = \pm \frac{1}{2}\delta_{ij}\phi_j(x_o) + \int_S F_{ij}(x_o, \xi)\phi_j(\xi) ds(\xi) \tag{26}$$

Equations (25) and (26) are now integrated along the boundary of the body, which is divided into boundary elements for this purpose. The resulting equations can be represented in matrix form as:

$$\{u^c\} = [G]\{\phi\} \tag{27a}$$

$$\{t^c\} = [F]\{\phi\} \tag{27b}$$

For a well posed problem, the number of known tractions and displacements will be exactly enough to find the unkown vector ϕ. This is accomplished by combining the equations as follows:

$$\{Y^c\} = [A]\{\phi\} \tag{28}$$

where $[A]$ is a system matrix and the individual components of the left-hand side can be obtained using equations (14*a*) and (14*b*)

$$\begin{Bmatrix} u^c \\ t^c \end{Bmatrix} = \begin{Bmatrix} u \\ t \end{Bmatrix} - \begin{Bmatrix} u^p \\ t^p \end{Bmatrix} \tag{29}$$

or

$$\{Y^c\} = \{Y\} - \{Y^p\} \tag{30}$$

where Y is a vector of given boundary tractions and displacements, Y^p contains the contribution from particular integrals computed at each point, and Y^c is a vector of "complementary" tractions and displacements. Equation (28) can now be solved for the ϕ vector. Having found ϕ, the unkown tractions and displacements can be obtained using equations (27*a*) and (27*b*) together with (14*a*) and (14*b*). The stresses at boundary points are then obtained from the tractions t_i, the displacements u_i, and the constitutive and the strain displacement relations in the usual manner (Banerjee and Butterfield, 1981; Cruse, 1973). The stresses at any in-

Table 1

	SCF	Percent error	Run time (s)
Analytical	4.32	--	--
IBEM	5.31	22.9	42.2
DBEM	4.32	0.0	34.3

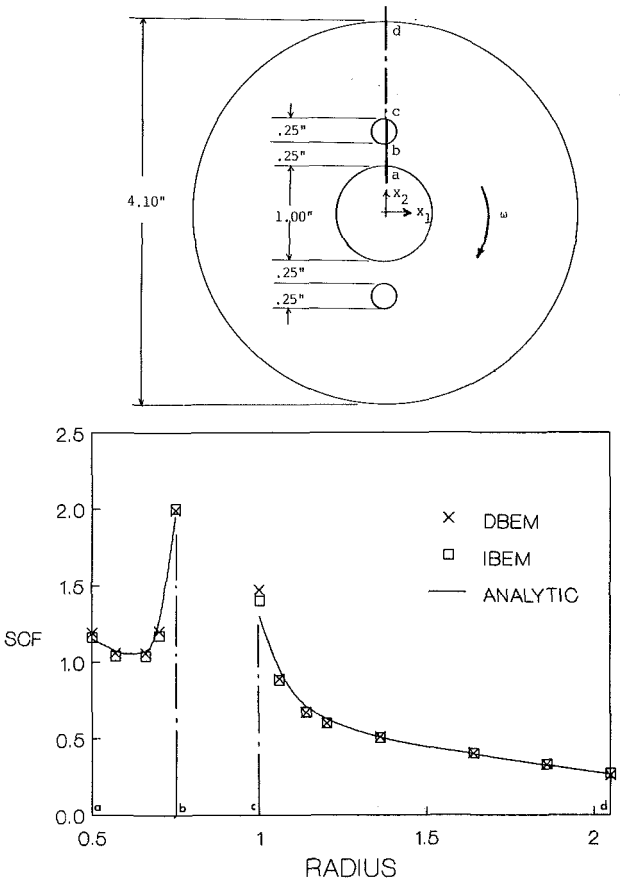


Fig. 2 Rotating disk with eccentric holes

terior point, if required, can be obtained from the discretized version of the equation

$$\sigma_{ij}(x) = \int_S T_{ijk}(x, \xi)\phi_k(\xi)ds(\xi) + \sigma^0_{ij}(x) \tag{31}$$

where $T_{ijk}(x, \xi)$ is defined explicitly in Appendix A. An excellent discussion of the treatment of the surface integral involved in equation (31) was given recently by Cox and Shugar (1985).

Numerical Examples

A Plate with a Circular Hole. Consider a plate in plane stress of width $2b$ with a centrally placed unloaded circular hole of radius a under uniform tension as shown in Fig. 1. Because of the symmetry in the problem, only a quarter of the geometry need be discretized where for comparison purposes, we take $a=1$, $b=2$. The plate is modeled with 16 quadratic elements for a total of 32 nodes. The nodes on the x_2 axis are constrained against displacing in the x_1 direction while those on the x_1 axis are fixed in the x_2 direction. A uniform traction σ_o is applied to the nodes at the far end of the model.

The analytical solution to this problem is found in Savin (1968). The stress distribution along the line of symmetry $c-d$ is plotted in Fig. 1 and the stress concentration factor at point c is compared to the accepted value in Table 1 where the run times on an HP-9000 desk-top minicomputer are also shown.

Table 2

	Maximum SCF	Percent error	Run time (s)
Analytical	1.95	--	--
IBEM	2.00	2.5	130.5
DBEM	1.99	2.0	100.6

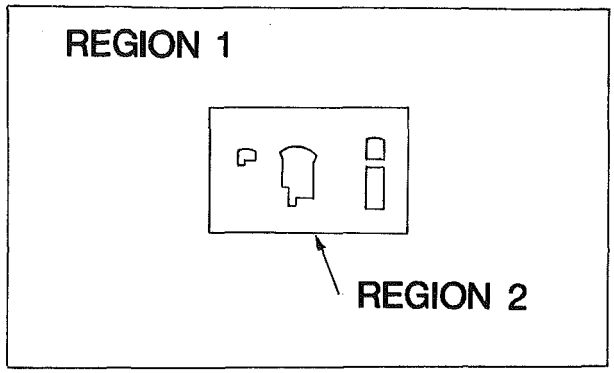


Fig. 3 Boundary element model of underground power station

As seen in Fig. 1, both boundary element methods give accurate results on the majority of the section, but the IBEM results, as is often the case, are poor near the sharp corners. Accurate results for point c could be obtained also with the IBEM by running the full geometry, thus eliminating the corners on the hole boundary.

Rotating Disk with Eccentric Holes. The problem of the stress concentration in a rotating disk with a central hole as well as two eccentric holes was investigated experimentally first by Hetényi (1939), and later theoretically by Ku (1960). The problem they considered is shown in Fig. 2.

Because of the symmetry of the problem, only one-half of the body is modeled. The discretization consists of 20 quadratic elements and 40 nodes. Again, nodes on the x_2 axis are constrained against displacement in the x_1 direction while nodes on the x_1 axis are constrained in the x_2 direction. The only loading applied is that due to rotation.

The analytic solution is given by Ku (1960) in the form of a stress concentration factor (SCF) which is defined as the ratio of the circumferential stress at any point on the disk with the two noncentral holes to the circumferential stress at the inner boundary of a rotating disk with the central hole only. For comparison purposes, the stress is converted to an SCF by dividing by the proper circumferential stress for a disk with a central hole only (as computed using Timoshenko and Goodier's, 1970, solution).

The SCF along a section through the holes is plotted in Fig. 2. The maximum SCF, at the inner edge of the small holes, is given in Table 2, along with the run times for comparisons.

Note again in this case both methods provide accurate results for interior points. Additionally, since the point of interest does not lie on a corner, both the IBEM and DBEM give excellent results.

Underground Power Station. As a final example we consider the self-weight stresses in the underground power station shown in Fig. 3. Since integrating from the larger distant outer elements to the smaller elements around the openings would reduce the accuracy and lower the efficiency, the problem is divided into two subregions. The DBEM with a total of 144 nodes and 72 isoparametric elements as well as 64 internal stress points are used to solve this plane strain problem. The nondimensionalized equivalent stress is plotted in Fig. 4. A similar problem is analyzed using the finite element method in Zienkiewicz (1977), although a direct comparison is not possible since detailed geometric and material data are not given.

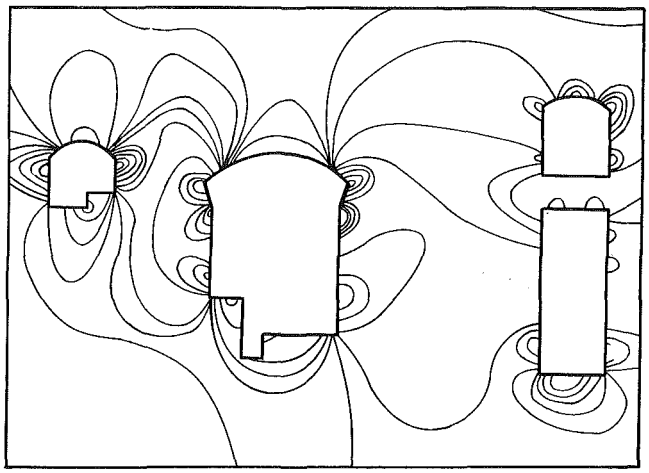


Fig. 4 Nondimensional equivalent stress: Region 2

Conclusion

The method of particular integrals proves to be an efficient and accurate method of calculating stresses due to body forces. This method, based on a theory of differential equations, while certainly not new, has never been fully explored for elastostatics with the BEM. The complementary solutions are obtained using boundary only discretization, while the particular solutions need only be evaluated at each boundary point. Thus, the use of the method of particular integrals has succeded in eliminating any extra integration involving the body force terms and has been demonstrated for complicated, multi-region body force problems.

Advanced formulations of the direct and indirect boundary element methods have been compared for a range of problems. In general, the results correlate well with accepted results. The DBEM formulation runs faster than the IBEM on an HP-9000 computer.

Acknowledgment

This research has been carried out at the State University of New York at Buffalo as part of an industry/university cooperative research project to study the characteristics of Harmonic Shapes for Optimum Design supported by National Science Foundation Grant No. CEE-8310956, with Drs. R. Richards, Jr. and G. S. Bjorkman, Jr. as Coprincipal Investigators.

References

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

Cox, J. V., and Shugar, T. A., 1985, "A Recursive Integration Technique for BEM in Elastostatics," Cruse, Pifko, and Armen, eds., *AMD*, Vol. 72, American Society of Mechanical Engineers.

Cruse, T. A., and Wilson, R. B., 1977, "Boundary Integral Equation Method for Elastic Fracture Mechanics Analysis," AFOSR-TR-78-0355, Pratt and Whitney Aircraft.

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

Cruse, T. A., 1973, "An Improved Boundary Integral Equation Method for Three-Dimensional Elastic Stress Analysis," *Int. Jour. Computers and Structures*, Vol. 4, pp. 741-757.

Danson, D. J., 1982, "A Boundary Element Formulation of Problems in Linear Isotropic Elasticity with Body Forces," *Proc. 3rd Int. Seminar on Recent Advances in Boundary Element Methods*, Irvine, CA, pp. 105-122.

Hetényi, M., 1939, "The Application of Hardening Resins to Three-Dimensional Photoelasticity," *Journal of Applied Physics*, Vol. 10, pp. 295-300.

Jaswon, M. A., and Maiti, M., 1968, "An Integral Equation Formulation of Plate Bending Problems," *Journal of Eng. Math.*, Vol. 2, pp. 83-93.

Kamiya, N., and Sawaki, Y., 1985, "An Efficient BEM for Some Inhomogeneous and Nonlinear Problems," *Proc. 7th Int. Conf. on Boundary Element Methods*, Lake Como, Italy, pp. 13-59 to pp. 13-68.

Ku, Ta-Cheng, 1960, "Stress Concentration in a Rotating Disk with a Central Hole and Two Additional Symmetrically Located Holes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 27, pp. 359-360.

Mukherjee, S., 1982, *Boundary Element Methods in Creep and Fracture*, Applied Science Publishers, London.

Nardini, D., and Brebbia, C. A., 1982, "A New Approach to Free Vibration Analysis using Boundary Elements," *Proc. of 4th Int. Seminar*, Southampton, England, pp. 313-326.

Rizzo, F. J., and Shippy, D. J., 1977, "An Advanced Boundary Integral Equation Method for Three-Dimensional Thermoelasticity," *International Journal for Numerical Methods in Engineering*, Vol. 11, pp. 1753-1768.

Savin, G. N., 1968, *Stress Distribution Around Holes*, NASA TT F-607, Naukova Dumka Press.

Sokolnikoff, I. S., 1956, *Mathematical Theory of Elasticity*, 2nd Ed., McGraw-Hill, New York.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd Ed., McGraw-Hill, New York.

Watson, J. O., 1979, "Advanced Implementation of the Boundary Element Method for Two and Three-Dimensional Elastostatics," *Developments in BEM*, Vol. 1, Banerjee, P. K., and Butterfield, R., eds., Applied Science Publishers, London, Chapter 3, pp. 31-63.

Zienkiewicz, O. C., 1977, *The Finite Element Method*, 3rd Ed., McGraw-Hill.

APPENDIX A

The fundamental singular solution for the displacement field in a two-dimensional elastostatic problem is given by:

$$u_i(x) = G_{ij}(x, \xi) e_j(\xi) \tag{A.1}$$

where

$$G_{ij}(x, \xi) = C_1 \left(C_2 \delta_{ij} \ln r - \frac{y_i y_j}{r^2} \right) + A_{ij} \tag{A.2}$$

and

$$C_1 = \frac{-1}{8\pi\mu(1-\nu)}$$

$$C_2 = 3 - 4\nu$$

A_{ij} = arbitrary constant tensor which can be determined by specifying that at any distance from the load point the displacements are zero

$$y_i = x_i - \xi_i$$

$$y_j = x_j - \xi_j$$

$$r^2 = y_i y_i$$

The stress state corresponding to the above displacement field can be deduced from the strain-displacement and stress-strain relations as:

$$\sigma_{ij}(x) = T_{ijk}(x, \xi) e_k(\xi) \tag{A.3}$$

The surface tractions at a point with an outward normal n_j are calculated from:

$$t_i(x) = \sigma_{ij}(x) n_j(x) = F_{ik}(x, \xi) e_k(\xi) \tag{A.4}$$

where

$$F_{ik}(x, \xi) = \left(\frac{C_3}{r^2} \right) \left[C_4 (n_k y_i - n_i y_k) + \left(C_4 \delta_{ik} + \frac{2 y_i y_k}{r^2} \right) y_j n_j \right]$$

and the stress kernels are:

$$T_{kij}(x, \xi) = \frac{a_1}{r} \left[\frac{a_2}{r} \left(\delta_{ik} y_j + \delta_{jk} y_i - \delta_{ij} y_k \right) + \frac{2 y_i y_j y_k}{r^3} \right] \tag{A.5}$$

$$E_{kij}(x,\,\xi)=\frac{a_3}{r^2}\left[\frac{n_\ell y_\ell}{r^2}\left\{2a_2\delta_{ij}y_k+2\nu(\delta_{ik}y_j+\delta_{jk}y_i)-\frac{8y_iy_jy_k}{r^2}\right\}+n_i\left(2\nu\,\frac{y_jy_k}{r^2}+a_2\delta_{jk}\right)+n_j\left(2\nu\,\frac{y_iy_k}{r^2}+a_2\delta_{ik}\right)+n_k\left(2a_2\,\frac{y_iy_j}{r^2}-a_4\delta_{ij}\right)\right] \tag{A.6}$$

with:

$$a_1 = \frac{1}{4\pi(1-\nu)}$$

$$a_2 = 1-2\nu$$

$$a_3 = \frac{\mu}{2\pi(1-\nu)}$$

$$a_4 = 1-4\nu$$

Analysis of Orthogonally Cracked Laminates Under Tension

Z. Hashin
Dept. of Solid Mechanics, Materials and Structures, Faculty of Engineering, Tel-Aviv University, Tel-Aviv, Israel 69978
Fellow ASME

The problems of stiffness reduction and stress analysis of cross-ply fiber composite laminates, where all plies are cracked in fiber directions, are treated by a variational method on the basis of the principle of minimum complementary energy. The Young's modulus obtained is a strict lower bound but is expected to be close to the true value on the basis of experience with a previous analysis. Approximate values of Poisson's ratio and internal stresses have been obtained. The latter reveal important tendencies of continued failure by delamination.

Introduction

The major damage which develops in laminates under static or cyclic loading is in the form of interlaminar and intralaminar cracks. The former develop gradually and slowly in between plies. The latter appear suddenly and in large numbers in plies in which the stresses reach critical values, perhaps defined by the first failure criteria of the plies. They are families of parallel cracks in fiber direction and their macroscopic effect is reduction of the inplane stiffness of the laminate.

The subject of the analysis of stiffness reduction of cracked laminates has received repeated attention but has mainly been concerned with cross-ply laminates, i.e., $[0_m^\circ/90_n^\circ]_s$ configurations in which only the 90° plies are cracked. The main methods of analysis employed are: a simple shear-lag method (Reifsnider and Talug, 1980; Reifsnider and Jamison, 1982), self-consistent approximation to assess ply stiffness reduction in conjunction with classical laminate analysis (Laws et al., 1983, 1985), and a variational method (Hashin, 1985). All of these methods give results which are in good to excellent agreement with experimental data. The last method, unlike the others, also provides useful estimates of the internal stresses in the cracked laminate.

In many cases of damage in laminates, several plies will be cracked and in particular adjacent plies. As a typical and relatively simple case we shall consider the problem of a $[0_m^\circ/90_n^\circ]_s$ laminate in which all plies have intralaminar cracks and which is thus orthogonally cracked. Because of the complicated interaction of the orthogonal cracks this problem is much more difficult than the one with plies cracked in only one direction. We are aware of only one publication in the literature (Highsmith and Reifsnider, 1986) on this subject, which is concerned with evaluation of stresses in a typical repeating cracked laminate element by numerical analysis.

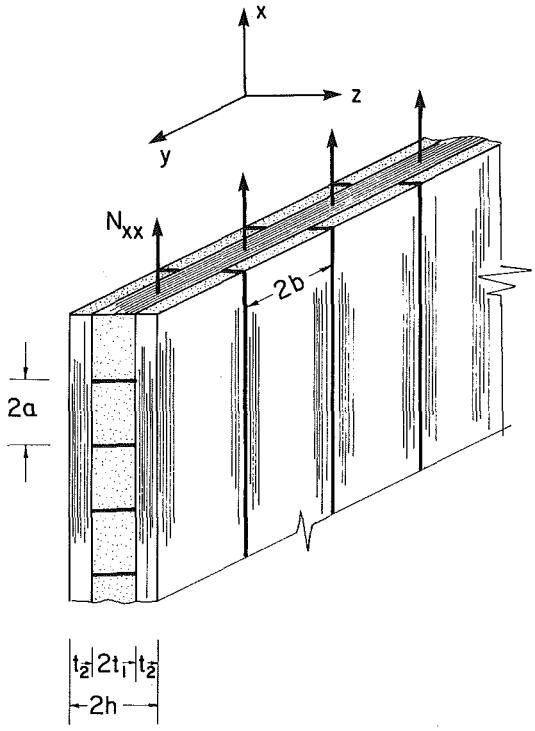


Fig. 1 Orthogonally cracked laminate

The present problem is closely related to that of a cracked ±45° symmetric laminate which is indeed merely a rotated 0°/90°. It is easily seen that uniaxial tension of a $[+45_m^\circ/-45_n^\circ]_s$ in bisector direction can be analyzed in terms equibiaxial tension and shear of a $[0_m^\circ/90_n^\circ]_s$ laminate. Indeed a typical case of orthogonal cracking is encountered for cyclic loading of a ±45° laminate. The problem of shearing of a cross-ply will be considered elsewhere. Here we are concerned with uniaxial tensile loading of an orthogonally cracked cross-ply. Our purpose is to evaluate stiffness reduction due to cracks and approximate local stresses. We shall do this by generalization of the variational method which has been developed in Hashin (1985).

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 2, 1987.

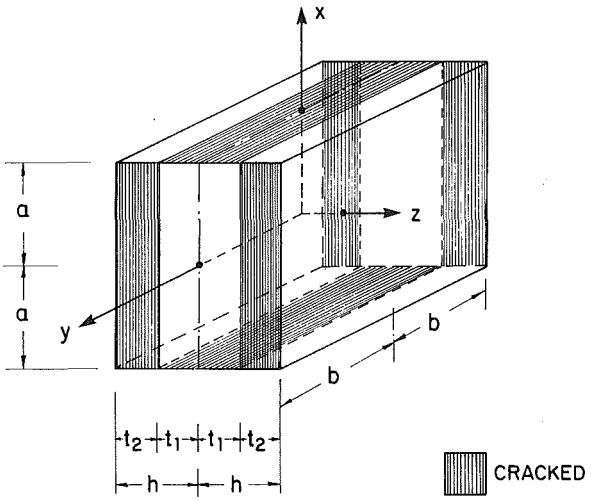


Fig. 2 Repeating element

Admissible Field Construction

Consider an orthogonally cracked laminate, Fig. 1, and a typical laminate element defined by intersecting crack pairs, Fig. 2. Let the laminate be subjected to constant membrane force

$$N_{xx} = 2\sigma^o h \tag{1}$$

and no other loads. Thus σ^o is the average stress $\bar{\sigma}_{xx}$ over the laminate and also over the thickness $2h$.

In the event that longitudinal and transverse crack families are each equidistant, the stresses and strains in each element (Fig. 2) are the same, the boundary and interface conditions of the repeating element can be formulated and the problem can be treated numerically in terms of finite elements, if desired. The problems associated with such an approach will be discussed at the end of the paper. Here we shall proceed differently; it is our purpose to construct admissible stress fields which satisfy equilibrium and all traction, boundary and interface conditions. These admissible fields will then be optimized in the context of the principle of minimum complementary energy to yield approximate stresses and lower bounds on stiffness.

The stresses in the undamaged laminate under the loading considered are

$$\begin{aligned} \sigma_{xx}^{o(1)} &= \sigma_x^{(1)} = k_x^{(1)}\sigma^o \qquad \sigma_{yy}^{o(1)} = \sigma_y^{(1)} = k_y^{(1)}\sigma^o \\ \sigma_{xx}^{o(2)} &= \sigma_x^{(2)} = k_x^{(2)}\sigma^o \qquad \sigma_{yy}^{o(2)} = \sigma_y^{(2)} = k_y^{(2)}\sigma^o \\ \sigma_{xy}^{o(1)} &= \sigma_{xy}^{o(2)} = 0 \end{aligned} \tag{2}$$

where the k coefficients are easily found from conventional laminate analysis.

Let $\tilde{\sigma}_{ij}$ be a three-dimensional admissible stress field within the cracked laminate. By definition such a stress field must satisfy equilibrium, traction continuity, and traction boundary conditions. For reasons of symmetry it is sufficient to consider one half of the element of Fig. 2 defined by $-a \leqq x \leqq a$, $-b \leqq y \leqq b$, $0 \leqq z \leqq h$, as will be understood from now on. We list required boundary and interface conditions

$$\tilde{\sigma}_{xz}^{(1)}(x,y,0) = 0 \qquad \tilde{\sigma}_{yz}^{(1)}(x,y,0) = 0 \tag{3}$$

$$\tilde{\sigma}_{xz}^{(2)}(x,y,h) = 0 \qquad \tilde{\sigma}_{yz}^{(2)}(x,y,h) = 0 \tag{4}$$

$$\tilde{\sigma}_{zz}^{(2)}(x,y,h) = 0$$

$$\begin{aligned} \tilde{\sigma}_{xz}^{(1)}(x,y,t_1) &= \tilde{\sigma}_{xz}^{(2)}(x,y,t_1) \\ \tilde{\sigma}_{yz}^{(1)}(x,y,t_1) &= \tilde{\sigma}_{yz}^{(2)}(x,y,t_1) \\ \tilde{\sigma}_{zz}^{(1)}(x,y,t_1) &= \tilde{\sigma}_{zz}^{(2)}(x,y,t_1) \end{aligned} \tag{5}$$

$$\begin{aligned} &\tilde{\sigma}_{xx}^{(1)}(\pm a,y,z) = \tilde{\sigma}_{xy}^{(1)}(\pm a,y,z) = \tilde{\sigma}_{xz}^{(1)}(\pm a,y,z) = 0 \\ &\tilde{\sigma}_{xy}^{(2)}(x,\pm b,z) = \tilde{\sigma}_{yy}^{(2)}(x,\pm b,z) = \tilde{\sigma}_{yz}^{(2)}(x,\pm b,z) = 0 \end{aligned} \tag{6}$$

Here and from now on superscripts 1, 2 indicate the plies, equations (3) are symmetry conditions, equations (4) are free surface conditions, equations (5) express traction interface continuity, and equations (6) are zero traction conditions on the crack surface. There are still needed traction conditions on the faces $x = \pm a$, $t_1 \leqq z \leqq h$ and $y = \pm b$, $0 \leqq z \leqq t_1$ which shall be considered further below.

The xy in-plane parts of the admissible stress fields are chosen in the form

$$\tilde{\sigma}_{xx}^{(1)} = \sigma_x^{(1)}[1 - \phi_1(x)] \qquad \tilde{\sigma}_{xx}^{(2)} = \sigma_x^{(2)}[1 - \phi_2(x)] \tag{7a}$$

$$\tilde{\sigma}_{yy}^{(1)} = \sigma_y^{(1)}[1 - \psi_1(y)] \qquad \tilde{\sigma}_{yy}^{(2)} = \sigma_y^{(2)}[1 - \psi_2(y)] \tag{7b}$$

$$\tilde{\sigma}_{xy}^{(1)} = \tilde{\sigma}_{xy}^{(2)} = 0 \tag{7c}$$

where ϕ and ψ are unknown functions. The physical significance of these assumptions will be discussed further below.

Force equilibrium of the undamaged laminate in the x and y directions requires

$$\begin{aligned} \sigma_x^{(1)} t_1 + \sigma_x^{(2)} t_2 &= \sigma^o h \\ \sigma_y^{(1)} t_1 + \sigma_y^{(2)} t_2 &= 0 \end{aligned} \tag{8}$$

Examining force equilibrium in the same directions in the cracked laminate, in terms of the admissible stresses (7), we find

$$\begin{aligned} \sigma_x^{(1)} t_1 \phi_1(x) + \sigma_x^{(2)} t_2 \phi_2(x) &= 0 \\ \sigma_y^{(1)} t_1 \psi_1(y) + \sigma_y^{(2)} t_2 \psi_2(y) &= 0 \end{aligned} \tag{9}$$

This reduces the number of unknown functions from four to two.

In view of equation (7c), the equilibrium equations for the admissible ply stresses assume the form

$$\begin{aligned} \tilde{\sigma}_{xx,x} + \tilde{\sigma}_{xz,z} &= 0 \\ \tilde{\sigma}_{yy,y} + \tilde{\sigma}_{yz,z} &= 0 \\ \tilde{\sigma}_{xz,x} + \tilde{\sigma}_{yz,y} + \tilde{\sigma}_{zz,z} &= 0 \end{aligned} \tag{10}$$

where commas denote partial differentiation.

We now define perturbation stresses $\sigma_{ij}^{(m)}$ for the plies, $m = 1, 2$, by

$$\tilde{\sigma}_{ij}^{(m)} = \sigma_{ij}^{o(m)} + \sigma_{ij}^{(m)} \tag{11}$$

where $\sigma_{ij}^{o(m)}$ are given by the undamaged laminate stresses (2). Insertion of equations (7) into (10) for each ply and systematic integration, with elimination of unknown residual functions by use of equations (3)–(5), and utilization of equations (9), leads to the results

$$\sigma_{xx}^{(1)} = -\sigma_x^{(1)}\phi(x) \tag{12a}$$

$$\sigma_{xx}^{(2)} = \frac{1}{\lambda}\,\sigma_x^{(1)}\phi(x) \tag{12f}$$

$$\sigma_{xz}^{(1)} = \sigma_x^{(1)}\phi'(x)z \tag{12b}$$

$$\sigma_{xz}^{(2)} = \frac{1}{\lambda}\,\sigma_x^{(1)}\phi'(x)(h-z) \tag{12g}$$

$$\sigma_{yz}^{(1)} = \sigma_y^{(1)}\dot{\psi}(y)z \tag{12d}$$

$$\sigma_{yz}^{(2)} = \frac{1}{\lambda}\,\sigma_y^{(1)}\dot{\psi}(y)(h-z) \tag{12i}$$

$$\sigma_{zz}^{(1)}=[\sigma_x^{(1)}\phi''(x)+\sigma_y^{(1)}\ddot{\psi}(y)]\frac{1}{2}(ht_1-z^2) \tag{12e}$$

$$\sigma_{zz}^{(2)}=\frac{1}{\lambda}[\sigma_x^{(1)}\phi''(x)+\sigma_y^{(1)}\ddot{\psi}(y)]\frac{1}{2}(h-z)^2 \tag{12j}$$

where prime and dot indicate x and y derivative, respectively, and

$$\phi=\phi_1 \qquad \psi=\psi_1$$
$$\lambda=t_2/t_1 \tag{13}$$

It follows from equations (6) and (9) that ϕ and ψ must satisfy the boundary conditions

$$\phi(\pm a)=1 \qquad \phi'(\pm a)=0$$
$$\psi(\pm a)=1 \qquad \dot{\psi}(\pm b)=0 \tag{14}$$

The above formulation determines the traction components on the faces $x = \pm a$, $t_1 \leqq z \leqq h$ and $y = \pm b$, $0 \leqq z \leqq t_1$. It is seen that the only nonvanishing traction component is

$$\tilde{\sigma}_{xx}^{(2)}(\pm a)=\sigma_x^{(1)}(1+t_1/t_2) \tag{15}$$

Thus the admissible tractions are continuous across all common faces of adjacent elements, regardless of their a, b dimensions. Therefore, the stresses (11)–(12) rigorously satisfy all conditions of admissibility.

Variational Formulation

Next we evaluate the complementary energy functional $\tilde{U}_c$ associated with the admissible stresses. In the present case where tractions are prescribed everywhere

$$\tilde{U}_c=\frac{1}{2}\int_V S_{ijkl}\tilde{\sigma}_{ij}\tilde{\sigma}_{kl}dV \tag{16}$$

where V is laminate volume and S_{ijkl} are local compliances. It has been shown in Hashin (1985) that for *any* cracked body

$$\tilde{U}_c=U_c^o+\tilde{U}_c' \tag{17}$$

where

$$U_c^o=\frac{1}{2}\int S_{ijkl}\sigma_{ij}^o\sigma_{kl}^o dV \tag{18a}$$

$$\tilde{U}_c'=\frac{1}{2}\int S_{ijkl}\sigma_{ij}'\sigma_{kl}'dV \tag{18b}$$

$$\tilde{\sigma}_{ij}=\sigma_{ij}^o+\sigma_{ij}' \tag{18c}$$

Here σ_{ij}^o are the actual stresses in the uncracked body, i.e., in the undamaged laminate; thus U_c^o is the actual stress energy of the undamaged laminate. In the present case it is rigorously true that

$$U_c^o=\frac{{\sigma^o}^2}{2E_x^o}V \tag{19}$$

where E_x^o is the Young's modulus in the x direction of the undamaged laminate. Furthermore, the stress energy U_c of the cracked laminate can be rigorously expressed in the form

$$U_c=\frac{{\sigma^o}^2}{2E_x}V \tag{20}$$

where E_x is the effective Young's modulus of the *cracked* laminate. Also from the principle of minimum complementary energy

$$U_c \leqq \tilde{U}_c \tag{21}$$

Combining equations (17) and (19)–(21) we have

$$\frac{1}{E_x}\leqq\frac{1}{E_x^o}+\frac{2\tilde{U}_c'}{V{\sigma^o}^2} \tag{22}$$

which provides a lower bound on E_x.

In the present case the stresses (12) are identified with the σ_{ij}' in equation (18c). In order to evaluate equation (18b) we list the local stress energy densities in the plies as referred to the common x, y, z system of the laminate

$$\begin{aligned}2W^{(1)}&=\sigma_{yy}^{(1)2}/E_A+(\sigma_{xx}^{(1)2}+\sigma_{zz}^{(1)2})/E_T-\sigma_{yy}^{(1)}(\sigma_{xx}^{(1)}\\&+\sigma_{zz}^{(1)})2\nu_A/E_A-\sigma_{xx}^{(1)}\sigma_{zz}^{(1)}2\nu_T/E_T\\&+(\sigma_{xy}^{(1)2}+\sigma_{yz}^{(1)2})/G_A+\sigma_{xz}^{(1)2}/G_T\\2W^{(2)}&=\sigma_{xx}^{(2)2}/E_A+(\sigma_{yy}^{(2)2}+\sigma_{zz}^{(2)2})/E_T-\sigma_{xx}^{(2)}(\sigma_{yy}^{(2)}\\&+\sigma_{zz}^{(2)})2\nu_A/E_A-\sigma_{yy}^{(2)}\sigma_{zz}^{(2)}2\nu_T(E_T\\&+(\sigma_{xy}^{(2)2}+\sigma_{xz}^{(2)2})/G_A+\sigma_{yz}^{(2)2}/G_T\end{aligned} \tag{23}$$

where the coefficients are in terms of the following properties of the unidirectional ply material

E_A = Axial Young's modulus (fiber direction)

ν_A = Axial Poisson's ratio

E_T = Transverse Young's modulus

ν_T = Transverse Poisson's ratio

G_A = Axial shear modulus

G_T = Transverse shear modulus

Suppose a typical element, Fig. 2, has in-plane dimensions $2a_m$ and $2b_n$. Then for half the laminate $0 \leqq z \leqq h$

$$\tilde{U}_{cn}'=\int_{-a_m}^{a_m}\int_{-b_n}^{b_n}\int_0^{t_1}W^{(1)}dxdydz+\int_{-a_m}^{a_m}\int_{-b_n}^{b_n}\int_{t_1}^{h}W^{(2)}dxdydz \tag{24a}$$

$$\tilde{U}_c'=\Sigma\,\tilde{U}_{cn}' \tag{24b}$$

which can now be evaluated in terms of equations (12). For simplicity in writing we shall perform detailed analysis for the case when $a_m = a$, $b_n = b$ and give results for the more general case of unequal intercrack distances at the end. When a_m and b_n are constant we can, without loss of generality, interpret all preceding energies as those stored in an element with dimensions $2a$, $2b$, h. Introducing equations (12) into (24a), performing all z integrations and using the nondimensional variables

$$\xi=x/t_1 \qquad \eta=y/t_1 \tag{25}$$

we have

$$\begin{aligned}2\tilde{U}_c'=\sigma^{o2}t_1^3\int_{-\rho_1}^{\rho_1}\int_{-\rho_2}^{\rho_2}&[k_x^2A_0\phi^2+2k_xk_yB_0\phi\psi\\&+k_y^2C_0\psi^2+A_1k_x^2\phi'^2+B_1k_y^2\dot{\psi}^2\\&+A_2k_x\phi(k_x\phi''+k_y\ddot{\psi})+B_2k_y\psi(k_x\phi''+k_y\ddot{\psi})\\&+C(k_x\phi''+k_y\ddot{\psi})^2]d\xi d\eta\end{aligned} \tag{26}$$

where prime and dot now denote ξ and η derivative, respectively, and

$$\begin{aligned}\phi&=\phi(\xi) & \psi&=\psi(\eta)\\ \rho_1&=a/t_1 & \rho_2&=b/t_1\\ k_x&=k_x^{(1)} & k_y&=k_y^{(1)}\end{aligned} \tag{27}$$

$$A_0 = 1/\lambda E_A + 1/E_T \qquad B_0 = -(1+1/\lambda)\nu_A/E_A$$
$$C_0 = 1/E_A + 1/\lambda E_T$$
$$A_1 = (\lambda/G_A + 1/G_T)/3 \qquad B_1 = (1/G_A + \lambda/G_T)/3$$
$$A_2 = [(3\lambda+2)\nu_T/E_T - \lambda\nu_A/E_A]/3 \qquad (28)$$
$$B_2 = [(3\lambda+2)\nu_A/E_A - \lambda\nu_T/E_T]/3$$
$$C = (\lambda+1)(3\lambda^2+12\lambda+8)/60E_T$$

and the boundary conditions (14) now assume the form

$$\phi(\pm\rho_1) = 1 \qquad \phi'(\pm\rho_1) = 0$$
$$\psi(\pm\rho_2) = 1 \qquad \dot{\psi}(\pm\rho_2) = 0 \qquad (29)$$

In view of (22) any functions ϕ and ψ which satisfy equations (29) will, when introduced into equations (24) and (22), provide an upper bound for $1/E_x$. To obtain the best bound $\tilde{U}_c'$ should be minimized. Using techniques of the calculus of variations we find that the minimum conditions of the integral (26) are the integro-differential equations

$$\frac{d^4\phi}{d\xi^4} + p_1\frac{d^2\phi}{d\xi^2} + q_1\phi + \frac{k_y}{k_x}\frac{B_0}{C}\frac{1}{2\rho_2}\int_{-\rho_2}^{\rho_2}\psi d\eta = 0 \qquad (30a)$$

$$\frac{d^4\psi}{d\eta^4} + p_2\frac{d^2\psi}{d\eta^2} + q_2\psi + \frac{k_x}{k_y}\frac{B_0}{C}\frac{1}{2\rho_1}\int_{-\rho_1}^{\rho_1}\phi d\xi = 0 \qquad (30b)$$

where

$$p_1 = \frac{A_2 - A_1}{C} \qquad q_1 = \frac{A_0}{C}$$
$$p_2 = \frac{B_2 - B_1}{C} \qquad q_2 = \frac{C_0}{C} \qquad (31)$$

Solution of the Integro-Differential Equations

Obviously the solution of equations (30) with (29) is insensitive to change of sign of independent variables. Therefore,

$$\phi(\xi) = \phi(-\xi) \qquad \psi(\eta) = \psi(-\eta) \qquad (32)$$

Now define the mean values

$$\frac{1}{2\rho_1}\int_{-\rho_1}^{\rho_1}\phi d\xi = \bar{\phi} \qquad \frac{1}{2\rho_2}\int_{-\rho_2}^{\rho_2}\psi d\eta = \bar{\psi} \qquad (33)$$

Then the general solutions of equations (30) can be written in the form

$$\phi = \phi_0(\xi) - m_1\bar{\psi} \qquad (34a)$$

$$\psi = \psi_0(\eta) - m_2\bar{\phi} \qquad (34b)$$

$$m_1 = \frac{k_y}{k_x}\frac{B_0}{A_0} \qquad m_2 = \frac{k_x}{k_y}\frac{B_0}{C_0} \qquad (34c)$$

where the terms with subscripts zero are general solutions of the homogeneous versions of equations (30) and the remaining terms are particular solutions. A homogeneous differential equation of type (30) has four independent solutions which may be arranged into two symmetric and two antisymmetric ones. In view of equations (32) only symmetric solutions need be retained. These are:

when $4q > p^2$

$$f(v) = Ch\alpha v\cos\beta v \qquad g(v) = Sh\alpha v\sin\beta v \qquad (35)$$
$$\alpha = q^{1/4}\cos(\theta/2) \qquad \beta = q^{1/4}\sin(\theta/2)$$
$$\tan\theta = \sqrt{4q/p^2 - 1}$$

when $4q < p^2$ $\qquad p < 0$

$$f(v) = Ch\alpha v \qquad g(v) = Ch\beta v \qquad (36)$$
$$\alpha,\beta = \sqrt{(-1 \pm \sqrt{1-4q/p^2})p/2}$$

where v is either ξ or η, and p and q are either of p_1, p_2 or q_1, q_2.

The case $p > 0$ is unlikely for the usual stiff fiber composites.

To solve the problem (29) and (30) let equations (34) be expressed in the form

$$\phi(\xi) = D_1 f_1(\xi) + F_1 g_1(\xi) - m_1\bar{\psi}$$
$$\psi(\eta) = D_2 f_2(\eta) + F_2 g_2(\eta) - m_2\bar{\phi} \qquad (37)$$

where f_1, g_1 indicate solutions in terms of p_1, q_1 and f_2, g_2—in terms of p_2, g_2. Introducing equations (37) into (29) we obtain sets of linear equations for the constants D, F with solutions

$$D_1 = \frac{(1+m_1\bar{\psi})g_1'(\rho_1)}{f_1(\rho_1)g_1'(\rho_1) - f_1'(\rho_1)g_1(\rho_1)}$$

$$F_1 = -\frac{(1+m_1\bar{\psi})f_1'(\rho_1)}{f_1(\rho_1)g_1'(\rho_1) - f_1'(\rho_1)g_1(\rho_1)} \qquad (38)$$

$$D_2 = \frac{(1+m_2\bar{\phi})\dot{g}_2(\rho_2)}{f_2(\rho_2)\dot{g}_2(\rho_2) - \dot{f}_2(\rho_2)g_2(\rho_2)}$$

$$F_2 = -\frac{(1+m_2\bar{\phi})\dot{f}_2(\rho_2)}{f_2(\rho_2)\dot{g}_2(\rho_2) - \dot{f}_2(\rho_2)g_2(\rho_2)}$$

Now average both sides of equations (37) as in (33). The result can be written in the form

$$\bar{\phi} + m_1\bar{\psi} = D_1\bar{f}_1 + F_1\bar{g}_1 = (1+m_1\bar{\psi})\omega_1$$
$$\bar{\psi} + m_2\bar{\phi} = D_2\bar{f}_2 + F_2\bar{g}_2 = (1+m_2\bar{\phi})\omega_2 \qquad (40)$$

where

$$\omega_1 = \frac{\bar{f}_1 g_1'(\rho_1) - f_1'(\rho_1)\bar{g}_1}{f_1(\rho_1)g_1'(\rho_1) - f_1'(\rho_1)g_1(\rho_1)}$$

$$\omega_2 = \frac{\bar{f}_2\dot{g}_2(\rho_2) - \dot{f}_2(\rho_2)\bar{g}_2}{f_2(\rho_2)\dot{g}_2(\rho_2) - \dot{f}_2(\rho_2)g_2(\rho_2)} \qquad (41)$$

Since expressions (41) are known, equations (40) can be solved for $\bar{\phi}$ and $\bar{\psi}$. Thus

$$\bar{\phi} = \frac{\omega_1 - m_1(1-\omega_1)\omega_2}{1 - m_1m_2(1-\omega_1)(1-\omega_2)}$$

$$\bar{\psi} = \frac{\omega_2 - m_2(1-\omega_2)\omega_1}{1 - m_1m_2(1-\omega_1)(1-\omega_2)} \qquad (42)$$

These results together with equations (30)–(39) determine the solution (37) of equations (29) and (30). Specific results for the two cases (35 and (36) are given in the Appendix.

The perturbation stresses (12) now assume the form

$$\sigma_{xx}^{(1)} = -\sigma_x^{(1)}\phi(\xi) \qquad \sigma_{xx}^{(2)} = \frac{1}{\lambda}\sigma_x^{(1)}\phi(\xi)$$

$$\sigma_{xz}^{(1)} = \sigma_x^{(1)}\frac{d\phi}{d\xi}z/t_1 \qquad \sigma_{xz}^{(2)}\frac{1}{\lambda}\sigma_x^{(1)}\frac{d\phi}{d\xi}(h-z)/t_1$$

$$\sigma_{yy}^{(1)} = -\sigma_y^{(1)}\psi(\eta) \qquad \sigma_{yy}^{(2)} = \frac{1}{\lambda}\sigma_y^{(1)}\psi(\eta) \qquad (43)$$

$$\sigma_{yz}^{(1)} = \sigma_y^{(1)}\frac{d\psi}{d\eta}z/t_1 \qquad \sigma_{yz}^{(2)} = \frac{1}{\lambda}\sigma_y^{(1)}\frac{d\psi}{d\eta}(h-z)/t_1$$

$$\sigma_{zz}^{(2)}=\left(\sigma_x^{(1)}\frac{d^2\phi}{d\xi^2}+\sigma_y^{(1)}\frac{d^2\psi}{d\eta^2}\right)(ht_1-z^2)/2t_1^2$$

$$\sigma_{zz}^{(2)}=\left(\sigma_x^{(1)}\frac{d^2\phi}{d\xi^2}+\sigma_y^{(1)}\frac{d^2\psi}{d\eta^2}\right)(h-z)^2/2t_1^2$$

Effective Young's Modulus E_x and Poisson's Ratio ν_{xy}

It is first necessary to evaluate the minimum value of the integral (26) and to introduce it into (22). Direct evaluation of the integral (26) in terms of the functions ϕ and ψ which have been determined above is a formidable undertaking which can fortunately be avoided. The best way to proceed is to integrate equation (30*a*) multiplied by ϕ from $-\rho_1$ to ρ_1 and, similarly, equation (30*b*) multiplied by ψ from $-\rho_2$ to ρ_2. These integrals obviously vanish and when transformed by integration by part with use of the boundary conditions (29) they can be identified with substantial parts of the integral (26). As a result of this procedure we find the simple result

$$U'_{c_{\min}}=-2t_1^3\sigma^{o^2}C[k_x^2\rho_2\phi'''(\rho_1)+k_y^2\rho_1\ddot{\psi}(\rho_2)] \tag{44}$$

To evaluate the third derivatives in equation (44) we evaluate the mean of equation (30*a*) with respect to ξ within $[-\rho_1, \rho_1]$ and of equation (30*b*) with respect to η within $[-\rho_2, \rho_2]$. Noting that the means of the second derivatives vanish because of equations (29) and taking into account equations (31) we obtain

$$\begin{aligned}-C\frac{\phi'''(\rho_1)}{\rho_1}&=A_0\bar{\phi}+\frac{k_y}{k_x}B_0\bar{\psi}\\ -C\frac{\dddot{\psi}(\rho_2)}{\rho_2}&=C_0\bar{\psi}+\frac{k_x}{k_y}B_0\bar{\phi}\end{aligned} \tag{45}$$

Substituting equations (45) into (44) and (27) we have

$$U'_{c_{\min}}=2t_1ab\sigma^{0^2}[(k_xA_0+k_yB_0)k_x\bar{\phi}+(k_xB_0+k_yC_0)k_y\bar{\psi}] \tag{46}$$

Thus the minimum energy has been expressed in terms of $\bar{\phi}$ and $\bar{\psi}$ which are explicitly determined by equations (42). Since the reference volume in the present case is

$$V=4abh=4ab(t_1+t_2) \tag{47}$$

we have from equations (46), (47), and (22)

$$\begin{aligned}\frac{1}{E_x}&\leqq\frac{1}{E_x^0}+\frac{1}{1+\lambda}(K_x\bar{\phi}+K_y\bar{\psi})\\ K_x&=(k_xA_0+k_yB_0)k_x\qquad K_y=(k_xB_0+k_yC_0)k_y\end{aligned} \tag{48}$$

This provides a rigorous lower bound on E_x in the event that all cracks are equidistant. If this is not the case let inter-crack distances be $2a_m$ in x direction and $2b_n$ in y direction, respectively. Then the admissible stress fields employed are still valid in each laminate element of dimensions $2a_m$, $2b_n$, h. The complementary energy functional is now equations (24) and each term in the summation is expressed in terms of functions ϕ_m and ψ_n, in the form equation (26), leading to extremum conditions of form (30) for each of these functions, subject to boundary condition of type (29)

$$\begin{aligned}\phi_m(\pm\rho_{1m})&=1\qquad\phi'_m(\pm\rho_{1m})=0\\ \psi_n(\pm\rho_{2n})&=1\qquad\psi'_n(\pm\rho_{2n})=0\end{aligned} \tag{49}$$

where

$$\rho_{1m}=a_m/t_1\qquad\rho_{2n}=b_n/t_1 \tag{50}$$

This defines means $\bar{\phi}_m$ and $\bar{\psi}_n$ which are precisely $\bar{\phi}$ and $\bar{\psi}$ expressed in terms of (50). Thus

$$U'_{c_{\min}}=-2t_1\sigma_0^2\sum^m\sum^n a_mb_n(K_x\bar{\phi}_m+K_y\bar{\psi}_n) \tag{51}$$

and the relevant laminate volume is now

$$V=4h\sum^m\sum^n a_mb_n \tag{52}$$

Introducing equations (51) and (52) into (22) we obtain again a lower bound for E_x.

This result can be put into more compact form if the inter-crack distances are random variables a, b, assuming values a_m, b_n. Then (50) are the values of associated random variables ρ_1 and ρ_2 with joint probability density function $P(\rho_1, \rho_2)$. In this event (22) will assume the form

$$\frac{1}{E_x}\leqq\frac{1}{E_x^0}+\frac{K_x<\rho_1\rho_2\phi(\rho_1)>+K_y<\rho_1\rho_2\psi(\rho_2)>}{(1+\lambda)<\rho_1\rho_2>} \tag{53}$$

where $<\ >$ indicates probability average. It can probably be assumed that ρ_1 and ρ_2 are independent random variables. In this event

$$P(\rho_1,\rho_2)=P_1(\rho_1)P_2(\rho_2) \tag{54}$$

and (53) assumes the form

$$\frac{1}{E_x}\leqq\frac{1}{E_x^0}+\frac{1}{1+\lambda}\left[K_x\frac{<\phi(\rho_1)>}{<\rho_1>}+K_y\frac{<\psi(\rho_2)>}{<\rho_2>}\right] \tag{55}$$

where

$$<\rho_1>=\int_{-\infty}^{\infty}P_1(\rho_1)\rho_1d\rho_1\quad<\rho_2>=\int_{-\infty}^{\infty}P_2(\rho_2)\rho_2d\rho_2 \tag{56}$$

and also

$$<\rho_1>=<a>/t_1\qquad<\rho_2>=<b>/t_2$$

The effective Poisson's ratio ν_{xy} is defined by

$$\nu_{xy}=-\frac{\bar{\epsilon}_{yy}}{\bar{\epsilon}_{xx}}\qquad\bar{\epsilon}_{xx}=\frac{\sigma^0}{E_x} \tag{57}$$

Since ν_{xy} cannot be bounded by variational methods we evaluate it in terms of the approximate stresses resulting from our variational treatment. For the case of equidistant cracks in the x and y direction

$$\bar{\epsilon}_{yy}=\frac{1}{4abh}\left\{\int_{-a}^{a}\int_{-b}^{b}\int_0^{t_1}\epsilon_{yy}^{(1)}dxdydz+\int_{-a}^{a}\int_{-b}^{b}\int_{t_1}^{h}\epsilon_{yy}^{(2)}dxdydz\right\} \tag{58}$$

It follows from the ply orientations and the stress-strain relations of the ply material that

$$\begin{aligned}\epsilon_{yy}^{(1)}&=-\sigma_{xx}^{(1)}\nu_A/E_A+\sigma_{yy}^{(1)}/E_A-\sigma_{zz}^{(1)}\nu_A/E_A\\ \epsilon_{yy}^{(2)}&=-\sigma_{xx}^{(1)}\nu_A/E_A+\sigma_{yy}^{(2)}/E_T-\sigma_{zz}^{(2)}\nu_T/E_T\end{aligned} \tag{59}$$

Next we introduce the stresses (11) and (12) into equations (59) and the resulting expressions into equation (58). We observe that: (1) The stresses $\sigma_{ij}^{0(m)}$ produce the Poisson's ratio ν_{xy}^0 of the undamaged laminate; (2) The contribution of σ_{zz} to equation (58) vanishes since

$$\int_{-a}^{a}\int_{-b}^{b}\phi''(x)dxdy=2b\int_{-a}^{a}\phi''(x)dx=2b[\phi'(x)]_{-a}^{a}$$

and this vanishes because of equations (14). A similar result is true for the integral of ψ over the rectangle. An easy calculation then yields the result

Table 1 Properties of unidirectional material. Moduli in GPa.

	E_A	E_T	G_A	G_T	ν_A	ν_T
Glass/Epoxy	41.7	13.0	3.40	4.58	0.300	0.420
Graphite/Epoxy	208.3	6.5	1.65	2.30	0.255	0.413

$$\nu_{xy} = \nu^0_{xy}\frac{E_x}{E^0_x} - \frac{k^{(1)}_y \bar{\psi}}{1+\lambda}\left(\frac{E_x}{E_T} - \frac{E_x}{E_A}\right) \tag{60}$$

where $\bar{\psi}$ is defined by equations (33). Evaluation of ν_{yx} is of course entirely analogous.

Results, Discussion and Conclusions

The approximate nature of the analysis which has been developed here is contained in the assumed form of the stresses (7) which certainly cannot be correct. Indeed from experience with crack field solutions one would expect local singularities at the crack tips, which are absent from equations (7). But it should be remembered that in the usual structural laminates we are dealing with very thin plies of thickness of order 0.2–0.4 mm. Typical glass and carbon fiber diameters are 0.01 mm. Assuming for the sake of argument that fibers are arranged in regular hexagonal arrays, we find that at the usual 0.60 fiber volume fraction the ply thickness can accommodate about 16–32 fibers. Now the basic assumption in laminate theory is that plies are homogeneous anisotropic with elastic properties equal to the effective elastic properties of the unidirectionally reinforced ply material. But it should be remembered that the concept of effective modulus is based on stress and strain averages over representative volume elements (RVE) which by definition must contain many fibers. It follows that a ply can accommodate perhaps 2–3 RVE sizes across its thickness. Therefore, a severe gradient of local average stress or strain through the thickness is impossible. This does not present a problem for the simple cases of undamaged laminates when the stresses are constant or linear across ply thickness, but the situation is very different in the case of an intralaminar crack. If it is assumed that the fields of the homogeneous ply model are averages over RVE, then the crack tip singularities of fracture mechanics and their associated large stress gradient cannot exist. If this assumption is not made, there remains the choice between two difficult alternatives: (*a*) The crack tip is in the matrix and recognizes the adjacent fibers as distinct heterogeneities, thus creating an intractable problem in crack mechanics; (*b*) The crack tip vicinity is governed by a nonlocal elasticity theory of unknown nature or with many unknown elastic constants. None of these alternatives is useful or practical and it would therefore appear that the simple assumptions (7), with their resultant gentle through-the-thickness variations of the stresses, are less severe than would perhaps appear at first sight.

All of these considerations apply also to finite element analysis. The number of finite elements allowable through the thickness must be severely limited and singular elements raise the same problems as the crack tip singularities discussed above.

Highsmith and Reifsnider (1986) have treated the problem of orthogonally cracked cross-plies in terms of a refined laminate theory advanced by Pagano (1978) in which stress variation through ply thickness is restricted to be linear. The treatment is of necessity restricted to equidistant cracks and requires extensive numerical analysis. Their results, Highsmith and Reifsnider (1986) show that the stress σ_{xx} depends only weakly on y, the stress σ_{yy} depends only weakly on x, and σ_{xy} is negligible. This is in accordance with our basic assumptions (7). Furthermore, the shear stress σ_{xz} practically does not depend on y and σ_{xz} practically does not depend on x and this agrees with (12 *b*,*d*,*g*,*i*).

Table 2 The effect of cracks on Young's modulus E_x and Poisson's ratio ν_{xy}

	Glass/Epoxy								Graphite/Epoxy			
	$t_1 = 3t_2$				$t_1 = t_2$				$t_1 = t_2$			
	$\rho_1 = \rho$		$\rho_1 = \rho_2 = \rho$		$\rho_1 = \rho$		$\rho_1 = \rho_2 = \rho$		$\rho_1 = \rho$		$\rho_1 = \rho_2 = \rho$	
ρ	$\rho_2 \to \infty$				$\rho_2 \to \infty$				$\rho_2 \to \infty$			
no cracks	$E^0_x = 20.30$GPa		$\nu^0_{xy} = .113$		$E^0_x = 27.57$GPa		$\nu^0_{xy} = .143$		$E^0_x = 107.6$GPa		$\nu^0_{xy} = .0154$	
	E_x/E^0_x	ν_{xy}	E_x/E^0_x	ν_{xy}	E_x/E^0_x	ν_{xy}	E_x/E^0_x	ν_{xy}	E_x/E^0_x	ν_{xy}	E_x/E^0_x	ν_{xy}
50	.980	.111	.980	.110	.990	.141	.990	.142	.999	.0154	.999	.0198
20	.951	.108	.951	.106	.976	.139	.975	.140	.997	.0154	.997	.0264
10	.907	.102	.906	.100	.953	.134	.951	.139	.994	.0153	.994	.0374
5	.830	.094	.829	.091	.910	.129	.907	.138	.989	.0152	.988	.0596
3	.745	.084	.743	.082	.859	.122	.853	.140	.982	.0151	.981	.0864
2	.661	.074	.658	.075	.813	.115	.804	.144	.975	.0150	.974	.108
1	.548	.061	.542	.073	.775	.109	.762	.149	.971	.0149	.969	.125
.50	.524	.058	.516	.075	.770	.108	.757	.150	.970	.019	.968	.127

For purpose of numerical evaluation we have used glass/epoxy and graphite/epoxy laminates with unidirectional ply properties as in Hashin (1985). These are given in Table 1.

Table 2 shows Young's modulus reduction and also Poisson's ratio for various laminates with equidistant cracks in both directions. Also shown by comparison are results of Hashin (1985) for the case when there are cracks only in the 90° ply. It is seen that the effect of the cracks in the 0° plies is quite small, no doubt because of the small value of transverse σ_{yy} stress in the 0° plies due to tension in 0°, i.e., x direction. This is also in agreement with the results of Highsmith and Reifsnider (1986). The situation is quite different for the Poisson's ratio. It is seen that cracks have significant effects

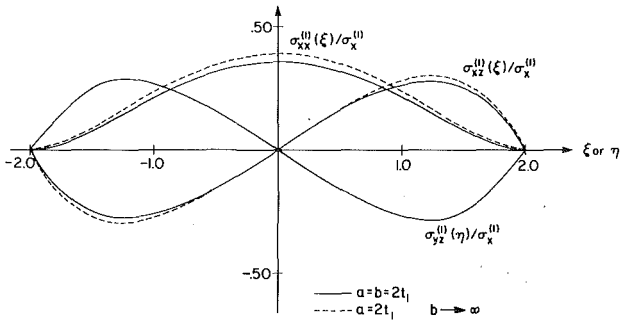


Fig. 3 $\sigma_{xx}^{(1)}$, $\sigma_{xz}^{(1)}$, and $\sigma_{yz}^{(1)}$ variation

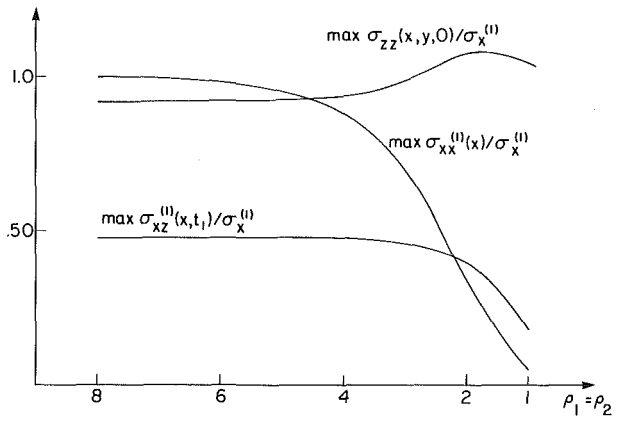


Fig. 4 Variation of maximum σ_{xx}, τ, and σ_{zz} with ρ

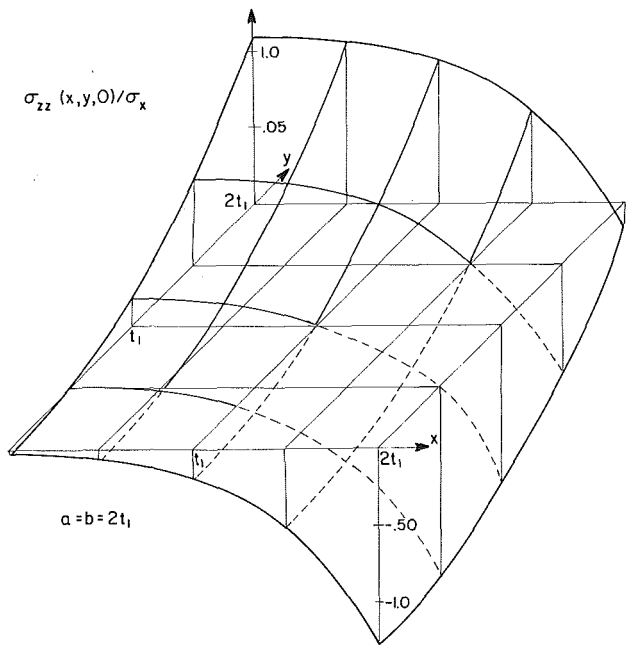


Fig. 5 Variation of σ_{zz} over midplane

on the value of this quantity and that the effect of orthogonal cracks can be very substantial as is evidenced in particular by the Graphite/Epoxy cross ply.

For the loading case of uniaxial tension considered here the stresses of major interest are: $\sigma_{xx}^{(1)}$ and $\sigma_{yy}^{(2)}$ which are transverse to the fibers in their respective layers and may thus produce intralaminar cracks; the interlaminar shear stresses σ_{xz} and σ_{yz}; the stress σ_{zz} normal to the plane of the laminate which is called peeling stresses when tensile. It is seen from the form of equations (12) and (43) that the shear stresses assume their maximum values at the interface of the plies and that σ_{zz} assumes its maximum value on the midplane $z = 0$. However, the value of σ_{zz} at the interface is also of importance, for the stress which produces interlaminar separation may be smaller than the one producing midplane transverse cracking.

We have evaluated stress distributions for $[0°/90°]_s$ graphite/epoxy laminates assuming equidistant cracks $a = b$, for various crack densities defined by different ratios $\rho = a/t_1$. We have observed the following significant phenomena:

(1) The tensile stress $\sigma_{yy}^{(2)}$ is much smaller than the tensile stress $\sigma_{xx}^{(1)}$ in the load direction. It can, therefore, be assumed that during loading no new intralaminar cracks will develop in the 0° plies in the y direction.

(2) The stresses $\sigma_{xx}^{(1)}$ and $\sigma_{xz}^{(1)}$ are close in value to the same stresses produced in the same laminate when only the 90 plies are cracked (Hashin, 1985). A comparison is shown in Fig. 3 for the case $a/t_1 = 2$.

(3) The maximum value of tensile transverse stress $\sigma_{xx}^{(1)}$ is midway between cracks, is equal to the corresponding stress $\sigma_x^{(1)}$ in the undamaged laminate when distances between cracks are large, and decreases monotonically with increasing crack density (decreasing ρ) (Fig. 4).

(4) The intralaminar shear stress σ_{yz} has significant values (Fig. 3) and therefore the important interlaminar shear stress is $\tau = \sqrt{\sigma_{xz}^2 + \sigma_{yz}^2}$. Figure 4 shows the value of $\tau_{\max}$ as a function of crack density. It is seen that this stress starts out with a value of $0.48\sigma_x^{(1)}$ which is certainly significant. When the vlaue of $\sigma_x^{(1)}$ is large enough to cause cracking, τ may be large enough to produce shear delamination. It is interesting to note that increasing crack density decreases this shear stress.

(5) The peeling stress σ_{zz} assumes significant values. Figure 5 shows the variation of midplane σ_{zz} over one quarter of the square defined by the intersection of two pairs of cracks. Observe the large tensile stress $1.08\sigma_x^{(1)}$ in the middle of the $y = a$ edge. When $\sigma_x^{(1)}$ is large enough to crack the 90° ply transversely, this stress is certain to produce a crack in the xy plane. Note that for the present laminate σ_{zz} at the ply interface is one half of the midplane value. Depending on interface strength such a stress may also produce interface separation. The largest value of σ_{zz} is always located on the edges $y = \pm a$. For small crack density the maximum is close to the cracks $x = \pm a$ and its location moves to the center of the edge with increasing crack density. Figure 4 shows the variation of max σ_{zz} with intercrack distance. It would be of interest to analyze various other cases and also biaxial loading, which is easily carried out by superposing the results of two uniaxial loads in the x and y directions.

Acknowledgment

Support of the Air Force Office of Scientific Research under Contract 85-0342 is gratefully ackowledged.

References

Hashin, Z., 1985, "Analysis of Cracked Laminates: A Variational Approach," *Mechanics of Materials*, Vol. 4, pp. 121–136.

Highsmith. A. L., and Reifsnider, K. L., 1986, "Internal Load Distribution Effects During Fatigue Loading of Composite Laminates," *ASTM*, STP 907, pp. 233–251.

Laws, N., Dvorak, G. J., and Hejazi, M., 1983, "Stiffness Changes in Unidirectional Composites Caused by Crack Systems," *Mechanics of Materials*, Vol. 2, pp. 123–137.

Laws, N., Dvorak, G. J., and Hejazi, M., 1985, "The Loss of Stiffness in Cracked Laminates," *Fundamentals of Deformation and Fracture*, Proc. IUTAM Eshelby Memorial Symposium, Cambridge University Press, pp. 119–128.

Pagano, N. J., 1978, "Stress Fields in Composite Laminates," *International Journal of Solids and Structures*, Vol. 14, pp. 385–400.

Reifsnider, K. L., and Jamison, R., 1982, "Fracture of Fatigue-loaded Composite Laminates," *International Journal of Fatigue*, pp. 187–197.

Reifsnider, K. L., and Talug, A., 1980, "Analysis of Fatigue Damage in Composite Laminates," *International Journal of Fatigue*, pp. 3–11.

APPENDIX

Denote for convenience

$$\bar{\phi}=\Phi_1 \qquad \bar{\psi}=\Phi_2$$

For case (35)

$$D_1=2(1+m_1\Phi_2)[\alpha_1 Ch(\alpha_1\rho_1)\sin(\beta_1\rho_1)+\beta_1 Sh(\alpha_1\rho_1)\cos(\beta_1\rho_1)]/M_1$$

$$F_1=2(1+m_1\Phi_2)[\beta_1 Ch(\alpha_1\rho_1)\sin(\beta_1\rho_1)-\alpha_1 Sh(\alpha_1\rho_1)\cos(\beta_1\rho_1)]/M_1$$

$$M_1=\alpha_1\sin(2\beta_1\rho_1)+\beta_1 Sh(2\alpha_1\rho_1)$$

$$\omega_1=\frac{2[Ch(2\alpha_1\rho_1)-\cos(2\beta_1\rho_1)]}{\rho_1(\alpha_1/\beta_1+\beta_1/\alpha_1)M_1}$$

To obtain D_2, F_2, and ω_2, simply replace 1 by 2 in the above equations. Then

$$\phi(\xi)=D_1 Ch(\alpha_1\xi)\cos(\beta_1\xi)+F_1 Sh(\alpha_1\xi)\sin(\beta_1\xi)-m_1\bar{\psi}$$

$$\psi(\eta)=D_2 Ch(\alpha_2\eta)\cos(\beta_2\eta)+F_2 Sh(\alpha_2\eta)\sin(\beta_2\eta)-m_2\bar{\phi}$$

For case (36)

$$D_1=(1+m_1\Phi_2)\beta_1 Sh(\beta_1\rho_1)/N_1$$

$$F_1=-(1+m_1\Phi_2)\alpha_1 Sh(\alpha_1\rho_1)/N_1$$

$$N_1=\beta_1 Ch(\alpha_1\rho_1)Sh(\beta_1\rho_1)-\alpha_1 Sh(\alpha_1\rho_1)Ch(\beta_1\rho_1)$$

$$\omega_1=(\beta_1/\alpha_1-\alpha_1/\beta_1)Sh(\alpha_1\rho_1)Sh(\beta_1\rho_1)/\rho_1 N_1$$

To obtain D_2, F_2, and ω_2, replace 1 by 2 in the above equations. Then

$$\phi(\xi)=D_1 Ch(\alpha_1\xi)+F_1 Sh(\beta_1\xi)-m_1\bar{\psi}$$

$$\psi(\eta)=D_2 Ch(\alpha_2\eta)+F_2 Sh(\beta_2\eta)-m_2\bar{\phi}$$

Effective Medium Approach to Matrix-Inclusion Type Composite Materials

S. Nomura
Department of Mechanical Engineering,
University of Texas at Arlington,
Arlington, TX 76019

This paper addresses a problem of finding the effective medium that exhibits the same overall response as a given composite material reinfored by unidirectional short-fibers (chopped fibers). The expression for the displacement field in composites is obtained by converting the equilibrium equation into an integro-differential equation using the quasi-static Green's function for a homogeneous medium. The "effective medium" is chosen that the ensemble averaged displacement field for the composite is equal to that of an equivalent homogenous medium. The "effective stiffness" and the "effective mass density" are defined as those properties of the effective medium. This is a first preliminary attempt to analyze the elasto-dynamic effect of matrix-inclusion type of composites. The obtained result for the effective stiffness is new and is not symmetrical with the interchange of the matrix phase and the fiber phase, unlike previous models. The result is also favorably compared with experimental data for spherical-inclusion reinforced composites.

Introduction

This paper introduces an analytical technique to find an equivalent homogeneous medium with a given composite material using a statistical approach with the intention of extending the method to elasto-dynamic analysis of composite materials. The composite material in this paper consists of a homogeneous matrix medium in which ellipsoidal-shaped inclusions (short-fibers, discontinuous fibers, chopped fibers) are embedded. Modeling short-fibers as ellipsoidal inclusions is advantageous because by changing the inclusion aspect ratio (major axis/minor axis), they can cover a wide range of fiber geometries from spherical particulates (aspect ratio = 1) to short fibers (intermediate aspect ratio) to long-continuous fibers (aspect ratio = ∞).

The study of finding the effective modulus of composites has a long history dating back to the beginning of the century and numerous solution methods have already been proposed as more composite materials are being used in industries (see a recent review of Hashin, 1983). The self-consistent mechanics of Hill (1965) and Budiansky (1965) was adopted to calculate effective elastic moduli of various types of composites including short-fiber composites by Laws and McLaughlin (1979) and Chou et al. (1980). The limitation of the self-consistent approximation used in those works is that it overestimates effective moduli of composites at higher fiber volume fractions due to the assumption that the immediate surrounding of a single fiber has the properties of the composite, not the matrix. Most available experimental data also show lower values of effective moduli than those predicted by the self-consistent method (e.g., Blumentritt, 1974). Christensen and Lo (1979) developed a model of three phase composites with long continuous fibers and spherical inclusions to take the matrix phase into account. However, their method cannot be directly applied to other shapes of fibers due to analytical complexity.

In this paper, the effective modulus and the effective mass density of fiber composite materials are obtained by introducing an analytical model based on the micromechanical approach overcoming some of the above difficulties inherent to existing solution techniques. In addition, this is a first preliminary attempt to treat elasto-dynamics of composites. An integral equation for the displacement field in the composite is derived by introducing the quasi-static Green's function for a homogeneous medium. An "effective medium" is sought that exhibits the same overall response as the composite for the same boundary condition by a self-consistent approximation by Berryman (1979) and Gubernatis et al. (1979). The above scheme yields a set of algebraic equations for the effective elastic modulus and the effective mass density. The obtained result yields lower values of effective moduli than those by previous models. They are also favorably compared with available experimental data for spherical inclusions as a special case.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 5, 1985; final revision April 2, 1987.

Paper No. 87-WA/APM-23.

Formulation

The elastic equilibrium equation for a particular composite material denoted by α without body force when the elastic modulus and the mass density are functions of position is expressed as

$$(C_{ijkl}(\mathbf{r},\alpha)u_k(\mathbf{r},t,\alpha)_{,l})_{,j} - \rho(\mathbf{r},\alpha)\partial^2 u_i(\mathbf{r},t,\alpha)/\partial t^2 = 0 \quad (1)$$

where $_{,i}$ denotes the partial derivative with respect to x_i, $C_{ijkl}(\mathbf{r}, \alpha)$ and $\rho(\mathbf{r}, \alpha)$ are anisotropic elastic moduli and mass densities of a particular composite denoted by α, respectively. The summation convention is used throughout.

The elastic modulus and the mass density can be decomposed into a reference part and a fluctuating part as

$$C_{ijkl}(\mathbf{r},\alpha) = C^*_{ijkl} + \delta C_{ijkl}(\mathbf{r},\alpha) \quad (2)$$

$$\rho(\mathbf{r},\alpha) = \rho^* + \delta\rho(\mathbf{r},\alpha) \quad (3)$$

where C^*_{ijkl} and ρ^* are a constant elastic modulus and a constant mass density of a reference medium which is not specified yet and $\delta C_{ijkl}(\mathbf{r}, \alpha)$ and $\delta\rho(\mathbf{r}, \alpha)$ are fluctuating quantities from those of the reference medium. Substitution of equations (2) and (3) into equation (1) yields

$$C^*_{ijkl}u_k(\mathbf{r},t,\alpha)_{,lj} - \rho^*\partial^2 u_i(\mathbf{r},t,\alpha)/\partial t^2 = -((\delta C_{ijkl}(\mathbf{r},\alpha)u_k(\mathbf{r},t,\alpha)_{,l})_{,j} - \delta\rho(\mathbf{r},\alpha)\partial^2 u_i(\mathbf{r},t,\alpha)/\partial t^2) \quad (4)$$

It is convenient to Fourier transform equation (4) as

$$U_k(\mathbf{r},\omega,\alpha) = \int_0^\infty u_k(r,t,\alpha)\exp(-i\omega t)\,dt \quad (5)$$

Equation (4) now becomes

$$C^*_{ijkl}U_k(\mathbf{r},\alpha)_{,lj} + \omega^2\rho^* U_i(\mathbf{r},\alpha) = -((\delta C_{ijkl}(\mathbf{r},\alpha)U_k(\mathbf{r},\alpha)_{,l})_{,j} + \omega^2\delta\rho(\mathbf{r},\alpha)U_i(\mathbf{r},\alpha)) \quad (6)$$

where for brevity, ω in U_k notation is dropped henceforth.

In order to solve equation (6), the time-reduced Green function (quasi-Green function) is introduced as

$$C^*_{ijkl}g_{kp}(\mathbf{r}-\mathbf{r}')_{,lj} + \omega^2\rho^* g_{ip}(\mathbf{r}-\mathbf{r}') = -\delta_{ip}\delta(\mathbf{r}-\mathbf{r}') \quad (7)$$

where δ_{ip} is the Kronecker delta and $\delta(\mathbf{r})$ is the Dirac delta function. With equation (7), equation (6) can be solved formally by regarding the right-hand side of equation (6) as an imaginary body force as

$$U_p(\mathbf{r},\alpha) = U^A_p(\mathbf{r}) + \int(\delta C_{ijkl}(\mathbf{r}',\alpha)U_k(\mathbf{r}',\alpha)_{,l'})_{,j'}g_{pi}(\mathbf{r}-\mathbf{r}')d\mathbf{r}' + \omega^2\int\delta\rho(\mathbf{r}',\alpha)U_i(\mathbf{r}',\alpha)g_{pi}(\mathbf{r}-\mathbf{r}')d\mathbf{r}' \quad (8)$$

where $U^A_p(\mathbf{r})$ is Fourier transform of a displacement field in a homogeneous medium with C^*_{ijkl} and ρ^* that satisfies the displacement boundary condition. The integral range in equation (8) is over the whole material points. It should be noted that by converting the differential equation into an integro-differential equation, the continuity of the displacement and traction across the phase boundary is automatically taken care of. Equation (8) can be rearranged using integration by parts to

$$U_p(\mathbf{r},\alpha) = U^A_p(\mathbf{r}) + \int\delta C_{ijkl}(\mathbf{r}',\alpha)U_k(\mathbf{r}',\alpha)_{,l'}g_{pi}(\mathbf{r}-\mathbf{r}')_{,j}d\mathbf{r}' + \omega^2\int\delta\rho(\mathbf{r}',\alpha)U_i(\mathbf{r}',\alpha)g_{pi}(\mathbf{r}-\mathbf{r}')d\mathbf{r}' \quad (9)$$

Note that equation (9) is formal because it contains $U_p(\mathbf{r}, \alpha)$ in both sides.

Effective Medium

When the composite has N kinds of fibers embedded in a homogeneous matrix phase, it is convenient to rearrange $\delta C_{ijkl}(\mathbf{r}, \alpha)$ and $\rho(\mathbf{r}, \alpha)$ as

$$\delta C_{ijkl}(\mathbf{r},\alpha) = (C^m_{ijkl} - C^*_{ijkl}) + \sum_{n=1}^{N}(C^n_{ijkl} - C^m_{ijkl})\Phi_n(\mathbf{r},\alpha) \quad (10)$$

$$\delta\rho(\mathbf{r},\alpha) = (\rho^m - \rho^*) + \sum_{n=1}^{N}(\rho^n - \rho^m)\Phi_n(\mathbf{r},\alpha) \quad (11)$$

where the quantities with the index m denote those associated with the matrix phase, C^m_{ijkl} and ρ^n are the elastic modulus and the mass density of the nth kind fiber, respectively, and $\Phi_n(\mathbf{r}, \alpha)$ is the characteristic function of the nth kind fiber phase defined as (Kröner, 1977; Nomura, 1982)

$$\Phi_i(\mathbf{r},\alpha) \equiv \begin{cases} 0 \text{ if } \mathbf{r} \notin i\text{th phase} \\ 1 \text{ if } \mathbf{r} \in i\text{th phase} \end{cases} \quad i = 1,2,\ldots,n \quad (12)$$

The ensemble average of a stochastic function $f(\mathbf{r}, \alpha)$ over the sample space is defined as

$$\langle f(\mathbf{r},\alpha)\rangle = \int f(\mathbf{r},\alpha)P(d\alpha) \quad (13)$$

where $P(d\alpha)$ is a probabilistic measure defined over the sample space where each element represents the actual composite (see Nomura, 1982, for detail). By taking the ensemble average of equation (9) after substituting equations (11) and (12), we have

$$\langle U_p(\mathbf{r},\alpha)\rangle = U^A_p(\mathbf{r}) + (C^m_{ijkl} - C^*_{ijkl})\int\langle U_k(\mathbf{r}',\alpha)_{,l'}\rangle g_{pi}(r-r')_{,j}d\mathbf{r}' + \sum_{n=1}^{N}(C^n_{ijkl} - C^m_{ijkl})\int\langle\Phi_n(\mathbf{r}',\alpha)U_k(\mathbf{r}')_{,l}\rangle g_{pi}(\mathbf{r}-\mathbf{r}')_{,j}d\mathbf{r}' + \omega^2(\rho^m - \rho^*)\int\langle U_i(\mathbf{r}')\rangle g_{pi}(\mathbf{r}-\mathbf{r}')d\mathbf{r}' + \omega^2\sum_{n=1}^{N}(\rho^n - \rho^m)\int\langle\Phi_n(\mathbf{r}',\alpha)U_i(\mathbf{r}')\rangle g_{pi}(\mathbf{r}-\mathbf{r}')d\mathbf{r}' \quad (14)$$

So far C^*_{ijkl} and ρ^* were not specified yet. We now adopt a self-consistent scheme that C^*_{ijkl} and ρ^* are chosen in such a way that the ensemble average of $U_p(\mathbf{r}, \alpha)$ is equal to $U^A_p(\mathbf{r})$. For such a situation, the "effective" elastic modulus C^*_{ijkl} and the "effective" mass density ρ^* are defined as the properties of the "equivalent" homogeneous medium (effective medium) that produces the same ensemble averaged displacement as the composite as

$$\langle U_p(\mathbf{r},\alpha)\rangle = U^A_p(\mathbf{r}) \quad (15)$$

We further adopt the following replacement of each quantity in equation (14) as

$$\langle U_k(\mathbf{r},\alpha)_{,l}\rangle = U^A_k(\mathbf{r})_{,1} \quad (16)$$

$$\langle\Phi_n(\mathbf{r},a)U_k(\mathbf{r},\alpha)_{,l}\rangle = V_n A_{klij}(\mathbf{r})U^A_i(\mathbf{r})_{,j} \quad (17)$$

$$\langle\Phi_n(\mathbf{r},\alpha)U_i(\mathbf{r},\alpha)\rangle = V_n U^A_i(\mathbf{r}) \quad (18)$$

where V_n is the volume fraction of the nth kind fiber, $A^{(n)}_{klij}$ is a proportionality factor of the displacement gradient inside the nth kind fiber when a single fiber is placed in a homogeneous medium with the displacement gradient $U^A_{i,j}$ applied at remote distances as a boundary condition.

Equation (16) is a direct consequence of equation (15) where differentiation on space is interchanged with the ensemble average. In equations (17)–(18), we adopt the "ergodic assumption" that the ensemble average can be replaced by the spatial average

$$\int f(\mathbf{r},\alpha)P(d\alpha) = 1/V\int f(\mathbf{r}+\mathbf{r}',\alpha)d\mathbf{r}' \quad (19)$$

where V is the volume of the composite. Equation (19) is expected to be valid for small ω. Equation (17) implies that the displacement gradient field inside each fiber is approximated by the displacement gradient field when a single fiber is placed in a homogeneous medium subjected to a boundary displacement gradient U^A_p. This approximation is called quasi-static approximation (Berryman, 1979) which is correct to $O(\omega^4)$

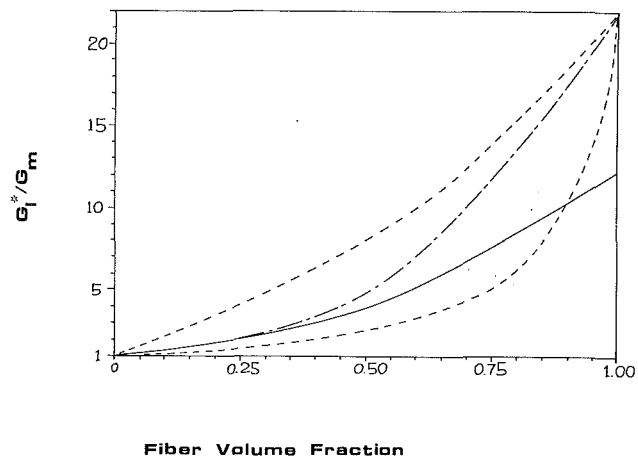


Fig. 1 The variation of axial shear modulus G_L^*/G_m with V_f for G_f/G_m = 21.538, ν_m = 0.4, and ν_f = 0.3 where G and ν are the shear modulus and Poisson's ratio, respectively, for composites with long continuous fibers ($\ell/d = \infty$); solid line: present theory; broken lines: upper and lower bounds of Hashin and Shtrikman (1965); chain-dotted line: conventional self-consistent model of Chou et al. (1980).

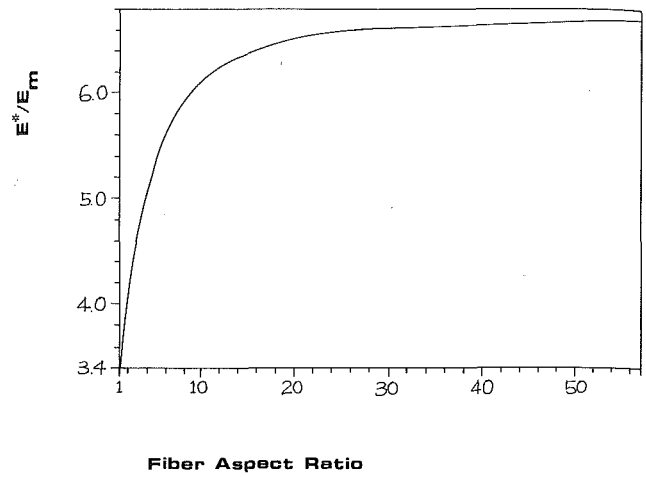


Fig. 2 The variation of E_L^* (axial Young's modulus)/E_m versus the fiber aspect ratios. V_f = 0.3, E_f/E_m = 20, ν_m = 0.4, and ν_f = 0.3 where V_f is the fiber volume fraction, E is the Young Modulus, and ν is the Poisson ratio.

exact as $\omega \to 0$ (Gubernatis et al., 1979). In equation (18), displacement field inside each fiber is approximated by a boundary displacement. Equations (17) and (18) can be justified because the displacement field is continuous through the composite while the displacement gradient has discontinuity across the phase boundary. Therefore, by combining equations (14)–(18), we obtain the following equations for C_{ijkl}^* and ρ^* as

$$\rho^m - \rho^* + \sum_{n=1}^{N} V_n(\rho^n - \rho^m) = 0 \tag{20}$$

$$C_{ijkl}^m - C_{ijkl}^* + \sum_{n=1}^{N} V_n A_{ijpq}^{(n)} (C_{pqkl}^n - C_{pqkl}^m) = 0 \tag{21}$$

Equation (20) can be solved for ρ^* for two-phase composites as

$$\rho^* = V_m \rho^m + V_i \rho^i \tag{22}$$

Although equation (21) can be derived in a simpler manner from elastostatic consideration, it should be noted that our approach makes elasto-dynamic analysis possible. Equation (22) shows that the effective mass density follows the rule of mixtures, which can be also derived by considering the balance of mass. Equation (21) yields a system of nonlinear algebraic equations since $A_{ijmn}^{(n)}$ is a function of C_{ijkl}^m, C_{ijkl}^*, and C_{ijkl}^n as well as the fiber aspect ratio (major axis/minor axis).

Numerical Calculations

The solutions to equation (21) for C_{ijkl}^* can be obtained numerically by iteration for several types of composites. The proportionality factor $A_{ijkl}^{(n)}$ was obtained by Lin and Mura (1973) when the shape of the fiber is ellipsoidal and the matrix phase is transversely-isotropic. Those expressions can be found in the book by Mura (1982) and is cited in modified form in the Appendix.

Figure 1 shows the comparisons of the effective axial shear modulus normalized by the matrix shear modulus, with the conventional self-consistent model (Chou et al., 1980) denoted by a chain-dotted line and upper and lower bounds, derived by Hashin and Shtrikman (1965), denoted by broken lines for the composite reinforced by long-continuous fibers. Although both the matrix phase and the fiber phases are isotropic, the whole composite renders transversely isotropic. It is seen that

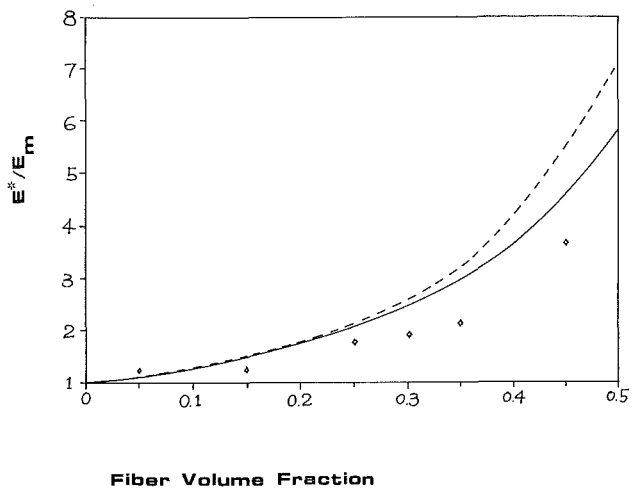


Fig. 3 The variation of E^*/E_m with V_f for E_f/E_m = 37, ν_m = 0.44, ν_f = 0.21 and ℓ/d = 1; solid line: present theory; broken line: conventional self-consistent model (Chou et al. 1980); solid circles: experimental data of Richard (1975).

the present approach predicts lower values than the conventional self-consistent method. The present prediction lies between the upper and lower bounds of Hashin and Shtrikman up to the fiber volume fraction of 0.95 and falls below the lower bound beyond this limit. However, it should be noted that the closest packing factor of the cylindrical fibers in the composites is about 0.9 and beyond this volume the physical significance is lost. Each curve beyond this point is shown just for completeness.

Figure 2 shows the variation of the effective axial Young modulus (normalized by the matrix Young modulus) with the fiber aspect ratios. As seen from Fig. 2, the effective axial Young's modulus exhibits the "rule of mixtures" as the fiber aspect ratio goes beyond 50. For all practical purposes, the fiber aspect ratio greater than 50 can be identified as the long-continuous fibers as far as the mechanical properties are concerned.

Figure 3 is the result of the present theory (solid line), the conventional self-consistent method (broken line), and the experimental data tested by Richard (1975) for the spherical inclusion composite (aspect ratio = 1). It is seen that the present prediction coincides very well with experimental data.

At high frequency regions, it is expected that the effective elastic modulus depends on the frequency. In such a case, the

transient effective quantity must be defined and treated. Further analysis will follow subsequently.

References

Aboudi, J., 1983, "The Effective Moduli of Short-Fiber Composites," *International Journal of Solids and Structures*, Vol. 19, pp. 693–707.

Berryman, J. G., 1979, "Theory of Elastic Properties of Composite Materials," *Applied Physics Letters*, Vol. 35, pp. 856–858.

Blumentritt, B. F., Vu, B. T., and Cooper, S. L., 1974, "The Mechanical Properties of Oriented Discontinuos Fibre-Reinforced and Thermoplastics I. Unidirectional Fiber Orientations," *Polymer Engineering and Science*, Vol. 14, pp. 633–645.

Budiansky, B., 1967, "On the Elastic Moduli of Some Heterogeneous Materials," *Journal of Mechanics and Physics of Solids*, Vol. 13, pp. 223–227.

Chou, T.-W., Nomura, S., and Taya, M., 1980, "A Self-Consistent Approach to the Elastic Stiffness of Short-Fiber Composites," *Journal of Composite Materials*, Vol. 14, pp. 178–188.

Christensen, R. M., and Lo, K. H., 1979, "Solutions for Effective Shear Properties in Three Phase Sphere and Cylinder Models," *Journal of Mechanics and Physics of Solids*, Vol. 27, pp. 315–330.

Gubernatis, J. E., Krumhansl, J. A., and Thomson, R. A., 1979, "Interpretation of Elastic Wave Scattering Theory for Analysis and Design of Flow Characterization Experiments. The Long Wave Limit," *Journal of Applied Physics*, Vol. 50, pp. 3338–3345.

Hashin, Z., and Shtrikman, S., 1965, "A Variational Approach to the Theory of the Elastic Behaviour of Multiphase Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 13, pp. 223–227.

Hashin, Z., 1983, "Analysis of Composite Materials – A Survey," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 481–505.

Hill, R., 1965, "A Self-Consistent Mechanics of Composite Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 13, pp. 213–222.

Kröner, E., 1977, "Bounds for Effective Elastic Moduli of Disordered Materials," *Journal of Mechanics and Physics of Solids*, Vol. 15, pp. 137–155.

Laws, N., and McLaughlin, R., 1979, "The Effect of Fibre Length on the Overall Moduli of Composite Materials," *Journal of Mechanics and Physics of Solids*, Vol. 27, pp. 1–13.

Lin, S. C., and Mura, T., 1973, "Elastic Fields of Inclusion in Anisotropic Media II," *Physica Status Solidi (a)*, Vol. 15, pp. 281–285.

Mura, T., 1982, *Micromechanics of Defects in Solids*, Martinus Nijhoff.

Nomura, S., 1982, "Statistical Aspects of Heterogeneous Materials," *Proceedings of Fourth International Conference on Composite Materials*, Vol. 2, pp. 1083–1089.

Nomura, S., and Chou, T.-W., 1984, "Bounds for Elastic Moduli of Multiphase Short-Fiber Composites," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 540–545.

Oshima, N., and Nomura, S., 1985, "A Method to Calculate Effective Modulus of Hybrid Composite Materials," *Journal of Composite Materials*, Vol. 19, pp. 287–293.

Richard, J. C., 1975, Society of Plastics Engineers, EPS Division Meeting, Akron, Ohio.

APPENDIX

The strain field inside an ellipsoidal inclusion surrounded by an unbounded and homogeneous medium can be solved by Eshelby's method. According to the solution by this method, the strain ϵ_{il} is uniform inside the inclusion and can be expressed as

$$\epsilon = [\mathbf{I} + \mathbf{S}\{(\mathbf{C}_m - \mathbf{C}_i)\mathbf{S} - \mathbf{C}_m\}^{-1}(\mathbf{C}_i - \mathbf{C}_m)]\langle\epsilon\rangle \qquad (A1)$$

where $\mathbf{I}$ is the identity tensor whose components are $I_{ijkl} = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})/2$. The inverse of the fourth rank tensor A_{ijkl} is understood as $A_{ijkl}A^{-1}_{klmn} = I_{ijmn}$. The fourth order tensors $\mathbf{C}_m$ and $\mathbf{C}_i$ are the elastic moduli of the surrounding medium and the inclusion, respectively. The second rank tensor $\langle\epsilon\rangle$ is the strain field applied to the material at infinity. The fourth order tensor, $\mathbf{S}$, is called the Eshelby tensor and its explicit form was given by Lin and Mura (1973) where the surrounding medium is transversly isotropic. It is assumed that the x_1 axis coincides with the major axis of the ellipsoidal inclusion expressed by $(x_1/t)^2 + (x_2)^2 + (x_3)^2 = 1$. The quantity, t, is the inclusion aspect ratio.

According to Lin and Mura (1973), it is convenient to decompose $\mathbf{S}$ as

$$S_{ikmn} = C^m_{jlmn}(M_{ikjl} + M_{kijl})/8\pi \qquad (A2)$$

The fourth order tensor $\mathbf{M}$ is a function of $\mathbf{C}_m$ and t. The nonzero components of $\mathbf{M}$ are given below:

$$\begin{aligned}
M_{1111} &= 4\pi\int_0^1 \Delta t^2x^2\{d(1-x^2)+ft^2x^2\}\{e(1-x^2)\\
&\quad + ft^2x^2\}dx\\
M_{2222} = M_{3333} &= \pi/2\int_0^1 \Delta(1-x^2)[\{f(1-x^2)\\
&\quad + ht^2x^2\}\{(3e+d)(1-x^2)+4ft^2x^2\}\\
&\quad - g^2t^2x^2(1-x^2)]dx\\
M_{2323} = M_{3232} &= \pi/2\int_0^1 \Delta(1-x^2)[\{f(1-x^2)\\
&\quad + ht^2x^2\}\{(e+3d)(1-x^2)+4ft^2x^2\}\\
&\quad - 3g^2t^2x^2(1-x^2)]dx\\
M_{2121} = M_{3131} &= \pi/2\int_0^1 \Delta t^2x^2[\{(d+e)(1-x^2)\\
&\quad + 2ft^2x^2\}\{f(1-x^2)+ht^2x^2\}\\
&\quad - g^2t^2x^2(1-x^2)]dx \qquad (A3)\\
M_{1313} = M_{1212} &= 2\pi\int_0^1 \Delta(1-x^2)\{d(1-x^2)\\
&\quad + ft^2x^2\}\{e(1-x^2)+ft^2x^2\}dx\\
M_{2233} &= \pi/2\int_0^1 \Delta(1-x^2)^2[g^2t^2x^2-(d-e)\{f(1\\
&\quad - x^2)+ht^2x^2\}]dx\\
M_{3311} = M_{2211} &= (-2\pi)\int_0^1 \Delta gt^2x^2(1-x^2)\{e(1\\
&\quad - x^2)+ft^2x^2\}dx
\end{aligned}$$

where

$$\begin{aligned}
\Delta^{-1} &= \{e(1-x^2)+ft^2x^2\}[\{d(1-x^2)\\
&\quad + ft^2x^2\}\{f(1-x^2)+ht^2x^2\}-g^2t^2x^2(1-x^2)]
\end{aligned}$$

It is known that there are five independent components for $\mathbf{C}_m$ when $\mathbf{C}_m$ is transversely isotropic. They are denoted as

$$\begin{aligned}
d &= C^m_{2222} = C^m_{3333}\\
e &= (C^m_{2222} - C^m_{2233})/2\\
f &= C^m_{1212} = C^m_{1313}\\
g &= C^m_{1122} = C^m_{1212}\\
h &= C^m_{1111}
\end{aligned} \qquad (A4)$$

D. G. Roddeman
Research Associate.

J. Drukker
Professor.

Department of Anatomy and Embryology,
University of Limburg,
6200 MD Maastricht,
The Netherlands

C. W. J. Oomens
Senior Research Associate.

J. D. Janssen
Professor.

Department of Mechanical Engineering,
Eindhoven University of Technology,
5600 MB Eindhoven, The Netherlands

The Wrinkling of Thin Membranes: Part I—Theory

A method to describe the stress situation in a wrinkled membrane is presented. In this paper it will be shown that a special deformation tensor can be chosen which leads to the correct stress state of a membrane after wrinkling when it is substituted in the constitutive equation. The method can be used for anisotropic membranes in geometrically and physically nonlinear analysis. The case of simple shear and stretching of a membrane is considered to illustrate the potency of the method.

Introduction

In modern technology thin membranes are often used as construction elements. Examples can be found in aircraft and spacecraft applications. The building industry also uses membranes (fabric constructions). The objective of our research is to study the force transmission from muscle to bone near the elbow joint. The connective tissue structures which connect contractile elements from the muscle to the bones often consist of thin membrane-like structures. Because of their geometry the membrane-like tissues will wrinkle easily. This wrinkling has much influence on the stress state and the force transmission.

Models describing the mechanical behavior of membranes are usually based on the assumption that membranes have zero flexural stiffness. In normal membrane theory, however, negative stresses are possible. A membrane theory which accounts for wrinkling does not allow any negative stress to appear. When a negative stress is about to appear the membrane will wrinkle. A model for the stress field after wrinkling is a so-called tension field. By definition a tension field is uniaxial in the sense that it has only one nonzero principal stress component. In the direction perpendicular to the lines of tension the membrane is wrinkled.

The modelling was apparently started by Wagner (1929). He tried to explain the behavior of thin metal webs and spars carrying a shear load well in excess of the initial buckling value. Many authors (for example Reissner, 1938; Kondo et al., 1955; Mansfield, 1970, 1977) contributed to the geometrically linear analysis. This modelling is based on the theorem that the lines of tension in a wrinkled membrane are exactly in a direction for which the strain energy of the membrane is at a maximum. In this situation the wrinkling of isotropic and anisotropic membranes can be described (Mansfield, 1977). In geometrically nonlinear theory the analysis is more complex. Wu (1981) presented a model describing the wrinkling of membranes in finite plane-stress theory. He modified the deformation tensor by introducing an extra parameter. The value of this parameter was determined by the condition that the stress in wrinkling direction is zero. The modification of the deformation tensor was chosen in a way that the principal Cauchy directions did not change because of the wrinkling, which is only true when the material is isotropic.

Since the connective tissue structures we study may show large deformations and anisotropy, it was necessary to develop a new model capable of dealing with these phenomena. In this paper a detailed discussion of the theory is given.

Theory

Let us assume that: (*a*) Plane-stress theory can be used; (*b*) Flexure of the membrane does not introduce stresses in the membrane; (*c*) The membrane is not able to support any negative stress. If a negative stress is about to appear the membrane will wrinkle at once.

Although not essential for the theory we restrict ourselves to materials which behave "Cauchy-elastic". Thus, if it is taken into account that constitutive equations have to be objective, we are allowed to write for the Cauchy stress tensor σ:

$$\sigma = 1/J\mathbf{F}.\mathbf{H}(\mathbf{E}).\mathbf{F}^c \qquad (1)$$

with: $\mathbf{F}$ the deformation tensor; $J = \det(\mathbf{F})$ the Jacobian of the deformation tensor; $\mathbf{E}$ the Green-Lagrange strain tensor; and $\mathbf{H}$ a tensor function of $\mathbf{E}$.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 15, 1986; final revision April 20, 1987.

Paper No. 87-WA/APM-24.

Notice that the deformation tensor $\mathbf{F}$ can only be used in equation (1) if it contains the real deformations of a membrane. If however, a theoretical model does not account for wrinkling, deformations $\mathbf{F}$ may occur which result in negative Cauchy stresses. In reality these deformations will not occur because the membrane will wrinkle. The exact shape of the membrane after wrinkling is not definable with our theory. However, it is possible to use a special modified deformation tensor in the constitutive equation which results in the real Cauchy stresses.

Consider a small part of a membrane (Fig. 1) in the neighborhood of position $\vec{x}$. Vector $\vec{a}$ is a vector tangent to the midsurface of the membrane. In mathematical terms the assumption that no negative stresses occur can be represented by:

$$\vec{a}\cdot\sigma\cdot\vec{a}\geq 0 \tag{2}$$

The tensor σ is the real Cauchy stress tensor.

Inequality (2) means that, in equilibrium, there can be no direction with negative Cauchy stress. There is an infinite number of inequality conditions because $\vec{a}$ is arbitrary. It can be proved, however, that the following finite number of inequality conditions are necessary and sufficient in order to satisfy equation (2):

$$\vec{n}_1\cdot\sigma\cdot\vec{n}_1\geq 0 \tag{3}$$

$$\vec{n}_2\cdot\sigma\cdot\vec{n}_2\geq 0 \tag{4}$$

$$\vec{n}_1\cdot\sigma\cdot\vec{n}_2=0 \tag{5}$$

where $\vec{n}_1$ and $\vec{n}_2$ are orthonormal vectors denoting the principal directions of the real Cauchy stress tensor. So, if the membrane wrinkles, these vectors give the a priori unknown directions of the principal Cauchy stresses in the wrinkled membrane. The two inequalities represent the condition that neither of the principal Cauchy stresses can be negative. These conditions are necessary because a negative principal stress would contradict the assumption that negative stresses are not possible. These conditions are sufficient because determination of the stress in an arbitrary direction $\vec{a}$:

$$\vec{a}\cdot\sigma\cdot\vec{a}=[(\vec{a}\cdot\vec{n}_1)\vec{n}_1+(\vec{a}\cdot\vec{n}_2)\vec{n}_2]\cdot\sigma\cdot[(\vec{a}\cdot\vec{n}_1)\vec{n}_1 + (\vec{a}\cdot\vec{n}_2)\vec{n}_2]=(\vec{a}\cdot\vec{n}_1)^2\vec{n}_1\cdot\sigma\cdot\vec{n}_1+(\vec{a}\cdot\vec{n}_2)^2\vec{n}_2\cdot\sigma\cdot\vec{n}_2 \tag{6}$$

always leads to a positive stress in direction $\vec{a}$ if conditions (3) and (4) are true.

Considering conditions (3), (4), and (5), the following situations are possible. If both the principle stresses are positive the membrane is taut. If both the principle stresses are zero the membrane is slack. The only possibility left is the situation in which one principal stress is zero (for example, in the $\vec{n}_1$ direction) and the other principal stress (in the $\vec{n}_2$ direction) is positive. In the latter case conditions (3), (4), and (5) are transformed into:

$$\vec{n}_1\cdot\sigma\cdot\vec{n}_1=0 \tag{7}$$

$$\vec{n}_2\cdot\sigma\cdot\vec{n}_2>0 \tag{8}$$

$$\vec{n}_1\cdot\sigma\cdot\vec{n}_2=0 \tag{9}$$

In this situation the membrane may be wrinkled. In Fig. 2 a wrinkled membrane part is shown. The deformed configuration of this membrane part, if it would not have wrinkled, is indicated by the dotted lines. The dotted membrane part would be the result of the deformation tensor $\mathbf{F}$ corresponding with a theory which does allow for negative principal stresses. In this figure $\vec{n}_1$ again is the a priori unknown direction in which the real principal Cauchy stress is zero. The problem to be solved is the determination of the principal stress in the direction of the unit vector $\vec{n}_2$. Because of the second assumption, the stresses in the membrane part stay the same if we straighten it (by flexure only) in the plane determined by $\vec{n}_1$ and $\vec{n}_2$ (Fig. 3).

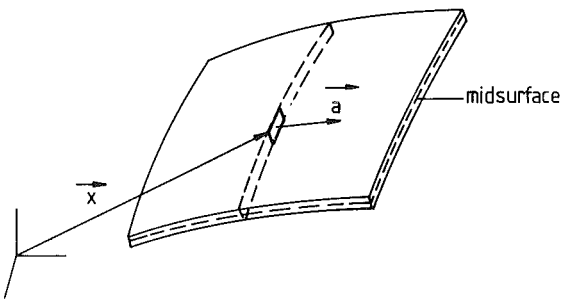


Fig. 1 Vector $\vec{a}$ is touching the midsurface of the membrane

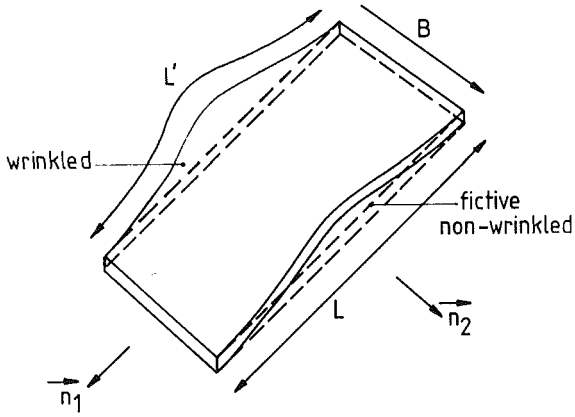


Fig. 2 A wrinkled membrane part with deformed length L' and deformed width B. Also the fictive nonwrinkled membrane part is shown (dotted lines) with length $L<L'$ and width B.

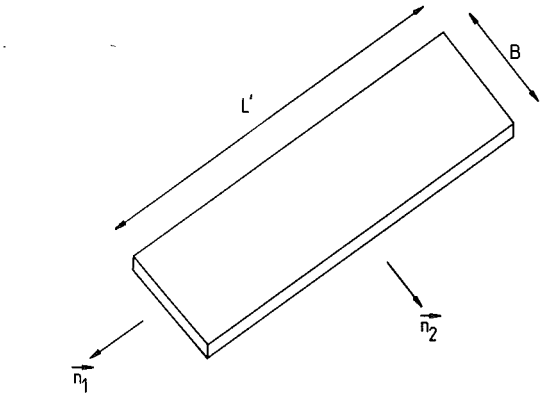


Fig. 3 The wrinkled membrane part straight in the plane determined by $\vec{n}_1$ and $\vec{n}_2$

The deformation tensor which would have given the membrane part of Fig. 3 is a deformation tensor which corresponds with the real stresses, because that membrane part contains the real stresses as we argued above. Since the membrane part of Fig. 3 has the same shape, but is only longer in the $\vec{n}_1$ direction in comparison with the fictive nonwrinkled part, that deformation tensor has to be of the form:

$$\mathbf{F}'=(\mathbf{I}+\beta\vec{n}_1\vec{n}_1)\cdot\mathbf{F} \tag{10}$$

with $\mathbf{I}$ denoting the unit tensor. The tensor $(\mathbf{I}+\beta\vec{n}_1\vec{n}_1)$ lengthens the fictive nonwrinkled membrane part to become just as long as the real wrinkled membrane part.

It should be noticed that, when the material is anisotropic, the principal directions after wrinkling in general differ from the principal directions in the fictive nonwrinkled situation.

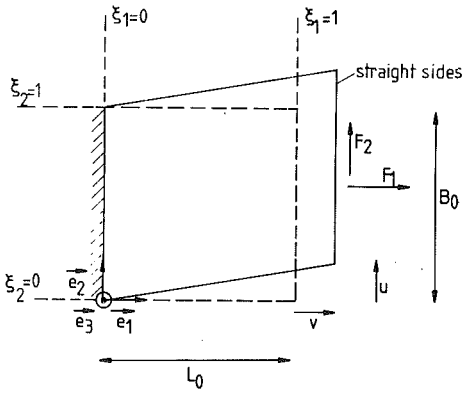


Fig. 4 Simple shear and stretching of a membrane

The parameter β and the direction of the principal frame have to be determined by using the coupled nonlinear conditions (7) and (9). The parameter β, which is never negative, is a measure of the wrinkliness of the membrane.

Summarizing we may state that the real stress state in a wrinkled membrane can be determined by using the modified deformation tensor $\mathbf{F}'$ in the constitutive equation:

$$\sigma(\mathbf{F}') = 1/J \quad \mathbf{F}' \cdot \mathbf{H}(\mathbf{E}') \cdot \mathbf{F}'^{c} \tag{11}$$

where:

$$\mathbf{E}' = 1/2(\mathbf{F}'^{c} \cdot \mathbf{F}' - \mathbf{I}); \quad J = \det(\mathbf{F}') \tag{12}$$

and where β and the direction of the principal frame have to be determined by making use of the equations (7) and (9).

At this point there is only one problem left. Suppose a deformation tensor $\mathbf{F}$ is given which, using constitutive equation (1), generates one or two negative principal Cauchy stresses. It is not immediately clear what the condition of the membrane shall be. To be able to determine whether the membrane is wrinkled or slack, the Green-Lagrange strain tensor based on $\mathbf{F}$ should be considered. When both the principal values of this strain tensor are negative, there are no directions at which material is stretched, so the membrane part is slack. If, however, in this situation, one of the principal values of this strain tensor is positive, there are directions at which material is stretched, so there are positive stresses, thus the membrane part is wrinkled.

Notice that no assumptions, such as geometrical linearity or isotropic material have been made. The preceding formulation is generally applicable.

Example: Simple Shear and Stretching of a Membrane

In this example a membrane is deformed by simple shear and stretching (Fig. 4). F_1 and F_2 are the forces required to shear the membrane over a distance u and stretch it over a distance v. Material coordinates are denoted by $\xi_i (0 \le \xi_i \le 1; i = 1,2,3)$. The initial position vector of a material point is given by:

$$\vec{x}_0 = \xi_2 B_0 \vec{e}_2 + \xi_1 L_0 \vec{e}_1 + \xi_3 D_0 \vec{e}_3 \tag{13}$$

where B_0 is the initial width, L_0 is the initial length and D_0 is the initial thickness of the membrane.

If no wrinkling takes place, the displacement field is given by:

$$\vec{u} = \xi_2 u \vec{e}_2 + \xi_1 v \vec{e}_1 + \xi_3 w \vec{e}_3 \tag{14}$$

where u is the shearing distance, v is the stretching distance, and w is the variation in the thickness of the membrane. This leads to the deformation tensor:

$$\mathbf{F} = \mathbf{I} + u/L_0 \vec{e}_2 \vec{e}_1 + v/L_0 \vec{e}_1 \vec{e}_1 + w/D_0 \vec{e}_3 \vec{e}_3 \tag{15}$$

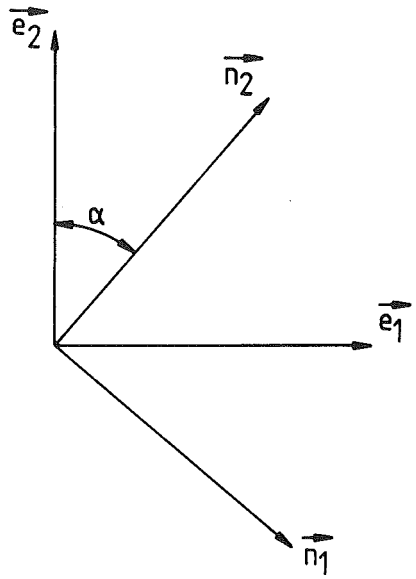


Fig. 5 The direction of the principal frame is indicated by the angle α

In simulating transversely isotropic material, the following representation of the tensor function $\mathbf{H}(\mathbf{E})$ with respect to the frame $\vec{e}_1, \vec{e}_2, \vec{e}_3$ is chosen:

$$H^e_{11} = \frac{E}{(1+v)(1-2v)}[(1-v)E^e_{11} + v(E^e_{22} + E^e_{33})] + \frac{(f-1)E}{(1-v^2)}E^e_{11} \tag{16}$$

$$H^e_{22} = \frac{E}{(1+v)(1-2v)}[(1-v)E^e_{22} + v(E^e_{11} + E^e_{33})] \tag{17}$$

$$H^e_{33} = \frac{E}{(1+v)(1-2v)}[(1-v)E^e_{33} + v(E^e_{11} + E^e_{22})] \tag{18}$$

$$H^e_{ij} = \frac{E}{1+v}E^e_{ij}, \quad i \text{ not equal } j \tag{19}$$

E^e_{ij} are the components of the Green-Lagrange strain tensor with respect to the frame $\vec{e}_1, \vec{e}_2, \vec{e}_3$. The factor f determines the measure of anisotropy of the membrane (if $f = 1$ then the membrane is isotropic, if $f > 1$ then $\vec{e}_1$ is the stiffest direction).

If the membrane wrinkles the real Cauchy stresses have to be determined by using the following set of equations:

$$\mathbf{F}' = (\mathbf{I} + \beta \vec{n}_1 \vec{n}_1) \cdot (\mathbf{I} + u/L_0 \vec{e}_2 \vec{e}_1 + v/L_0 \vec{e}_1 \vec{e}_1 + w/D_0 \vec{e}_3 \vec{e}_3) \tag{20}$$

$$\sigma(\mathbf{F}') = 1/J \mathbf{F}' \cdot \mathbf{H}(\mathbf{E}') \cdot \mathbf{F}'^{c} \tag{21}$$

where:

$$J = \det(\mathbf{F}'); \quad \mathbf{E}' = 1/2(\mathbf{F}'^{c} \cdot \mathbf{F}' - \mathbf{I}) \tag{22}$$

and in the equations (16) to (19) E^e_{ij} is replaced by E'^e_{ij}.

Denoting the direction of the principal frame by the angle α (Fig. 5), the parameters α, β, and w have to be determined by using conditions (23) and (24):

$$\vec{n}_1 \cdot \sigma(\mathbf{F}') \cdot \vec{n}_1 = 0 \tag{23}$$

$$\vec{n}_1 \cdot \sigma(\mathbf{F}') \cdot \vec{n}_2 = 0 \tag{24}$$

and the plane-stress condition:

$$\vec{e}_3 \cdot \sigma(\mathbf{F}') \cdot \vec{e}_3 = 0 \tag{25}$$

Using equations (20) to (22) with the conditions (23) to (25) leads to three nonlinear coupled equations in the unknowns α,

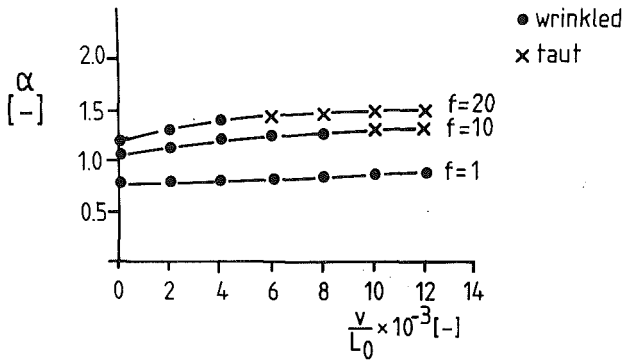


Fig. 6 The angle α as a function of the stretching distance v for different values of the anisotropy parameter f (see equation (16))

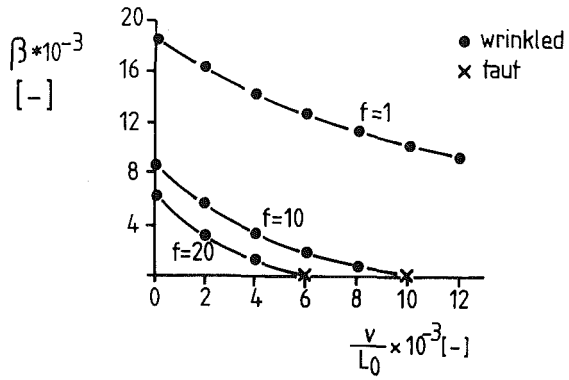


Fig. 7 The parameter β as a function of the stretching distance v for different values of the anisotropy parameter f

β, and w:

$$h_1(\alpha,\beta,w)=0 \tag{26}$$

$$h_2(\alpha,\beta,w)=0 \tag{27}$$

$$h_3(\alpha,\beta,w)=0 \tag{28}$$

The unknowns α, β, and w can be solved numerically by using equations (26) to (28), for example by means of a Newton-Raphson procedure.

We have chosen the following numerical values for the model parameters:

$L_0=100$[mm]; $B_0=100$[mm]; $D_0=1$[mm];

$$E=100[\text{N/mm}^2]; \quad \upsilon=0.3$$

The membrane is sheared with $u=5$ [mm] and then stretched until $v=1.2$ [mm]. Results for the geometrically nonlinear analysis are given in Figs. 6 to 9.

In Fig. 6 it can be seen that if the membrane stiffens in $\vec{e}_1$ direction the lines of tension tend to this direction. Notice that if the stretching is strong enough the wrinkles are pulled out of the membrane. This is the point at which the parameter β becomes zero (Fig. 7). If the membrane stiffens, the forces necessary to the deform the membrane increase (Fig. 8 and Fig. 9).

For isotropic materials ($f=1$) the formulation of Wu (1981) would generate the same results because, in this situation, equation (10) degenerates to the same modification of the deformation tensor as Wu proposes in order to describe the wrinkling of isotropic membranes. The results for anisotropic materials, however, can only be found by making use of the theoretical model given above.

Similar figures can be found in a geometrical linear analysis. These results turn out to be the same as results found by geometrical linear formulations, for example the formulation of Mansfield (1977).

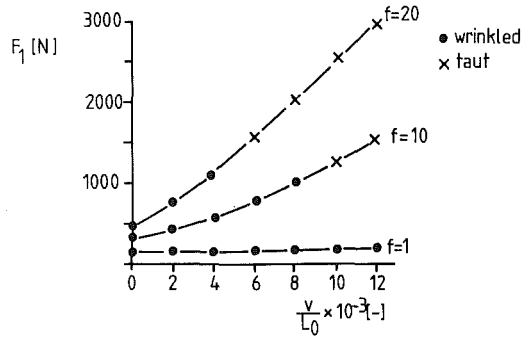


Fig. 8 The force F_1 as a function of the stretching distance v for different values of the anisotropy parameter f

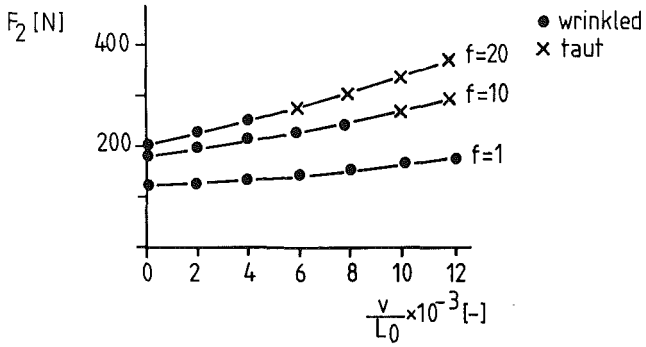


Fig. 9 The force F_2 as a function of the stretching distance v for different values of the anisotropy parameter f

Conclusions

The model given above is easy to understand and pretends to describe the general situation of the wrinkling of anisotropic membranes in geometrically nonlinear analysis. However, the preceding analysis of wrinkling membranes is only theoretical. A confrontation between this theory and experiments still has to be done.

In situations in which no analytical solutions to problems are known we want to use numerical approximation methods. We choose to use the Finite Element Method. Based on the theoretical method, a membrane element which is able to wrinkle has been developed. We plan to discuss this element in a paper to follow.

Acknowledgments

This study was supported by a grant from the Netherlands Foundation for Biological Research (BION).

References

Danielson, D. A., Natarajan, S., 1975, "Tension Field Theory and the Stress in Stretched Skin," *J. Biomechanics*, Vol. 8, pp. 135–142.

Kondo, K., Iai, T., Moriguri, S., Murasaki, T., 1955, *Memoirs of the Unifying Study of the Basic Problems in Engineering by Means of Geometry*, Vol. 1, pp. 417–441.

Mansfield, E. H., 1970, "Load Transfer Via a Wrinkled Membrane," *Proc. Roy. Soc. Lond.*, Series A 316, pp. 269–289.

Mansfield, E. H., 1977, "Analysis of Wrinkled Membranes With Anisotropic and Nonlinear Elastic Properties," *Proc. Roy. Soc. Lond.*, Series A 353, pp. 477–498.

Miller, R. K., Hedgepeth, J. M., Weingarten, V. I., Das, P., Kahyai, S., 1985, "Finite Element Analysis of Partly Wrinkled Membranes," *Computers and Structures*, Vol. 20, pp. 631–639.

Wagner, H., 1929, "Flat Sheet Metal Girders With Very Thin Metal Web," *Z. Flugtechn. Motorluftschiffahrt*, Vol. 20, pp. 200–314.

Wu, C. H., 1978, "Nonlinear Wrinkling of Nonlinear Membranes of Revolution," ASME Journal of Applied Mechanics, Vol. 45, pp. 533–538.

Wu, C. H., Canfield, T. R., 1981, "Wrinkling in Finite Plane-Stress Theory," *Q. Appl. Math.*, pp. 179–199.

D. G. Roddeman
Research Associate.

J. Drukker
Professor.

Department of Anatomy and Embryology,
University of Limburg,
6200 MD Maastricht, The Netherlands

C. W. J. Oomens
Senior Research Associate.

J. D. Janssen
Professor.

Department of Mechanical Engineering,
Eindhoven University of Technology,
5600 MB Eindhoven, The Netherlands

The Wrinkling of Thin Membranes: Part II—Numerical Analysis

Using a weighted residual method, a geometrically and physically nonlinear membrane element is derived, which can be used in the analysis of anisotropic membranes. What is special about the formulation is that the wrinkling behavior of the element is incorporated. If wrinkling occurs the stress situation in the element is determined by making use of a modified deformation tensor. A structure may have completely slack regions, leading to a singular stiffness matrix. Because of this we have chosen to use a restricted step method for the iterative solution procedure. A simple shear test is used to compare numerical and analytical results which show good agreement.

Introduction

Force transmission from muscle to bone near the elbow joint takes place by means of connective tissues which connect contractile elements from the muscle to the bones. They often consist of thin membrane-like structures. Because of their geometry, the membrane-like tissues will wrinkle easily. This wrinkling may have much influence on the stress state and the force transmission. In Roddeman et al. (1987) a mechanical model of wrinkling membranes has been presented. Wrinkling is accounted for by replacing a given deformation tensor, which would result in negative Cauchy stresses in the membrane, by a modified deformation tensor which results in the correct stress situation. With this model the wrinkling of anisotropic membranes in geometrically nonlinear analysis can be described. In this paper a membrane element will be derived which can be used for the modelling of thin structures.

The Equilibrium Conditions

The equilibrium conditions in local form may be formulated as:

$$\vec{\nabla}\cdot\boldsymbol{\sigma}=\vec{0} \tag{1}$$

where $\vec{\nabla}$ is the gradient operator with respect to the deformed configuration and $\boldsymbol{\sigma}$ is the real Cauchy stress tensor which is symmetric: $\boldsymbol{\sigma}=\boldsymbol{\sigma}^c$.

For a finite element model an integral form of equation (1) is needed. Taking the inner product of the local equilibrium conditions with an arbitrary vector function $\vec{h}$ and integrating the result over the deformed volume V leads to:

$$\int_V (\vec{\nabla}\cdot\boldsymbol{\sigma})\cdot\vec{h}dV=0 \tag{2}$$

The integral form of equation (2) is equivalent to the equilibrium conditions (1), as long as $\vec{h}$ is arbitrary. Partial integration and the application of Gauss' theorem leads to:

$$\int_V \left[\boldsymbol{\sigma}:(\vec{\nabla}\vec{h})^c\right]dV=\int_A \vec{n}_A\cdot[\boldsymbol{\sigma}\cdot\vec{h}]dA \tag{3}$$

where $\vec{n}_A$ denotes the outward unit normal on the deformed surface A of the body. Since the volume V and the surface A are a priori unknown, equation (3) is transformed to the initial configuration, for which the volume V_0 and the surface A_0 are known:

$$\int_{V_0} [\boldsymbol{\sigma}:((\mathbf{R}^{-c}\cdot\vec{\nabla}_0)\vec{h})^c]JdV_0=\int_{A_0}\vec{n}_A\cdot[\boldsymbol{\sigma}\cdot\vec{h}]J^a dA_0 \tag{4}$$

with $J=dV/dV_0$; $J^a=dA/dA_0$; and $\vec{\nabla}_0$ is the gradient operator with respect to the initial configuration.

$\mathbf{R}$ denotes the real deformation tensor of the body. It should be noticed that in a wrinkled membrane with no flexural stiffness $\mathbf{R}$ is indefinite. It can be shown that an equivalent formulation of equation (4) is:

$$\int_{V_0} [(\mathbf{R}^{-1}\cdot J\boldsymbol{\sigma}):(\vec{\nabla}_0\vec{h})^c]dV_0=\int_{A_0}\vec{n}_A\cdot[\boldsymbol{\sigma}\cdot\vec{h}]J^a dA_0 \tag{5}$$

The term $\mathbf{R}^{-1}\cdot J\boldsymbol{\sigma}$ in equation (5) is called the first Piola-Kirchhoff stress tensor $\boldsymbol{\pi}$. If a membrane wrinkles (in a principal direction, for example given by unit vector $\vec{n}_1$) the exact shape in wrinkling direction is not known, and so the real deformation tensor $\mathbf{R}$ is unknown. It can be demonstrated however that $\mathbf{R}$ may be replaced by a similar deformation tensor $\mathbf{S}$, which is known.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 15, 1986; final revision April 20, 1987.
Paper No. 87-WA/APM-25.

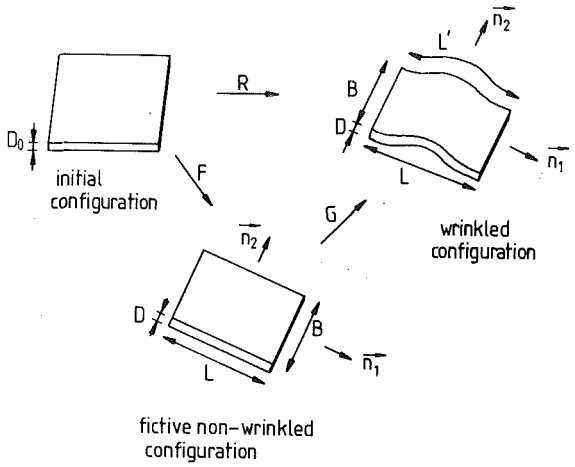


Fig. 1 The total deformation *R* is divided into *F* and *G*

Consider the deformation of a wrinkled membrane, as a deformation from the initial configuration to an imaginary nonwrinkled membrane (with corresponding deformation tensor **F**), followed by the wrinkling of the membrane, with corresponding deformation tensor **G** (Fig. 1). So, we write:

$$\mathbf{R} = \mathbf{G} \cdot \mathbf{F} \tag{6}$$

where:

$$\mathbf{G} = \vec{n}_2\vec{n}_2 + g_{11}\vec{n}_1\vec{n}_1 + g_{13}\vec{n}_1\vec{n}_3 + g_{31}\vec{n}_3\vec{n}_1 + g_{33}\vec{n}_3\vec{n}_3 \tag{7}$$

The frame $\vec{n}_1, \vec{n}_2, \vec{n}_3$ is the principal Cauchy frame where $\vec{n}_3$ is perpendicular to the plane of the fictive nonwrinkled membrane. The unknown terms $g_{ij}\vec{n}_i\vec{n}_j$ only cause deformations in the plane determined by $\vec{n}_1$ and $\vec{n}_3$ and define the wrinkles of the membrane part. Using the fact that there are only stresses in the $\vec{n}_2$ direction, it is obvious that:

$$J\sigma = (J\sigma_{22})\vec{n}_2\vec{n}_2 \tag{8}$$

and it can be shown that:

$$\begin{aligned}\pi &= \mathbf{R}^{-1} \cdot (J\sigma) = \mathbf{F}^{-1} \cdot \mathbf{G}^{-1} \cdot (J\sigma) \\ &= \mathbf{F}^{-1} \cdot \{\vec{n}_2\vec{n}_2 + g_{11}\vec{n}_1\vec{n}_1 + g_{13}\vec{n}_1\vec{n}_3 + g_{31}\vec{n}_3\vec{n}_1 \\ &\quad + g_{33}\vec{n}_3\vec{n}_3\}^{-1} \cdot (J\sigma_{22})\vec{n}_2\vec{n}_2 = \mathbf{F}^{-1} \cdot (J\sigma_{22})\vec{n}_2\vec{n}_2 = \mathbf{F}^{-1} \cdot (J\sigma)\end{aligned} \tag{9}$$

In a similar way the deformation tensor **F** can be regarded to be a deformation from the initial configuration to a nonwrinkled membrane, still with initial thickness D_0, with corresponding deformation tensor **S**, followed by a change of the thickness from D_0 to its real value D:

$$\mathbf{F} = (\vec{n}_1\vec{n}_1 + \vec{n}_2\vec{n}_2 + D/D_0\vec{n}_3\vec{n}_3) \cdot \mathbf{S} \tag{10}$$

After substituting equation (10) in (9) it is found that:

$$\begin{aligned}\pi = \mathbf{F}^{-1} \cdot (J\sigma) &= \mathbf{S}^{-1} \cdot (\vec{n}_1\vec{n}_1 + \vec{n}_2\vec{n}_2 + D/D_0\vec{n}_3\vec{n}_3)^{-1} \cdot (J\sigma_{22})\vec{n}_2\vec{n}_2 \\ &= \mathbf{S}^{-1} \cdot (J\sigma_{22})\vec{n}_2\vec{n}_2\end{aligned}$$

thus:

$$\pi = \mathbf{S}^{-1} \cdot (J\sigma) \tag{11}$$

It is easy to show that the last equation is also true in case the membrane does not wrinkle. The advantage of equation (11) is that **S** does not contain variations in the membrane thickness anymore, which will prove to be useful for the linearization of the equations to follow.

The integral form of the equilibrium conditions now can be written as:

$$\int_{V_0} \pi : (\vec{\nabla}_0\vec{h})^c dV_0 = \int_{A_0} \vec{k}_0 \cdot \vec{h} dA_0 \tag{12}$$

where π is given by equation (11) and $\vec{k}_0 = J^a\vec{n}_A \cdot \sigma$, which is the force on the deformed surface transformed to the undeformed surface.

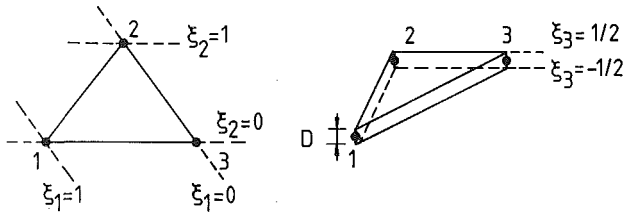


Fig. 2 A three-noded element. ξ_1, ξ_2, and ξ_3 are material coordinates.

An estimated solution of the equilibrium condition in integral form (12) is marked with a superscript *. In general the estimated solution will not satisfy equation (12) exactly, so a better solution is searched for. The difference between the estimated solution and a solution satisfying equilibrium condition (12) is indicated by a delta, so:

$$\pi = \pi^* + \delta\pi$$
$$\mathbf{S} = \mathbf{S}^* + \delta\mathbf{S}$$

, etc. Thus equation (12) may also be written as:

$$\int_{V_0} (\pi^* + \delta\pi) : (\vec{\nabla}_0\vec{h})^c dV_0 = \int_{A_0} (\vec{k}_0^* + \delta\vec{k}_0) \cdot \vec{h} dA_0 \tag{13}$$

Until now equation (13) is perfectly equivalent to the equilibium conditions (1). The difference between the displacement field satisfying the equilibrium conditions and the estimated displacement field is regarded to be the primary unknown of equation (13). Let us assume that the estimated solution is close to the real solution. Then equation (13) can be linearized with respect to the primary unknown.

According to equation (11), the difference in π in the estimated solution and equilibrium is approximately given by:

$$\delta\pi \approx \delta(\mathbf{S}^{-1}) \cdot (J\sigma)^* + \mathbf{S}^{*-1} \cdot \delta(J\sigma) \tag{14}$$

Since:

$$\delta(\mathbf{S}^{-1}) \approx -\mathbf{S}^{*-1} \cdot \delta\mathbf{S} \cdot \mathbf{S}^{*-1} \tag{15}$$

it is easily shown that linearization of equation (13) leads to:

$$\begin{aligned}&\int_{V_0} [\{(\vec{\nabla}_0\vec{h})^c \cdot \mathbf{S}^{*-1}\} : \{\delta(J\sigma)\} \\ &\qquad - \{\mathbf{S}^{*-1} \cdot (J\sigma)^* \cdot (\vec{\nabla}_0\vec{h})^c\} : \{\mathbf{S}^{*-1} \cdot \delta\mathbf{S}\}] dV_0 \\ &- \int_{A_0} \delta\vec{k}_0 \cdot \vec{h} dA_0 \approx -\int_{V_0} \pi^* : (\vec{\nabla}_0\vec{h})^c dV_0 + \int_{A_0} \vec{k}_0^* \cdot \vec{h} dA_0\end{aligned} \tag{16}$$

We will solve (16) by means of the Finite Element Method. The observed mechanical system is divided into a number of elements of finite dimensions. Often it is possible to define all kinds of elements and to give a general derivation of the equations without considering the type of the element. In the present case however it is convenient to have an element with constant strains and stresses. Otherwise the element might be divided into wrinkling and nonwrinkling zones, which would make the analysis unfeasible. Thus, we have decided to use a triangular, three-noded, constant strain element and specialize the derivation for this element. In Fig. 2 the element is shown. Equation (16) will be analyzed for one element (thus V_0 is the initial volume of the element and A_0 the initial surface).

The position of a material point of the element is given by:

$$\vec{x} = \psi_k\vec{x}_k + \xi_3 D\vec{n}_3 \tag{17}$$

where ψ_k are shape functions, ξ_3 is the material coordinate in the direction perpendicular to the plane of the membrane, and $\vec{x}_k$ are position vectors of the nodal points. D denotes the thickness of the membrane. The Einstein summation convention is used, i.e., when an index occurs twice in a product term, this implies summation with respect to all its possible values. Normally the possible values of an index are 1, 2, and 3. However, indices which are Greek characters only can take the values 1 and 2.

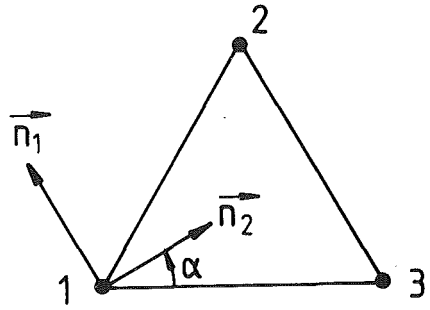


Fig. 3 The direction of the principal Cauchy frame is indicated by the angle α

The shape functions depend on the material coordinates ξ_1 and ξ_2:

$$\psi_1 = \xi_1; \psi_2 = \xi_2; \psi_3 = 1 - \xi_1 - \xi_2 \tag{18}$$

First, an expression will be derived for the increment $\delta\mathbf{S}$ (equation (16)). Base vectors in the initial configuration are given by:

$$\vec{c}_{01} = \partial\vec{x}_0/\partial\xi_1 = \vec{x}_{01} - \vec{x}_{03} \tag{19}$$

$$\vec{c}_{02} = \partial\vec{x}_0/\partial\xi_2 = \vec{x}_{02} - \vec{x}_{03} \tag{20}$$

$$\vec{c}_{03} = \partial\vec{x}_0/\partial\xi_3 = D_0\vec{n}_{03} \tag{21}$$

where the subscript 0 denotes values in the initial configuration. The reciprocal vectors are given by:

$$\vec{\gamma}_{01} = 1/c\vec{c}_{02}{}^*\vec{c}_{03} \tag{22}$$

$$\vec{\gamma}_{02} = 1/c\vec{c}_{03}{}^*\vec{c}_{01} \tag{23}$$

$$\vec{\gamma}_{03} = 1/c\vec{c}_{01}{}^*\vec{c}_{02} \tag{24}$$

where

$$c = \vec{c}_{03}\cdot(\vec{c}_{01}{}^*\vec{c}_{02}) \tag{25}$$

It can be derived that $\mathbf{F}$ is given by:

$$\mathbf{F} = \partial\psi_k/\partial\xi_\tau\vec{x}_k\vec{\gamma}_{0_\tau} + D\vec{n}_3\vec{\gamma}_{03} \tag{26}$$

From equations (10) and (26) it follows that:

$$\mathbf{S} = \partial\psi_k/\partial\xi_\tau\vec{x}_k\vec{\gamma}_{0_\tau} + D_0\vec{n}_3\vec{\gamma}_{03} \tag{27}$$

Using equation (27) leads to:

$$\delta\mathbf{S} = \partial\psi_k/\partial\xi_\tau\delta\vec{u}_k\vec{\gamma}_{0_\tau} + D_0\delta\vec{n}_3\vec{\gamma}_{03} \tag{28}$$

where $\vec{u}_k$ is the nodal displacement of node k. Secondly, expressions will be derived for the increment $\delta\vec{n}_3$ (equation (28)) and the increment $\delta(J\boldsymbol{\sigma})$ (equation (16)).

In Roddeman et al. (1987) it is shown that the real stresses in a wrinkled membrane can be determined by modifying the deformation tensor in the constitutive equation:

$$(J\boldsymbol{\sigma}) = \mathbf{F}'\cdot\mathbf{H}(\mathbf{E}')\cdot\mathbf{F}'^c \tag{29}$$

with:

$$\mathbf{F}' = (\mathbf{I} + \beta\vec{n}_1\vec{n}_1)\cdot\mathbf{F} \tag{30}$$

where $\mathbf{I}$ is the unit tensor. The direction of the principal Cauchy frame is denoted by the angle α (see Fig. 3).

The parameters α and β and the element's thickness D are determined numerically by using the coupled conditions:

(*a*) there is no stress in wrinkling direction:

$$\vec{n}_1\cdot(J\boldsymbol{\sigma})\cdot\vec{n}_1 = 0 \tag{31}$$

(*b*) the frame $\vec{n}_1$, $\vec{n}_2$, $\vec{n}_3$ is the principal frame:

$$\vec{n}_1\cdot(J\boldsymbol{\sigma})\cdot\vec{n}_2 = 0 \tag{32}$$

(*c*) membranes are in a state of plane stress:

$$\vec{n}_3\cdot(J\boldsymbol{\sigma})\cdot\vec{n}_3 = 0 \tag{33}$$

Using equations (29) to (33), it is possible to derive expressions for the increments in $\vec{n}_3$ and $J\boldsymbol{\sigma}$, which can be represented by:

$$\delta\vec{n}_3 \approx \mathbf{N}^*_{3k}\cdot\delta\vec{u}_k \tag{34}$$

$$\delta(J\boldsymbol{\sigma}) \approx {}^3\boldsymbol{\sigma}^*_k\cdot\delta\vec{u}_k \tag{35}$$

Thus, increments in these terms are depending on increments in the primary unknown nodal displacements, via equations (34) and (35). The second order tensor $\mathbf{N}^*_{3k}$ (in equation (34)) and the third order tensor ${}^3\boldsymbol{\sigma}^*_k$ (in equation (35)) depend on the current estimation of the nodal displacements. Using equations (34) and (35) the discretized form of equation (16) may be written as:

$$\int_{V_0}[\{(\vec{\nabla}_0\vec{h})^c\cdot\mathbf{S}^{*-1}\}:\{{}^3\boldsymbol{\sigma}^*_k\cdot\delta\vec{u}_k\} - \{\mathbf{S}^{*-1}\cdot(J_\sigma)^* \cdot(\vec{\nabla}_0\vec{h})^c\}:\{S^{*-1}\cdot(\partial\psi_k/\partial\xi_\tau\delta\vec{u}_k\vec{\gamma}_{0_\tau} + D_0\mathbf{N}^*_{3k} \cdot\delta\vec{u}_k\vec{\gamma}_{03})\}]dV_0 - \int_{A_0}\delta\vec{k}_0\cdot\vec{h}dA_0 \approx -\int_{V_0}\pi^*:(\vec{\nabla}_0\vec{h})^c dV_0 + \int_{A_0}\vec{k}^*_0\cdot\vec{h}dA_0 \tag{36}$$

Because of the approximation for the displacement field it is no longer possible to fulfill exactly equation (36) and the boundary conditions. However, an approximation can be made by choosing well determined weighting functions for $\vec{h}$. Often good results can be obtained by using the shape functions, introduced in equation (17), as weighting functions. So we introduce:

$$\vec{h} = \psi_l\vec{h}_l \tag{37}$$

where $\vec{h}_l$ is the value of the function $\vec{h}$ in the nodal point l.

Using equation (37) and the fact that all terms in the volume integral in expression (36) are constant over the volume, expression (36) may be replaced with:

$$\vec{h}_l\cdot[V_0\partial\psi_l/\partial\xi_v\vec{e}_p\vec{\gamma}_{0_v}\cdot S^{*-1}\cdot{}^3\sigma^{*rc}_k\cdot\vec{e}_p - V_0\partial\psi_l/\partial\xi_v\partial\psi_k/\partial\xi_\tau\vec{\gamma}_{0_\tau}\cdot\vec{e}_p(J\boldsymbol{\sigma})^*\cdot\mathbf{S}^{*-c}\cdot\vec{e}_p\vec{\gamma}_{0_v}\cdot\mathbf{S}^{*-1} - V_0D_0\partial\psi_l/\partial\xi_v\vec{\gamma}_{03}\cdot\vec{e}_p(J\boldsymbol{\sigma})^*\cdot\mathbf{S}^{*-c}\cdot\vec{e}_p\vec{\gamma}_{0_v}\cdot\mathbf{S}^{*-1} \cdot\mathbf{N}^*_{3k}]\cdot\delta\vec{u}_k - \int_{A_0}\delta\vec{k}_0\cdot\vec{h}dA_0 \approx -\int_{V_0}\pi^*:(\vec{\nabla}_0\vec{h})^c dV_0 + \int_{A_0}\vec{k}^*_0\cdot\vec{h}dA_0 \tag{38}$$

in which use has been made of the identity:

$$({}^3\mathbf{A}^{rc}\cdot\vec{w})\cdot\vec{v} = ({}^3\mathbf{A}\cdot\vec{v})\cdot\vec{w}$$

where $\vec{w}$ and $\vec{v}$ are arbitrary vectors. Furthermore $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$ denote an orthonormal set of vectors. The first term on the right of equation (38) will be written as:

$$\int_{V_0}\pi^*:(\vec{\nabla}_0\vec{h})^c dV_0 = \vec{h}_l\cdot[V_0\vec{e}_p\partial\psi_l/\partial\xi_v\vec{\gamma}_{0_v}\cdot\pi^*\cdot\vec{e}_p] = \vec{h}_l\cdot\vec{f}^*_l \tag{39}$$

in which the vector $\vec{f}^*_l$ is defined, which is the equivalent nodal force in node l due to the stresses in the membrane. We only consider constant forces on the nodal points. Thus:

$$\int_{A_0}\delta\vec{k}_0\cdot\vec{h}dA_0 = 0 \tag{40}$$

and:

$$\int_{A_0}\vec{k}^*_0\cdot\vec{h}dA_0 = \vec{h}_l\cdot\vec{f}_l \tag{41}$$

where $\vec{f}_l$ is the prescribed nodal force on node l.

According to equations (39) to (41), expression (38) may be written as:

$$\vec{h}_l\cdot\mathbf{K}^*_{lk}\cdot\delta\vec{u}_k \approx \vec{h}_l\cdot[\vec{f}_l - \vec{f}^*_l] \tag{42}$$

where $\mathbf{K}^*_{lk}$ contains the term $[V_0 \ldots\ldots\ldots \mathbf{N}^*_{3k}]$ from equation (38). Equation (42) is satisfied by arbitrary $\vec{h}_l$ if:

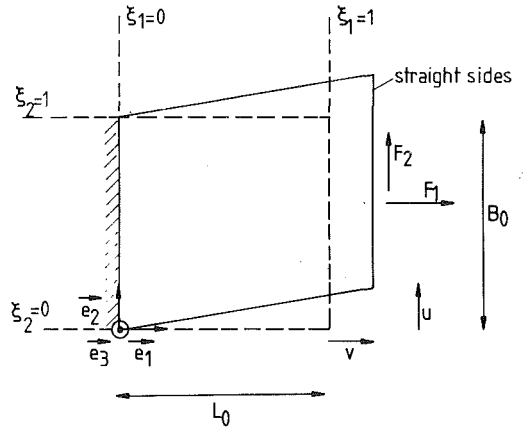


Fig. 4 Simple shear and stretching of a membrane

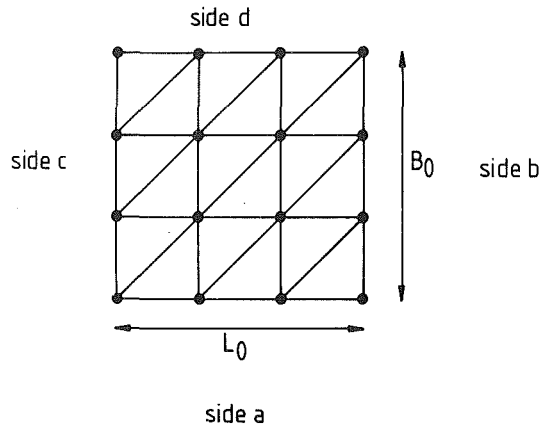


Fig. 5 An element mesh with 18 elements

$$\mathbf{K}^*_{lk} \cdot \delta \vec{u}_k \approx \vec{f}_l - \vec{f}^*_l \tag{43}$$

The matrix representation of the first term on the left side with respect to a certain base is called the stiffness matrix of the element. The right side contains the residual nodal forces.

Logical Structure of the Element

Given a new estimation of the nodal displacements, the element decides on the following criteria whether it is taut, wrinkled, or slack. If both the principal Cauchy stresses in an analysis without wrinkling are non-negative, the element is taut. In this situation the stiffness matrix and the equivalent nodal forces are determined by normal analysis (i.e., without wrinkling terms). If at least one of the principal Cauchy stresses in the analysis without wrinkling would be negative, and one principal Green-Lagrange strain in the analysis without wrinkling would be positive, the element is wrinkled. Then the new stiffness matrix and equivalent nodal forces are determined on the basis of wrinkling analysis. Otherwise the element is slack and the stiffness matrix and the equivalent nodal forces only contain zeros.

The Iterative Process

Choosing a base, assemblage of the equations (43) of all the elements of the body and elimination of prescribed nodal displacements leads to:

$$\underline{\underline{K}}^* \cdot \delta \underline{u} = \underline{f} - \underline{f}^* \tag{44}$$

with

$\underline{\underline{K}}^*$ the stiffness matrix of the structure;
$\delta \underline{u}$ the total column of incremental free nodal displacements;
$\underline{f}$ the total column of prescribed forces on the body;
$\underline{f}^*$ the total column of equivalent nodal forces due to the stresses in the elements.

Since there may be completely slack regions in the structure, the total stiffness matrix may be singular and Newton-Raphson iteration cannot be used. A globally convergent iterative procedure, based on a minimization problem, which can be used, for example, is a restricted step method where the column $\delta \underline{u}$ is the solution of:

$$(2\underline{\underline{K}}^{*T} \cdot \underline{\underline{K}}^* + \theta I) \cdot \delta \underline{u} = 2\underline{\underline{K}}^{*T} \cdot (\underline{f} - \underline{f}^*) \tag{45}$$

The parameter θ is chosen each iteration step such that $(2\underline{\underline{K}}^{*T} \cdot \underline{\underline{K}}^* + \theta I)$ is positive definite (which is true if θ is positive) and such that the norm of $(\underline{f} - \underline{f}^*)$ decreases. For a theoretical background see Fletcher (1980).

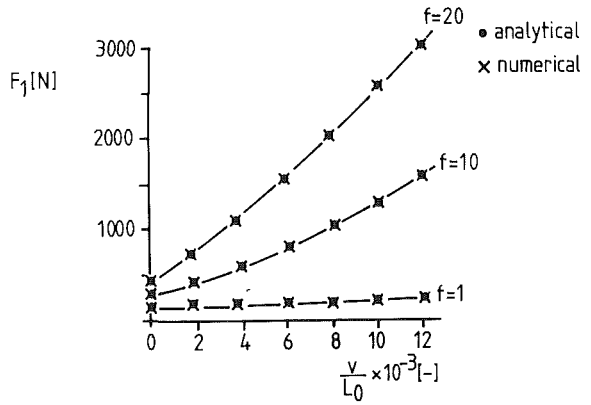


Fig. 6 The force F_1 as a function of the stretching distance v for different values of the anisotropy parameter f

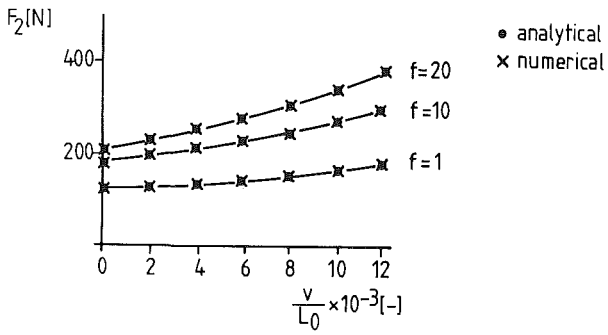


Fig. 7 The force F_2 as a function of the stretching distance v for different values of the anisotropy parameter f

Test Problems

In Roddeman et al. (1987) the analytical solution of the wrinkling of a membrane deformed by a simple shear and stretching is derived. In Fig. 4 the simple shear test is illustrated.

Again we have studied the behavior of the membrane when, with constant simple shear, the membrane is stretched. An element mesh with 18 elements is used (Fig. 5). The nodes on the sides of the mesh are prescribed displacements according to the analytical solution, so the nodal displacements in the inner area should converge to the analytical solution and the analytical results should be regained. Figure 6 and Fig. 7 compare analytical and numerical results. Tranversally isotropic material is used for the membrane. The parameter f determines the measure of anisotropy (see Roddeman et al., 1987). If $f=1$ then the membrane is isotropic, if $f>1$ then $\vec{e}_1$ is the stiffest direction. It can be seen that the agreement between the analytical and numerical results is very good.

Conclusions

The Finite Element Method can be used in geometrically nonlinear analysis of anisotropic membranes which wrinkle. Test problems show good agreement between analytical and numerical results. Since completely slack regions in the structure lead to a singular stiffness matrix, the iterative procedure should be chosen with care. Comparison between numerical and experimental results still has to be done. This is part of our present work and will be published later.

Acknowledgments

This study was supported by a grant from the Netherlands Foundation for Biological Research (BION).

References

Fletcher, R., 1980, *Practical Methods of Optimization: Unconstrained Optimization*, Vol. 1, Wiley.

Roddeman, D. G., Oomens, C. W. J., Janssen, J. D., and Drukker, J., 1987, "The Wrinkling of Thin Membranes: Part 1 – Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54.

Bisynchronous Torsional Vibrations in Rotating Shafts

O. Bernasconi
Research Engineer,
Swiss Federal Institute of Technology,
Lausanne, Switzerland CH-1015

In this study, the intrinsic behavior of rotating shafts with residual unbalance is considered. The longitudinal component of the angular momentum caused by synchronous precession (whirling) induces torsional vibrations with a frequency of twice the rotation frequency (bisynchronous). The nonlinear term which represents this coupling is characteristic of the asymmetrical aspect of rotating shaft kinematics. This result has been obtained analytically and confirmed experimentally.

1 Introduction

The effect of the mutual influence between transverse and torsional vibrations of a turbogenerator shaft on the stability of its motion has been investigated by Tondl (1965). Taking into account a synchronizing torque acting on the rotor of the alternator, the author shows that there are speed intervals where the vibrations due to the residual unbalance of turbine and generator rotors are unstable. Broniarek (1968) rightly shows the importance of the kinematic properties of the system considering the possibility of using Euler's angles in the case of rotating shafts. The disturbances in the system are caused only by the rotor unbalance. The main result is that torsional vibrations are excited parametrically by transverse vibrations. However, the quantitative expression of this result is complicated and does not allow concrete conclusions to be obtained. A shaft with a continuous mass distribution and a residual eccentricity has also been considered by Gasch et al. (1979). As in the Laval-Jeffcott rotor, which is not subjected to the gyroscopic effect, torsional vibrations can be caused only by one nonlinear coupling term, proportional to the eccentricity. It is shown that this term does not oscillate and therefore can cause only quasi-static shaft torsion, in a first approximation. Inversely, Kellenberger (1980) shows that in the presence of an oscillator driving torque, transverse vibrations can be induced by torsional vibrations if these are sufficiently large. However, a numerical application makes it possible to consider that these combined oscillations are not dangerous, because they have no particular influence on the dynamic behavior of shafts. Cohen and Porat (1985) consider an unbalanced gyroscopic rotor, driven by a shaft rigid to bending and flexible to torsion. The nonlinear equations of motion are investigated by an asymptotic method. This approach does not yield a closed solution, but permits determination of the stability conditions at the combination resonances and gives an indication that in certain cases, superimposed damping causes considerable enlargement of the instability zones.

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 11, 1986; final revision May 4, 1987.

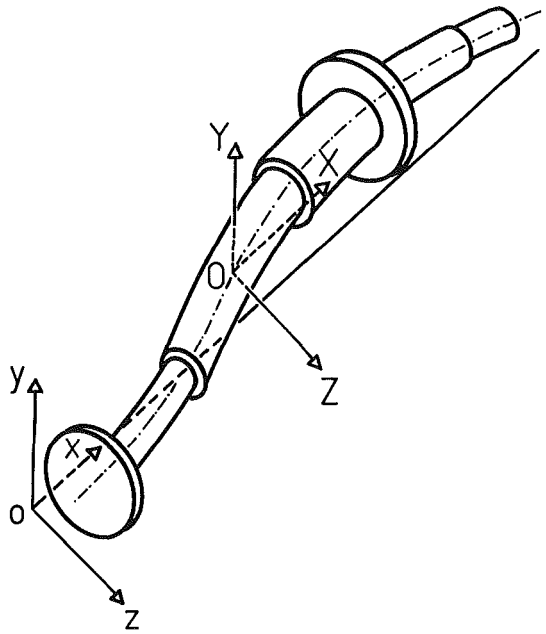


Fig. 1 Model of rotating shafts: *oxyz*: fixed (inertial) Cartesian frame *OXYZ*: Cartesian frame translating in O in relation to *oxyz*

The aim of this paper is to show a coupling phenomenon related to the gyroscopic effect, which consequently is able to induce torsional vibrations whose frequency is double that of the rotational frequency (i.e., bisynchronous). This result is obtained analytically and verified experimentally. This coupling problem requires a careful examination of the kinematic aspect of rotors.

2 Equations of Motion

The rotating shaft used as a reference in this paper has distributed mass $M(x)$, flexural stiffness $EI(x)$, torsional stiffness $GI^p(x)$, and eccentricity $e(x)$ (Fig. 1). With regard to transverse vibrations, the elastic (isotropic) behavior of deformable parts with circular cross-section is governed by the Euler-Bernoulli beam theory.

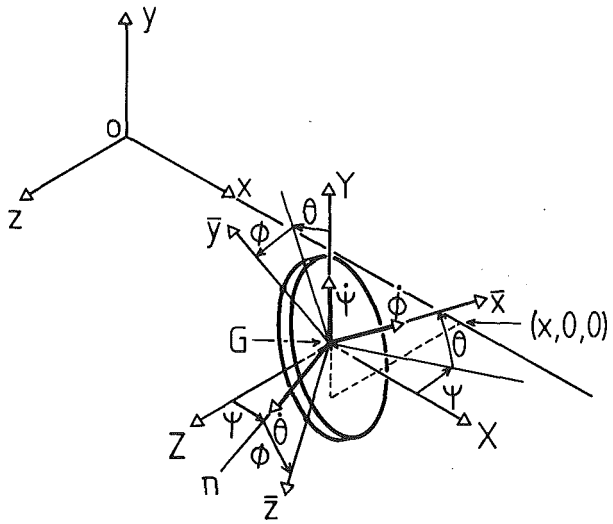


Fig 2 A disk of infinitesimal thickness *dx*, located by the displacements y^G and z^G of its mass center and by angles ψ, θ, and ϕ

As in the book written by Lalanne et al. (1983), the kinematic description of the shaft is performed cautiously. Any disk of the shaft is located first by defining a fixed (inertial) Cartesian frame *oxyz*; its position corresponds to the end of the shaft in equilibrium. Its axis *ox* coincides with the geometric centers curve of the model (or theoretical spin axis). Another Cartesian frame *GXYZ*, translating in relation to *oxyz*, is placed at the mass center *G* of the disk (Fig. 2). A third Cartesian frame $G\bar{x}\bar{y}\bar{z}$ is rigidly bound to the disk and follows all its movements. The axis $G\bar{x}$ remains constantly perpendicular to the faces of the disk and coincides with *GX* and *ox* in equilibrium.

The three Cartesian frames of reference are defined above in the same way as in the study of the dynamic behavior of airplanes and ships, where reference is made to the "nautical and aeronautical" axis (Lur'é, 1968). The orientation of the moving frame $G\bar{x}\bar{y}\bar{z}$ is defined by the angles ϕ, θ, and ψ, called, respectively, angle of roll, angle of pitch, and angle of yaw in the case of airplanes or ships. The angle ϕ characterizes the rotation of the shaft itself.

It is very important to take care that, contrary to appearances, from a dynamics point of view angles ψ and θ do not play a symmetrical role. In fact, the angular velocity vector $\dot{\psi}$ is situated on the *GY* axis which is constantly parallel to *oy* when the angular velocity vector $\dot{\theta}$ is on the moving axis *Gn*. This is valid even for small deviations from equilibrium of the disk, when the angles θ and ψ can be considered as small. This fact is frequently ignored in many studies of rotating shafts, even in recent ones (Cohen and Porat, 1985). The angle ϕ includes the "global" rotation of the shaft; it varies from zero to infinity. The a priori dissociation of rotational torsional vibrations is not only pointless when formulating the problem, but also limiting, as shown in the work done by Bernasconi and François (1982).

According to Bernoulli's assumption, plane and circular cross-sections of beam subjected to torsion and pure flexion remain plane. The model (Fig. 1) can thus be divided into an infinity of infinitesimal disks *dx*. The investigation of the behavior of one single disk makes it possible to write, after integration, the energy of the whole system.

The inertia tensor matrix (diagonal) of a disk is expressed simply in the moving axis $G\bar{x}\bar{y}\bar{z}$ bound to the disk. The influence of the eccentricity on the moment of inertia of the principal axis $G\bar{y}$ and $G\bar{z}$ is neglected. In the same moving axis, the most general form of the instantaneous angular velocity vector can be simplified because the angles θ and ψ remain small. Introducing the relationship between the position of the geometric center *O* and the mass center *G* and taking into account that the rotations can be put equal to displacement slopes, the kinetic energy of the complete model is obtained

$$T=\int_o^l \left\{\frac{1}{2}M(\dot{y}^2+\dot{z}^2)+\frac{1}{2}[J(\dot{y}'^2+\dot{z}'^2)+J^p\dot{\phi}^2] -(J^p-Me^2)\dot{\phi}\dot{z}'y'+Me\dot{\phi}[\dot{z}\cos(\phi+\alpha)-\dot{y}\sin(\phi+\alpha)]\right\}dx \quad (1)$$

where *J* is the diametral mass moment of inertia, J^p the longitudinal (polar) mass moment of inertia (both per unit length), *e* and α the eccentricity amplitude and phase of the eccentricity. Taking into account the assumptions concerning the elastic behavior of the model, its potential energy is equal to:

$$U=\int_o^l \frac{1}{2}[GI^p\phi'^2+EI(y''^2+z''^2)]dx \quad (2)$$

As the shaft is subjected only to the rotational torque (per unit length) $C(x,t)$, the last energy form of the system is the work *W* of the external torque:

$$W=\int_o^l C(x,t)\phi(x,t)dx \quad (3)$$

The application of Hamilton's principle allows the equations of motion to be obtained, writing Euler's equations for the variational problem. These equations form a nonlinear partial differential system which characterizes the basic aspect of the dynamic behavior of circular cross-section shafts with a residual unbalance, for transverse and rotational vibrations, when the shaft is subjected to an external driving torque $C(x,t)$:

$$\begin{aligned}
&M\ddot{y}+\frac{\partial^2}{\partial x^2}(EIy'')-\frac{\partial}{\partial x}J(\ddot{y}'+2\dot{\phi}\dot{z}') \\
&\qquad =Me\frac{\partial}{\partial t}[\dot{\phi}\sin(\phi+\alpha)] \\
&M\ddot{z}+\frac{\partial^2}{\partial x^2}(EIz'')-\frac{\partial}{\partial x}J(\ddot{z}'-2\dot{\phi}\dot{y}') \\
&\qquad =-Me\frac{\partial}{\partial t}[\dot{\phi}\cos(\phi+\alpha)] \qquad (4) \\
&J^p\ddot{\phi}-\frac{\partial}{\partial x}(GI^p\phi')=C+J^p\frac{\partial}{\partial t}(\dot{z}'y') \\
&\qquad +Me[\ddot{y}\sin(\phi+\alpha)-\ddot{z}\cos(\phi+\alpha)]
\end{aligned}$$

The unknowns of this differential system are the displacements *y* and *z* of the geometric centers curve as well as the rotation angle ϕ (including torsion). In the general case of stepped shafts, the derivatives should be understood in the sense of distributions (Bernasconi, 1986). This differential system includes a large number of particular cases found in the literature.

The combination of eccentricity with the transverse motion of the mass center *G* produces an "inertia moment" which appears on the right-hand side of the third equation. It can be seen that this coupling term can be calculated in the following way:

$$Mdx\ddot{\vec{r}}\wedge\vec{e}=\vec{i}_x Medx[\ddot{y}\sin(\phi+\alpha)-\ddot{z}\cos(\phi+\alpha)]$$

The second coupling term in the rotation vibrations equation is the one which is at the origin of the phenomenon described in this paper. It results from the term of the kinetic energy expression which causes the gyroscopic effect. It can be split up as follows:

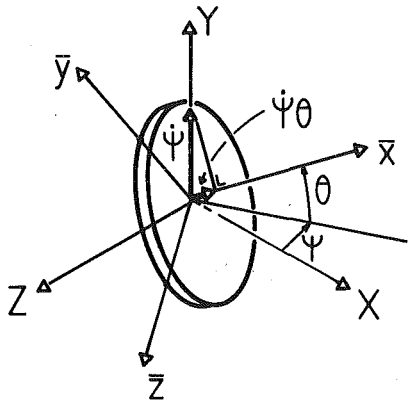


Fig. 3 Longitudinal component of the angular velocity vector $\dot{\psi}$

$$J^p \frac{\partial}{\partial t}(\dot{z}'y') = -J^p \frac{\partial}{\partial t}(\dot{\psi}\theta)$$

$$\cong -J^p \frac{\partial}{\partial t}(\dot{\psi}\sin\theta) = -\frac{\partial}{\partial t}(J^p \dot{\psi}_{\bar{x}}) = -M_{\bar{x}}$$

$\dot{\psi}_{\bar{x}}$ and $M_{\bar{x}}$ are the projections of $\vec{\dot{\psi}}$ and $\vec{M}$ on the $\bar{x}$ axis (Fig. 3). The equation, therefore, expresses the rate of change considered in time of the longitudinal component of the angular momentum, which is not caused by the rotation of the shaft but by the rotations during precessions. In accordance with the angular momentum equation, it induces a longitudinal moment. It, therefore, characterizes a typical dynamic effect, but is also a reminder that the angles θ and ψ, even if they are assumed to be small, do not play an equivalent role. This fact is partially dissimulated by equaling these angles to the displacement slopes.

3 Torsional Vibrations Induced by Transverse Vibrations

If there is a coupling it means that transverse vibrations induce torsional vibrations and vice-versa. In this paper, we examine only the possibility of transverse vibrations creating torsional vibrations.

If the transverse inertia and gyroscopic effect are neglected in the first two equations of the differential system (4), then $\ddot{y}$ and $\ddot{z}$ are removed from the equations and introduced in the eccentricity coupling term of the last equation, the differential system becomes

$$M\ddot{y} + \frac{\partial^2}{\partial x^2}(EIy'') = B\frac{\partial}{\partial t}[\dot{\phi}\sin(\phi+\alpha)]$$

$$M\ddot{z} + \frac{\partial^2}{\partial x^2}(EIz'') = -B\frac{\partial}{\partial t}[\dot{\phi}\cos(\phi+\alpha)]$$

$$J^p\ddot{\phi} - \frac{\partial}{\partial x}(GI^p\phi') = C + J^p\frac{\partial}{\partial t}(\dot{z}'y') \tag{5}$$

$$+ e\left[\frac{\partial^2}{\partial x^2}(EIz'')\cos(\phi+\alpha) - \frac{\partial^2}{\partial x^2}(EIy'')\sin(\phi+\alpha)\right]$$

with $B(x)$ the amount of unbalance per unit length ($=Me$). The method chosen consists of uncoupling the two first nonlinear equations of this last differential system (5), which is possible using a small approximation, as shown by Gasch et al. (1979).

Whatever movement is made by the shaft, the unknown functions can be expanded into orthogonal series in the modal bases concerning flexion and torsion,

$$y(x,t) = \sum_{k=1}^{\infty} T_k^y(t)V_k(x)$$

$$z(x,t) = \sum_{k=1}^{\infty} T_k^z(t)V_k(x)$$

$$\phi(x,t) = \sum_{m=0}^{\infty} T_m(t)\Phi_m(x) \tag{6}$$

the functions V_k and Φ_m being the mode shapes of the flexural and torsional vibrations. Taken to be distinct of the natural frequencies, the required ordinary differential equations can be obtained:

$$J_k^f\ddot{T}_k^y + R_k^f T_k^y = J_k^b\frac{d}{dt}[\dot{\phi}_o\sin(\phi_o+\alpha_k^b)]$$

$$J_k^f\ddot{T}_k^z + R_k^f T_k^z = -J_k^b\frac{d}{dt}[\dot{\phi}_o\cos(\phi_o+\alpha_k^b)]$$

$$J_m\ddot{T}_m + R_m T_m = \int_o^l C\Phi_m dx + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} J_{mij}^p\frac{\partial}{\partial t}(\dot{T}_i^z T_j^y) \tag{7}$$

$$+ \sum_{n=1}^{\infty}\omega_n^{f2} J_{nm}^b[T_n^z\cos(\phi_o+\alpha_{nm}^b) - T_n^y\sin(\phi_o+\alpha_{nm}^b)]$$

The terms J_k^b, α_k^b, J_{nm}^b, α_{nm}^b and J_{mij}^p are defined in the Appendix. The function $\phi_o(t)$ represents the rigid-body mode. Thus the three coupled partial differential equations are replaced by an infinity or ordinary differential equations; those which represent the tranverse motion are uncoupled.

The general start-up problem of rotating shafts and their passing through critical speeds is in many cases of great practical interest. The calculation of the rigid-body mode is immediate if the coupling terms opposite the "useful" terms are neglected. The first two equations, or family of equations, of the above differential system are those for elementary oscillators subjected to sinusoidal actions whose amplitude and frequency vary slowly, if rotation vibrations are neglected. If the rotation torque is constant, the particular solutions of these equations can be written as follows:

$$T_k^y = \hat{T}_k(t)\cos[\phi_o(t) + \alpha_k^b + \bar{\alpha}_k(t)]$$

$$T_k^z = \hat{T}_k(t)\sin[\phi_o(t) + \alpha_k^b + \bar{\alpha}_k(t)] \tag{8}$$

The functions $\hat{T}_k(t)$ and $\bar{\alpha}_k(t)$ vary slowly. The kth modal phase difference increases from zero to π; its value is $\pi/2$ when the excitation frequency of the system is near that of a natural flexural frequency.

It is now possible to evaluate the nonlinear coupling terms which can induce torsional vibrations. As regards the eccentricity coupling term:

$$\begin{aligned} T_n^z\cos(\phi_o+\alpha_{mn}^b) - T_n^y\sin(\phi_o+\alpha_{mn}^b) &= \hat{T}_n(t)[\sin(\phi_o+\alpha_n^b+\bar{\alpha}_n(t))\cos(\phi_o+\alpha_{mn}^b) \\ &\quad - \cos(\phi_o+\alpha_n^b+\bar{\alpha}_n(t))\sin(\phi_o+\alpha_{mn}^b)] \\ &= \hat{T}_n(t)\sin[\alpha_n^b+\bar{\alpha}_n(t)-\alpha_{mn}^b] \end{aligned} \tag{9}$$

Taking into account the nature of the functions $\hat{T}$ and $\bar{\alpha}$, which vary slowly, it is clear that this coupling term does not have an oscillatory component. It is, therefore, not able, in a first approximation, to induce torsional vibrations.

When the machine starts up, the transverse modes predominate one after the other. Then the torsional vibrations induced by the gyroscopic coupling terms are defined, in a first approximation, by the equation:

$$J_m\ddot{T}_m + R_m T_m = -J_{mnn}^p\hat{T}_n^2(t)\dot{\phi}_o^2(t)\sin 2[\phi_o(t) + \alpha_n^b + \bar{\alpha}_n(t)] \tag{10}$$

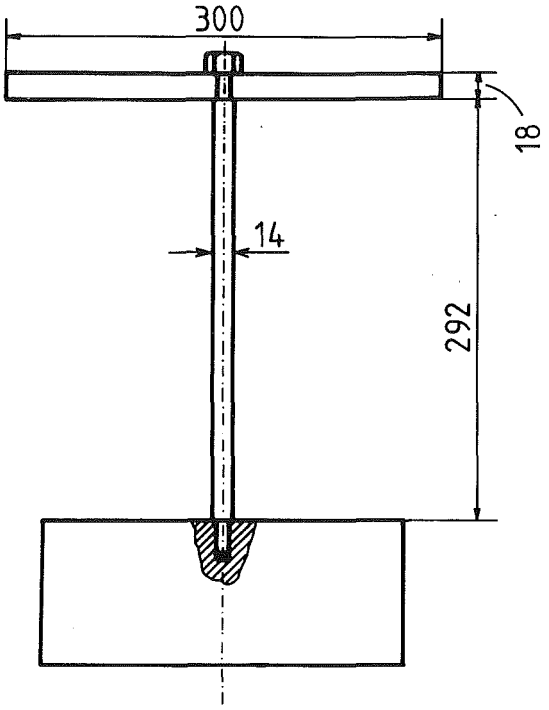


Fig. 4 Characteristics of the vertical experimental device: material: steel; density: 7850 Kg/m^3; elasticity modulus: 210 GPa; shear modulus: 82 GPa; mass of the disk: 10 Kg; longitudinal (polar) mass moment of inertia of the disk: 0.112 Kg m^2 moment of inertia of the support (spindle and driving system): 0.314 Kg m^2. Measured value of the natural frequencies: flexion ($\Omega = 0$): 9.27 Hz; torsion: 17.78 Hz.

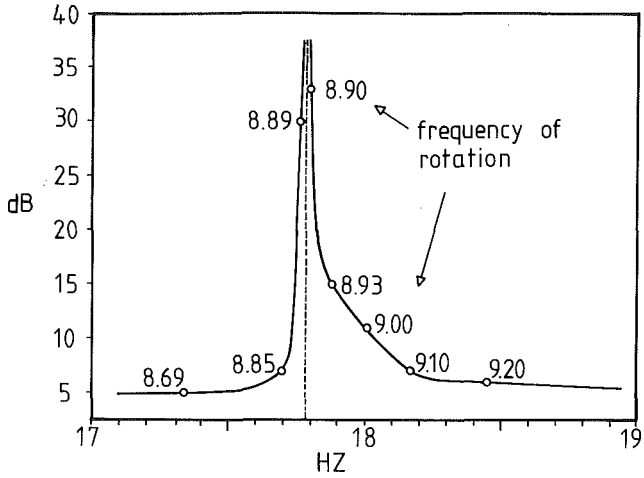


Fig. 5 Spectral representation of the measurement of the bisynchronous component of the torsional vibrations for several steady-state values of the frequency of rotation (parameter)

Without forgetting the modal superposition (6), it can be seen that the excitation amplitude is proportional to the square of the speed of rotation. Finally and most important, the excitation frequency corresponds to twice the speed of rotation. It induces bisynchronous torsional vibrations.

Indeed, when the shaft moves in synchronous precession, the motion of a disk is such that the angles θ and ψ follow sinusoidal laws with an angular frequency equal to the frequency of rotation Ω, and with a phase difference $\pi/2$. With reference to Fig. 3, it is easy to be convinced that $\dot{\psi}$ is "modulated" in a certain way at the frequency Ω. The component $J^p\dot{\psi}\theta$ of the angular momentum, therefore, oscillates with a frequency of twice the rotation frequency; the torsional motion also results.

4 Verification by Experiment

The question is how important are these bisynchronous torsional vibrations in reality? To answer this question, an experimental assembly has been constructed (Fig. 4) in the form of a disk, the longitudinal moment of inertia of which is relatively high. It is situated at one end of a vertical steel shaft. This shaft is imbedded in the front-end of a machine-tool spindle, which is very rigid rotating system with a low perturbation level. The whole rotating system is driven by a direct current electric motor through a highly elastic driving belt. The speed of rotation of the system can be varied from 0 to 600 rpm (0 to 10 Hz).

The shaft is equipped with strain gauges. Torsional vibrations are measured by four strain-gauges with grids set at an angle of ± 45 deg in relation to generating lines. A corresponding full-bridge assembly makes it possible to eliminate all the flexural vibrations. These are obtained using one single strain-gauge in the direction of a generating line. The signals are transmitted to the analyzers by telemetry.

In order to have an appropriate configuration of the system for our purpose, the system has been approximately tuned into the internal resonance. Indeed, the natural frequency of the torsional vibration model (17.8 Hz) is close to the double natural frequency of the lateral vibration mode. The measurement of the natural frequency of the flexural vibration mode when the shaft is not turning gives 9.3 Hz, the critical speed of the rotation calculation corresponds to 10.5 Hz.

The experiment shows that the bisynchronous excitation of the torsional vibrations has a considerable effect only when its frequency is in the region of the natural frequency. In other words, when the speed of rotation is near, or equal to, half the natural frequency of the torsional vibrations, their amplitude increases visibly and spectacularly. Recording the spectrum of the torsional vibrations using an HP-5420 Frequency Analyzer makes it possible to extract the amplitude corresponding to twice the constant speed of rotation. This was done with different values of the rotation frequency and the amplitudes are shown in Fig. 5.

In a similar way, with a very slow rotation sweep speed between 7 and 10 Hz and by memorizing the maxima, we obtain a "frequency response curve" shown in Fig. 6. During the experiment it is easy to see that the peak near the natural frequency (17.78 Hz) appears only when the rotation frequency is near 8.9 Hz. It is not the result of a deterministic or random external excitation because the highly elastic band between the electric motor and the machine-tool spindle is very flexible. The corresponding rotational (rigid-body) natural frequency is less than 1 Hz.

5 Conclusion

Torsional vibrations are considered in the general context of the dynamic behavior of rotating shafts. The kinematic description of the model, an aspect often neglected in the literature, is dealt with carefully. The torsional vibrations described in this paper are caused by transverse vibrations. The coupling is intrinsically related to the dynamic behavior of rotating shafts subjected to the gyroscopic effect. The kinetic energy expression of the system is absolutely classic, except that it includes an eccentricity term, and that the rotation angle is unknown, as are the displacements. The equations of motion of the continuous shaft are obtained using Hamilton's principle. This differential system includes a large number of particular cases found in the literature. It comprises two nonlinear coupling terms in the rotation vibrations equation. The first is due to the residual unbalance and the second gives the variation rate expressed in time of the longitudinal compo-

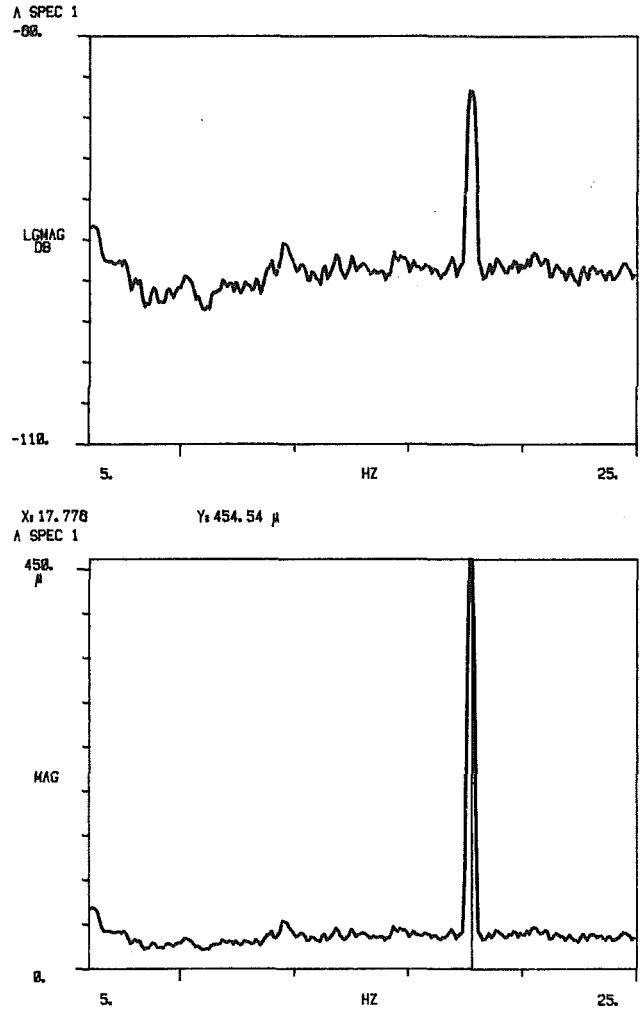


Fig. 6 Spectral representation of the bisynchronous component measurement of the torsional vibrations with a rotation speed sweep from 7 to 10 Hz

nent of the angular momentum resulting from the precession (whirling) of the shaft. If the gyroscopic effect is not taken into account a priori, this last term does not appear.

For the sake of simplicity, only the univocal possibility for transverse vibrations to induce torsional vibrations by coupling is investigated. It is shown that the gyroscopic coupling term plays a more important role in the dynamic behavior of shafts than the eccentricity one. The first can induce torsional vibrations whose frequency is equal to twice the rotation frequency.

Finally, this result is verified experimentally. In order to increase the effect of the second order gyroscopic nonlinear term, the experimental device has been approximately tuned into the internal resonance-natural frequency of the torsional vibration mode is close to the double natural frequency of the lateral vibration mode. In this situation, measurements show that a shaft subjected to the gyroscopic effect can induce extensive torsional vibrations when the rotation frequency is equal to half the natural frequency of the torsional vibrations. It is a resonance phenomenon in which the torsional vibration amplitude depends essentially on the damping of the corresponding mode.

References

Bernasconi, O., 1986, "Solution for Torsional Vibrations of Stepped Shafts Using Singularity Functions," *Int. J. Mech. Sci.*, Vol. 28, No. 1, pp. 31–39.

Bernasconi, O., François, M., 1982, "Etude du Couple Produit par une Roue Pelton à Plusieurs Injecteurs," *Ingénieurs et Architectes Suisses*, Vol. 24, pp. 351–354.

Broniarek, C., 1968, "Investigations of the Coupled Flexural-Torsional Vibration of Rotors with Continuous Parameters," *Zagadninia Drgan Nieliniowych*, No. 9, pp. 169–228.

Cohen, R., Porat, I., 1985, "Coupled Torsional and Transverse Vibration of Unbalanced Rotor," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 701–705.

Gasch, R., Markert, R., Pfützner, H., 1979, "Acceleration of Unbalanced Flexible Rotors through the Critical Speeds," *J. of Sound and Vibration*, Vol. 63, No. 3, pp. 393–409.

Kellenberger, W., 1980, "Résonances Combinées Forcées de l'Arbre Tournant—Couplage de Flexion et de Torsion," *Revue Brown-Bovery*, Vol. 2–80, pp. 117–121.

Lalanne, M., Berthier, P., Der Hagopian, J., 1983, *Mechanical Vibrations for Engineers*, Wiley, New York, pp. 114–118.

Lur'é, L., 1968, *Mécanique Analytique I et II*, Librairie Universitaire, Louvain, Belgique.

Tondl, A., 1965, *Some Problems of Rotor Dynamics*, Chapman and Hall, London.

APPENDIX

The following terms are defined as follows:
Amplitude and phase of the kth modal unbalance:

$$J_k^b = \left[\left(\int_o^l B \cos \alpha V_k dx \right)^2 + \left(\int_o^l B \sin \alpha V_k dx \right)^2 \right]^{1/2}$$

$$\alpha_k^b = \text{Arctg} \frac{\int_o^l B \sin \alpha V_k dx}{\int_o^l B \cos \alpha V_k dx}$$

Amplitude and phase of the nth modal unbalance weighted by the mth torsional mode:

$$J_{nm}^b = \left[\left(\int_o^l B \sin \alpha \Phi_m V_n dx \right)^2 + \left(\int_o^l B \sin \alpha \Phi_m V_n dx \right)^2 \right]^{1/2}$$

$$\alpha_{nm}^b = \text{Arctg} \frac{\int_o^l B \sin \alpha \Phi_m V_n dx}{\int_o^l B \cos \alpha \Phi_m V_n dx}$$

Mth longitudinal modal moment of inertia weighted by the ith and jth flexural modes:

$$J_{mij}^p = \int_o^l J^p \Phi_m V_i' V_j' dx$$

K. Kumar[1]
Visiting Associate Professor.

J. E. Cochran, Jr.
Professor.

Department of Aerospace Engineering,
Auburn University,
Auburn University, AL 36849

Closed-Form Analysis for Elastic Deformations of Multilayered Strands

Closed-form solutions are developed for elastic deformation characteristics of multilayered strands under tensile and torsional loads. These analytical results are successfully applied to obtain expressions for the effective extensional and torsional moduli of rigidity for the strands. Finally, a simple design criterion is established for "nonrotating" cables.

Introduction

The uses of stranded cables in conventional engineering applications are well recognized. For rope-selection and handling, practicing engineers have long depended upon extensive experimental results, such as those compiled by Scoble (1920–1928) that began to appear in the early 1900s. More recently, the Wire Rope Board and federal agencies (1980) have utilized the available empirical data to provide general guidelines for rope selection. Attempts to correlate the experimental data have met with some success (Drucker and Tachau, 1945), yet these fall short of an in-depth understanding of the effects of various wire-rope parameters on their static and dynamic behavior. In view of the recently proposed applications of the wire ropes for augmentation of structural damping in large space structures, the importance of such theoretical static and dynamic investigations has increased considerably.

In many of the earlier theoretical static analyses, simplifying assumptions of questionable validity have been made. For example, Hall (1951) assumed that each wire in a stranded cable is subjected to simple tension. Hruska (1951–1953) criticized this analysis for neglecting friction altogether. He also disagreed with Hall's assertion that the wire core is less stressed than the helical wires wrapped around it. His improved analysis led to expressions for tangential, radial and tensile forces in the wires. Leissa (1959) and Starkey and Cress (1959) were the first to recognize the importance of contact stresses. Bert and Stein (1962) and Machida and Durelli (1973) examined the effects of bending and twisting moments. Chi (1976) looked into the problem of extension of wire ropes with fixed-end strands. However, in his analysis, the effects of the contact forces were neglected.

Costello and Phillips (1973, 1974) adopted a more basic approach to investigate the static behavior of the cables. They treated the cables as groups of separate curved rods (Love, 1944) in the form of helices. Huang (1975, 1978) also made use of this approach in his studies of extension of yarns and cables. Costello and associates (1976–1985) developed the theory of cables further to study the various aspects of static deformation characteristics for cables with complex cross sections. The study was extended to account for the effects of bending moments (Costello and Buston, 1982) as well.

Costello's analysis, although based on conventional linear theory of deformations, involves several nonlinear geometric deformation relations. The complete set of governing equations is solved simultaneously using Newton-Raphson algorithm. That severely limits the utility of the analysis, however, as it depends on computational results obtained for particular cases of cable data. For design, it would be far more desirable to have the solution in an analytical form that can provide an insight into the influence of the various wire rope parameters on its deformation characteristics. Such analytical results may be of even greater significance for the future investigations of wire rope dynamics. In an earlier study, the present authors obtained such approximate analytical results for the simple case of a single strand wire rope with fibrous core (Kumar and Cochran, 1987). A comparison of these closed-form solutions with the corresponding numerical results obtained by Costello and associates established the validity of the approach used in developing those solutions. In this paper, we utilize the same approach to obtain the analytical solutions for the various deformation characteristics of a more complex wire rope model that has a metallic wire core and is wrapped around by a number of layers of helical wires each having its own direction and magnitude of lay.

Analysis

We consider a single strand cable with a straight metallic wire core wrapped around by successive layers of helical wires (Fig. 1). It is assumed that the diameter of the core wire is large enough to prevent the helical wires in a layer from

[1]On leave from I.I.T., Kanpur, India.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 23, 1986; final revision April 7, 1987.

Paper No. 87-WA/APM-19.

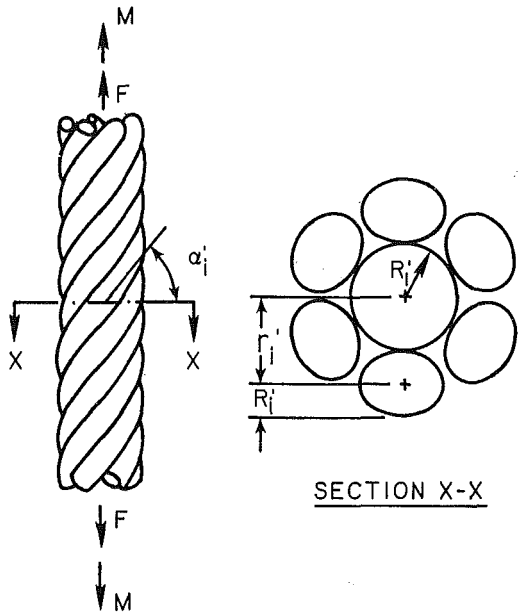


Fig. 1 Rope geometry showing metallic core wrapped around by layers of helical wires

touching each other, although these wires remain in contact with those in the adjacent layers. An axial force, say F, and an axial moment, say M, when applied to the rope cause it to undergo elongation as well as "twist." Here, our objective is to obtain the approximate analytical relations governing the elastic deformations of this single strand cable model.

Let us assume the cable to be made up of n successive layers. Let the ith layer have m_i helical wires, each of radius R_i and a right lay angle α_i. Let the helix radius of the ith layer be denoted by r_i. Since the metallic core is treated as the first layer in the analysis, we have

$$m_1 = 1, \quad \alpha_1 = \pi/2, \quad r_1 = 0.$$

Also from simple geometrical considerations

$$r_2 = R_1 + R_2$$

$$r_3 \simeq r_2 + R_2 + R_3 = R_1 + 2R_2 + R_3$$

and, in general,

$$r_i \simeq R_1 + 2\sum_{j=2}^{i-1} R_j + R_i; \quad i = 2, 3, \ldots, n. \tag{1}$$

Strain Relations. Let ϵ and β denote, respectively, the longitudinal and rotational strains of the wire rope, while the variables ξ_i, β_i are taken to represent the corresponding strains in the helical wires in the ith layer. Then, the analysis of the deformed configuration of the wire-helix in the ith layer yields the following relation between the longitudinal strains ξ_i and ϵ:

$$\epsilon = (1 + \xi_i)(\sin \alpha_i'/\sin \alpha_i) - 1; \quad i = 1, 2, \ldots n. \tag{2}$$

Here, the deformed configuration parameters are distinguished by using primes. This expression is linearized in $\Delta\alpha_i$ to get

$$\xi_i = \epsilon - \Delta\alpha_i \cot \alpha_i; \quad i = 1, 2, \ldots, n \tag{3}$$

where

$$\Delta\alpha_i = \alpha_i' - \alpha_i.$$

Similar analysis for the rotational strain in the helical wires yields

$$\beta_i = (r_i/r_i')/(1 + \epsilon)\cot \alpha_i' - \cot \alpha_i; \quad i = 1, 2, \ldots n. \tag{4}$$

The consideration of changes in the various wire-radii as a result of elongation leads to the following general expression for the deformed helix radii:

$$r_i' \simeq R_1' + 2\sum_{j=2}^{i-1} R_j' + R_i'$$

$$= R_1(1 - \nu\xi_1) + 2\sum_{j=2}^{i-1} \{R_j(1 - \nu\xi_j)\} + R_i(1 - \nu\xi_i);$$

$$i = 2, 3, \ldots, n.$$

where ν stands for the Poisson's ratio of the wire material. Through some simple algebraic manipulation, we now show

$$(r_i'/r_i) \simeq 1 + \delta_i$$

where

$$\delta_i = -\nu\xi_i + 2\nu\sum_{j=2}^{i-1} [(R_j/r_i)\cot \alpha_j \Delta\alpha_j] + \nu(R_i/r_i)\cot \alpha_i \Delta\alpha_i.$$

On substituting for (r_i'/r_i) and δ_i from these relations into equation (4), collecting $\Delta\alpha_i$ terms together and performing considerable algebraic manipulation, we obtain the following first-order result:

$$\Delta\alpha_i = \epsilon(1 + \nu)(1 - \nu_i)\sin \alpha_i \cos \alpha_i \Big[1 - 2\sum_{j=2}^{i-1} \{(r_j/r_i)\nu_j\}\Big] - \beta(r_i/R)(1 - \nu_i)\sin^2\alpha_i \Big[1 - 2\cot \alpha_i \sum_{j=2}^{i-1} \{\nu_j(r_j/r_i)^2 \tan \alpha_j\}\Big] \tag{5}$$

where

$$\nu_i = \nu(R_i/r_i)\cos^2 \alpha_i$$

and

$R =$ the radius of the wire rope.

Force and Moment Relations. This part of the analysis is based on the fundamental theory of wire ropes developed by Costello and Phillips. For completeness, the equations governing the internal forces and moments in helical wires are stated below (Love, 1944; Phillips and Costello, 1973):

$$\begin{aligned}
(dT_{i_n}/ds) &- T_{i_b}\tau_i' + T_{i_a}\lambda_i' + Q_{i_n} &= 0\\
(dT_{i_b}/ds) &- T_{i_a}\kappa_i' + T_{i_n}\tau_i' + Q_{i_b} &= 0\\
(dT_{i_a}/ds) &- T_{i_n}\lambda_i' + T_{i_b}\kappa_i' + Q_{i_a} &= 0\\
(dG_{i_n}/ds) &- G_{i_b}\tau_i' + G_{i_a}\lambda_i' - T_{i_b} + N_{i_n} &= 0\\
(dG_{i_b}/ds) &- G_{i_a}\kappa_i' + G_{i_n}\tau_i' + T_{i_n} + N_{i_b} &= 0\\
(dG_{i_a}/ds) &- G_{i_n}\lambda_i' + G_{i_b}\kappa_i' + N_{i_a} &= 0
\end{aligned}$$

where

T_i, G_i = the resultant internal and external forces and moments, respectively, in helical wires in the ith layer.

Q_i, N_i = the resultant contact forces and moments per unit length, respectively.

V_n, V_b, V_a = the components of the vector V along the wire normal, binormal and axial directions, respectively; $V = T_i, G_i, Q_i$ and N_i.

G_{i_n}, G_{i_b}, $G_{i_a} = (\pi R_i^4 E/4)\ (\kappa_i' - \kappa_i)$, $(\pi R_i^4 E/4)(\lambda_i' - \lambda_i)$ and $[\pi R_i^4 E/\{4(1 + \nu)\}](\tau_i' - \tau_i)$, respectively.

$$T_{i_a} = \pi R_i^2 E \xi_i.$$

E = modulus of rigidity of wire material.

κ_i, κ_i' = the components of initial and final curvatures of the centerline of the helical wire in normal direction, respectively; both are zero.

λ_i, λ_i' = the components of initial and final curvatures of the centerline of the helical wire along the binormal direction, which are $(\cos^2 \alpha_i/r_i)$ and $(\cos^2 \alpha_i'/r_i')$, respectively.

τ_i, τ_i' = $(\sin \alpha_i \cos \alpha_i/r_i)$, $(\sin \alpha_i' \cos \alpha_i'/r_i')$; the initial and final angles of twist per unit length, respectively.

α_i, r_i = helix angle and helix radius of the wire in the ith layer.

s = distance along the centerline of the deformed wire.

We now assume deformations to be small, tension along the length of the cable to be constant, and bending moments introduced through line-contacts to be zero; i.e., $N_{i_n} = N_{i_b} = 0$. Then the governing force and moment equations may be reduced to:

$$Q_{i_n} = T_{i_a}(\cos^2\alpha_i'/r_i') - T_{i_b}(\sin \alpha_i' \cos \alpha_i'/r_i')$$
$$T_{i_b} = (G_{i_a}(\cos^2\alpha_i'/r_i') - G_{i_b}(\sin\alpha_i'\cos \alpha_i'/r_i')$$
$$Q_{i_b} = Q_{i_a} = N_{i_a} = 0.$$

Now, ignoring the pairs of nonzero but equal and opposite contact forces denoted by Q_i (as these cancel each other and hence do not affect the overall rope equilibrium), the total axial force and moment on the helical wire can be written as

$$F_i = (T_{i_a} \sin \alpha_i' + T_{i_b} \cos \alpha_i')$$
$$M_i = (G_{i_a} \sin \alpha_i' + G_{i_b} \cos \alpha_i' + T_{i_a} r_i' \cos \alpha_i' - T_{i_b} r_i' \sin \alpha_i').$$

Incorporating the first-order approximations, as before, in conjunction with the approximate analytical results obtained above, we find that:

$$T_{i_a} \simeq \pi R_i^2 E(\epsilon - \Delta\alpha_i \cot \alpha_i)$$
$$G_{i_a} \simeq [\pi R_i^4 E/\{4(1+\nu)r_i\}][\Delta\alpha_i \cos 2\alpha_i - \delta_i \sin \alpha_i \cos \alpha_i]$$
$$G_{i_b} \simeq [\pi R_i^4 E/\{4r_i\}][-\Delta\alpha_i \sin 2\alpha_i - \delta_i \cos^2 \alpha_i]$$
$$T_{i_b} \simeq [\pi R_i^4 E/\{4r_i^2\}][\Delta\alpha_i \cos^2 \alpha_i(1 - \nu_f \cos 2\alpha_i) + \nu_f \delta_i \sin \alpha_i \cos^3 \alpha_i]$$

where

$$\nu_f = \nu/(1+\nu)$$

On substituting these in the expression for F_i, summing the forces F_i over all the wires and performing suitable algebraic manipulations, we obtain the following dimensionless force deformation relation:

$$\bar{F} = \epsilon \sum_i [m_i R_i^2 \sin \alpha_i \{1 - (1+\nu)p_i \cos^2\alpha_i\}]/\sum_i [m_i R_i^2]$$
$$+ \beta \sum_i [m_i R_i^2 p_i (r_i/R) \sin^2\alpha_i \cos \alpha_i \{1 - \nu \cot \alpha_i \sum_{j=2}^{i-1} [R_j/r_i)^2 \sin 2\alpha_j]\}]/\sum_i [m_i R_i^2]$$

where

$$\bar{F} = (F/AE)$$
$$p_i = (1-\nu_i)[1 - (1/4)(R_i/r_i)^2(1 - \nu_f \cos 2\alpha_i)\cos^2 \alpha_i].$$

Or simply,

$$\bar{F} = F_\epsilon \epsilon + F_\beta \beta \tag{6}$$

where

$$F_\epsilon = \sum_i [m_i R_i^2 \sin \alpha_i \{1 - (1+\nu)p_i \cos^2 \alpha_i\}]/\sum_i [m_i R_i^2] \tag{7}$$

$$F_\beta = \sum_i \Big[m_i R_i^2 p_i (r_i/R) \sin^2 \alpha_i \cos \alpha_i \Big\{1 - \nu \cot \alpha_i \sum_{j=2}^{i-1} [(R_j/r_i)^2 \sin 2\alpha_j] \Big\}\Big]/\sum_i m_i R_i^2]. \tag{8}$$

The terms F_ϵ, F_β represent the dimensionless longitudinal and rotational strain derivatives of the axial force.

Similarly, on substituting the expressions for the various terms obtained earlier and making suitable simplifying approximations, the expressions for the moments can be written as

$$4M_i r_i/(\pi R_i^4 E) = \epsilon \cos \alpha_i \Big[4(r_i/R_i)^2 - 1 + \nu_f \sin^4 \alpha_i - (1+\nu)(q_i - 1 + \nu_f \sin^4 \alpha_i)\Big\{1 - 2\sum_{j=2}^{i-1} (\nu_j r_j/r_i)\Big\}\Big]$$
$$+ \beta (r_i/R) \sin \alpha_i \Big[q_i \Big\{1 - 2\sum_{j=2}^{i-1} [\nu_j (r_j/r_i)^2 \tan \alpha_j \cot \alpha_i]\Big\} + 2(1 - \nu_f \sin^4 \alpha_i) \sum_{j=2}^{i-1} \{\nu_j (r_j/r_i)^2 (\sin \alpha_j/\sin \alpha_i)\}\Big]$$

where

$$q_i = (1-\nu_i)[4(r_i/R_i)^2 \cos^2 \alpha_i + 1 - \cos^4 \alpha_i + \nu_f \sin^4 \alpha_i \cos^2 \alpha_i].$$

By adding up the contribution of the internal moments of all the wires in the rope and dividing the total moment M by ER^3, the expression for the corresponding dimensionless moment, $\bar{M} = M/ER^3$, can be written as

$$\bar{M} = M_\epsilon \epsilon + M_\beta \beta \tag{9}$$

where

$$M_\epsilon = (\pi/4) \sum_i \Big[m_i \bar{R}_i^3 (R_i/r_i) \cos \alpha_i \Big[4(r_i/R_i)^2 - 1 + \nu_f \sin^4 \alpha_i - (1+\nu)(q_i - 1 + \nu_f \sin^4 \alpha_i) \Big\{1 - 2\sum_{j=2}^{i-1} [\nu_j (r_j/r_i)]\Big\}\Big]\Big] \tag{10}$$

$$M_\beta = (\pi/4) \sum_i m_i \bar{R}_i^4 \sin \alpha_i \Big[q_i \Big\{1 - 2\sum_{j=2}^{i-1} [\nu_j (r_j/r_i)^2 \tan \alpha_j \cot \alpha_i]\Big\} + 2(1 - \nu_f \sin^4 \alpha_i) \sum_{j=2}^{i-1} [\nu_j (r_j/r_i)^2 (\sin \alpha_j/\sin \alpha_i)]\Big] \tag{11}$$

and

$$\bar{R}_i = R_i/R.$$

To achieve a clearer insight into the effect of various wire rope parameters on the force and moment derivatives, namely, F_ϵ, F_β, M_ϵ and M_β, the above expressions were simplified using suitable approximations. For practical values of ν and α_j, the terms involving $\nu_j\ (r_j/r_i)$ and $\nu_j\ (r_j/r_i)^2$, usually ~ 0.01, can be neglected in comparison to unity. Incorporating this approximation reduces the above analytical results to:

$$F_\epsilon = \sum_i [m_i R_i^2 \sin\alpha_i(\sin^2\alpha_i - \nu\cos^2\alpha_i)]/\sum_i [m_i R_i^2] \tag{12}$$

$$F_\beta = \sum_i [m_i R_i^2 (r_i/R)\sin^2\alpha_i\cos\alpha_i]/\sum_i [m_i R_i^2] \tag{13}$$

$$M_\epsilon = \pi\sum_i [m_i \bar{R}_i^3 \cos\alpha_i\{(r_i/R_i)(\sin^2\alpha_i - \nu\cos^2\alpha_i) - (1/4)(R_i/r_i)\}] \tag{14}$$

$$M_\beta = \pi\sum_i [m_i \bar{R}_i^4 \sin\alpha_i\{(r_i/R_i)^2\cos^2\alpha_i + (1/4)(1+\nu_f\sin^4\alpha_i\cos 2\alpha_i)\}]. \tag{15}$$

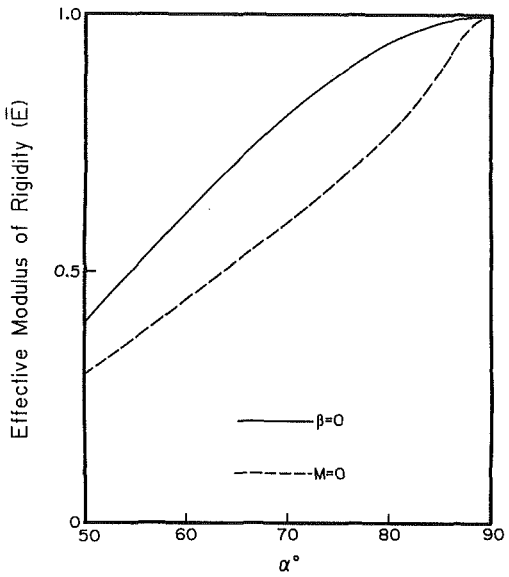


Fig. 2 Modulus of ridigity as affected by the helix angle

Results and Discussion

The analytical results obtained here provide considerable insight into the static behavior of strands. The results are rather general and apply to any arbitrary end-conditions. The analysis fully covers the strands with an arbitrary number of layers having right as well as left lay angles. However, for each of the layers with left lay, its helix angle α_i must be replaced by the corresponding obtuse angle $(\pi - \alpha_i)$ in the above expressions. The method of computation is illustrated through an example. Considered here is a three-layered cable with the following data:

$R_1 = R_2 = R_3 = a$;

$m_1 = 1, \quad m_2 = 6, \quad m_3 = 12$;

$\alpha_1 = \pi/2, \quad \alpha_2 = \alpha$(right lay); $\alpha_3 = \pi - \alpha$ (left lay).

Strictly speaking, R_2 has to be slightly less than R_1 in order to preserve the geometry of the multilayered strand under consideration; however, for practical computational purposes, the radius R_2 has been taken to be the same as R_1. On substituting these data in equations (12)–(15), we obtain the following results:

$$F_\epsilon = (1/19)[1 + 18\sin\alpha\{1 - (1+\nu)\cos^2\alpha\}]$$

$$F_\beta = -(36/95)\sin^2\alpha\cos\alpha$$

$$M_\epsilon = -(36\pi/125)(\sin^2\alpha - \nu\cos^2\alpha)\cos\alpha$$

$$M_\beta = (216\pi/625)[\sin\alpha\cos^2\alpha + (1/48)\sin\alpha(1+\nu_f\sin^4\alpha\cos 2\alpha) + (1/864)/(1+\nu)].$$

These analytical results are of considerable significance. We can directly make use of these expressions to study the rotational and extensional strains and hence the relative influence of the various cable parameters such as the number, size, and lay angle of helical wires in each layer on the cable deformation characteristics.

Next, we demonstrate the usefulness of the analytical results by applying these to study some important extensional and torsional stiffness characteristics of the wire ropes.

Applications

Effective Modulus of Rigidity. Two commonly encountered conditions are investigated separately. In the first one, the ends are constrained against rotation while in the second, the cable ends are free to rotate.

Case 1. No end rotation, i.e., $\beta = 0$. In this practical situation, the equations (6) and (9) reduce to:

$$\bar{F} = F_\epsilon\epsilon \tag{16}$$

$$\bar{M} = M_\epsilon\epsilon. \tag{17}$$

Equation (16) directly yields the expression for the effective modulus of rigidity ($\bar{E}$):

$$(\bar{E})_{\beta=0} = F_\epsilon = \sum_i [m_i R_i^2 \sin\alpha_i(\sin^2\alpha_i - \nu\cos^2\alpha_i)]/\sum_i [m_i R_i^2]. \tag{18}$$

Equation (17) provides the axial moment required to prevent any twist when the rope undergoes the axial strain ϵ.

Case 2. Cable ends free to rotate; i.e., $M = 0$. In several applications, the cables are not restrained against rotation. For this end condition, the axial moment can be assumed to be zero. Hence, to study this case, we substitute $M = 0$ in equation (9) which yields

$$\beta = -(M_\epsilon/M_\beta)\epsilon.$$

This in conjunction with equation (6) leads to the following expression for the extensional modulus of rigidity:

$$(\bar{E})_{M=0} = F_\epsilon - (F_\beta M_\epsilon/M_\beta).$$

On substituting the expressions for F_ϵ, F_β, M_ϵ, and M_β into the above two relations, we get:

$$\beta = -\epsilon\sum_i [m_i \bar{R}_i^3\cos\alpha_i\{(r_i/R_i)(\sin^2\alpha_i - \nu\cos^2\alpha_i) - (1/4)(R_i/r_i)\}] / \sum_i [m_i \bar{R}_i^4 \sin\alpha_i\{(r_i/R_i)^2\cos^2\alpha_i + (1/4)(1+\nu_f\sin^4\alpha_i\cos 2\alpha_i)\}]$$

and

$$(\bar{E})_{M=0} = \sum_i [m_i R_i^2 \sin\alpha_i (\sin^2\alpha_i - \nu\cos^2\alpha_i)] / \sum_i (m_i R_i^2]$$

$$- \sum_i [m_i \bar{R}_i^2 (r_i/R) \sin^2\alpha_i \cos\alpha_i]$$

$$\sum_i [m_i \bar{R}_i^3 \cos\alpha_i \{ (r_i/R_i)(\sin^2\alpha_i - \nu\cos^2\alpha_i) - (1/4)(R_i/r_i) \}] / \Big[\sum_i [m_i \bar{R}_i^2] .$$

$$\sum_i [m_i \bar{R}_i^4 \sin\alpha_i \{ (r_i/R_i)^2 \cos^2\alpha_i + (1/4)(1+\nu_f \sin^4\alpha_i \cos 2\alpha_i) \}] \Big] . \quad (19)$$

An application of these equations is illustrated through the earlier example of the three-layered cable for which we obtain the following results:

$$(\bar{E})_{\beta=0} \simeq (1/19)[1 + 18\sin\alpha(\sin^2\alpha - \nu\cos^2\alpha)]$$

$$(\bar{E})_{M=0} \simeq (1/19)[1 + \sin\alpha(\sin^2\alpha - \nu\cos^2\alpha) \{ 12\sin\alpha\cos^2\alpha + (3/8)\sin\alpha(1+\nu_f \sin^4\alpha\cos 2\alpha) + (1/48)/(1+\nu) \} / \{ \sin\alpha\cos^2\alpha + (1/48)\sin\alpha(1 + \nu_f \sin^4\alpha\cos 2\alpha) + (1/864)/(1+\nu) \}] .$$

The results of analysis for the case under consideration are presented in Fig. 2. It is evident that, for both the end conditions, the modulus of rigidity increases with α. Furthermore, the modulus of rigidity is smaller for the rope under free end conditions.

Effective Torsional Rigidity of Cables. On substituting for ϵ from equation (6) into (9), we get

$$\bar{M} = (M_\epsilon/F_\epsilon)\bar{F} + (M_\beta - M_\epsilon F_\beta/F_\epsilon)\beta .$$

This equation shows that the plot of dimensionless moment $\bar{M}$ versus rotational strain β for any specified tensile force represented by $\bar{F}$ is a straight line as expected. The common slope of this family of straight lines defines the torsional stiffness. The expression for the torsional rigidity in the dimensionless form can be written as

$$\bar{G} = M_\beta - M_\epsilon F_\beta / F_\epsilon .$$

On substituting the expressions for F_ϵ, F_β, M_ϵ, and M_β, we get

$$\bar{G} = \pi \Big[\sum_i m_i \bar{R}_i^4 \sin\alpha_i \{ (r_i/R_i)^2 \cos^2\alpha_i + (1/4)(1+\nu_f \sin^4\alpha_i \cos 2\alpha_i) \} - \sum_i [m_i \bar{R}_i^3 \cos\alpha_i \{ (r_i/R_i)(\sin^2\alpha_i - \nu\cos^2\alpha_i) - (1/4)(R_i/r_i) \}] \sum_i [m_i \bar{R}_i^3 (r_i/R_i) \sin^2\alpha_i \cos\alpha_i] / \sum_i [m_i \bar{R}_i^2 \sin\alpha_i (\sin^2\alpha_i - \nu\cos^2\alpha_i)] \Big] . \quad (21)$$

For the cable data assumed earlier, it is relatively easy to show that

$$\bar{G} = (216\pi/625)[\sin\alpha\cos^2\alpha \{ 1 + 12\sin\alpha(\sin^2\alpha - \nu\cos^2\alpha) \} / \{ 1 + 18\sin\alpha(\sin^2\alpha - \nu\cos^2\alpha) \}$$

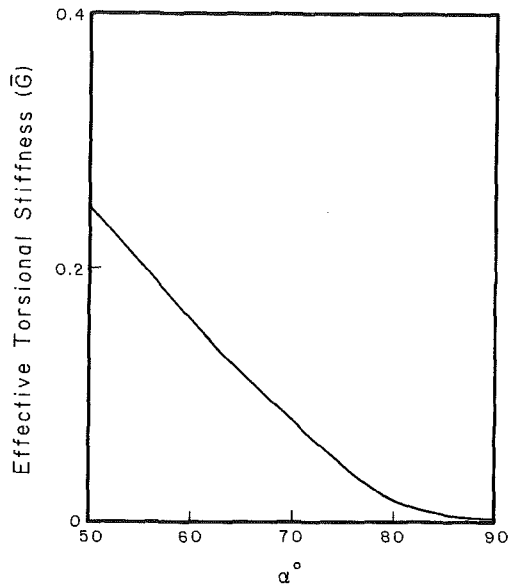


Fig. 3 Torsional stiffness as affected by the helix angle

$$+ (1/48)\sin\alpha(1+\nu_f \sin^4\alpha\cos 2\alpha) + (1/864)/(1+\nu)] . \quad (22)$$

Figure 3 illustrates the effect of α on the torsional stiffness. For the practical values of α in the range 50 deg $\leq \alpha \leq$ 90 deg, the torsional stiffness continuously decreases as α approaches 90 deg. In the limit, α = 90 deg, the dimensionless parameter $\bar{G}$ becomes rather small attaining a value of 0.0018. These analytical results can be utilized readily for generating design data for any general single strand cable configuration with metallic wire core.

Nonrotating Rope. When the rope is loaded by a tensile force, in general, its extension is accompanied by rotation. A suitable axial moment is therefore required if the rope under tension is to be held in the "unrotated" configuration. This moment must balance the sum of the end moments contributed by the inside and outside layers. If the inside and the outside layers have the opposite lays, their individual twisting moment contributions will be of opposite signs and neutralize each other under certain specific conditions. Such a rope where the pure tensile force causes no end rotation is called "nonrotating." It is easy to see that the criterion for "nonrotating" rope can be stated simply by

$$M_\epsilon = 0 .$$

Referring to equation (16), we have

$$\sum_i [m_i \bar{R}_i^3 \cos\alpha_i \{ (r_i/R_i)(\sin^2\alpha_i - \nu\cos^2\alpha_i) - (1/4)(R_i/r_i) \}] = 0 . \quad (23)$$

By applying this to a common particular case of a cable having a central metallic core with two layers – an inner and outer – we get

$$\cos\alpha_3 = -(m_2/m_3)(\bar{R}_2/\bar{R}_3)^3 \cos\alpha_2 [(r_2/R_2)(\sin^2\alpha_2 - \nu\cos^2\alpha_2) - (1/4)(R_2/r_2)]/[(r_3/R_3)(\sin^2\alpha_3 - \nu\cos^2\alpha_3) - (1/4)(R_3/r_3)] . \quad (24)$$

To first order in α_3, this expression can be approximated by

$$\cos\alpha_3 \simeq -(m_2/m_3)(R_2/R_3)^2 (r_2/r_3)\cos\alpha_2 .$$

On substituting this into the right side of the more accurate expression for $\cos\alpha_3$, we get

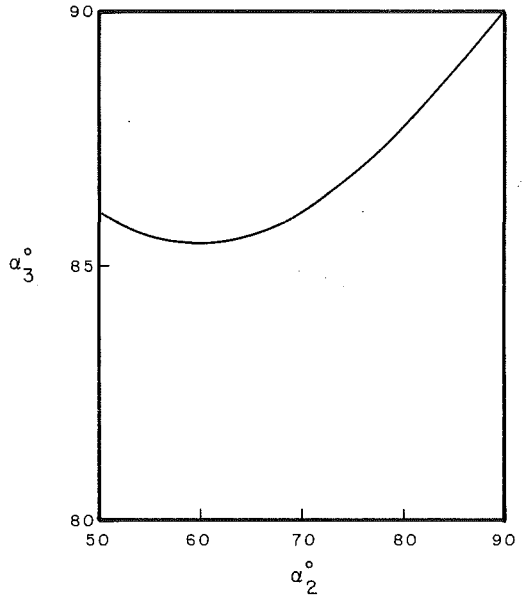


Fig. 4 The choice of α_3 as affected by α_2 for the "nonrotating" cable

$$\cos\alpha_3 \simeq -(m_2/m_3)(R_2/R_3)^2(r_2/r_3)\cos\alpha_2[1 - (1+\nu)\cos^2\alpha_2 - (1/4)(R_2/r_2)^2]/[1 - (1+\nu)(m_2/m_3)^2(R_2/R_3)^4(r_2/r_3)^2\cos^2\alpha_2 - (1/4)(R_3/r_3)^2]. \quad (25)$$

Now, we apply this result to a more specific, three-layered cable with the following assumed data:

$$R_1 = R_2 = R_3 = a;$$
$$m_1 = 1, \quad m_2 = 6, \quad m_3 = 12;$$
$$\alpha_1 = \pi/2, \quad \alpha_2 = \alpha_2.$$

By substituting these data in equation (24), the expression for α_3 can be written as

$$\cos\alpha_3 = -\cos\alpha_2[15 - 16(1+\nu)\cos^2\alpha_2]/[63 - 4(1+\nu)\cos^2\alpha_2].$$

Further iterations were unnecessary as the residual errors involved here would be much smaller than those introduced through the earlier approximations.

For a better appreciation of the analytical results obtained here, a plot of α_3 versus α_2 for the "nonrotating" rope is presented in Fig. 4. The plot shows how the choice of α_2 changes that of α_3 for the "nonrotating" rope. It is evident that as α_2 approaches $\pi/2$, so does α_3. Furthermore, in general, α_3 would be much closer to $\pi/2$ than α_2. The results of this investigation should prove to be quite useful in designing a "nonrotating" cable.

Conclusion

The dimensionless analytical force and torsional moment relations developed here for multilayered cable with metallic wire core are rather general. The explicit form of results is found to be useful in predicting the rope stiffness against elongation as well as rotation. The effects of the layout of layers, number of wires in each layer, and of course the direction and magnitude of lay angles on these important deformation characteristics are clearly brought out. The simple form of the analytical results is found to be equally useful in the analysis of "nonrotating" cables. Finally, the examples considered here demonstrate the computational ease and effectiveness with which the closed-form solution can be utilized in various studies.

Acknowledgment

The authors wish to acknowledge the support of this work by NASA (under Grant NAG8-532 through Marshall Space Flight Center, Huntsville, Alabama) and the Auburn University Engineering Experiment Station.

References

Bert, C. W., and Stein, R. A., 1962, "Stress Analysis of Wire Rope in Tension and Torsion," *Wire and Wire Products*, Vol. 37, pp. 769–770, pp. 772–816.

Chi, M., 1974, "Analysis of Multi-Wire Strands in Tension and Combined Tension and Torsion," *Developments in Theoretical and Applied Mechanics, Proceedings of the Seventh Southeastern Conference on Theoretical and Applied Mechanics*, Vol. 7, pp. 599–639.

Code of Federal Regulations, 1980, Vol. 30, Mineral Resources, pp. 357.19–39.

Costello, G. A., and Phillips, J. W., 1973, "Contact Stresses in Thin Twisted Rods," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, pp. 629–630.

Costello, G. A., and Phillips, J. W., 1974, "A More Exact Theory for Twisted Wire Cables," *ASCE Journal of the Engineering Mechanics Division*, Vol. 100, pp. 1096–1099.

Costello, G. A., and Phillips, J. W., 1976, "Effective Modulus of Twisted Wire Cables," *ASCE Journal of the Engineering Mechanics Division*, Vol. 102, pp. 171–181.

Costello, G. A., and Sinha, S. K., 1977a, "Torsional Stiffness of Twisted Wire Cables," *ASCE Journal of the Engineering Mechanics Division*, Vol. 103, pp. 766–770.

Costello, G. A., and Sinha, S. K., 1977b, "Static Behavior of Wire Rope," *ASCE Journal of the Engineering Mechanics Division*, Vol. 103, pp. 1011–1022.

Costello, G. A., 1978, "Analytical Investigation of the Wire Rope," *Applied Mechanics Review*, Vol. 31, pp. 897–900.

Costello, G. A., and Miller, R. E., 1979, "Lay Effect of Wire Rope," *ASCE Journal of the Engineering Mechanics Division*, Vol. 105, pp. 597–608.

Costello, G. A., and Miller, R. E., 1980, "Static Response of Reduced Rotation Rope," *ASCE Journal of the Engineering Mechanics Division*, Vol. 106, pp. 623–631.

Costello, G. A., and Buston, G. J., 1982, "Simplified Bending Theory for Wire-Rope," *ASCE Journal of the Engineering Mechanics Division*, Vol. 108, pp. pp. 219–227.

Costello, G. A., 1983, "Stresses in Multilayered Cables," *Journal of Energy Resources Technology*, Vol. 105, pp. 337–340.

Drucker, D. C., and Tachau, H., 1945, "A New Design Criterion for Wire Rope," ASME JOURNAL OF APPLIED MECHANICS, Vol. 67, pp. A-33–A-38.

Hall, H. M., 1951, "Stresses in Small Wire Ropes," *Wire and Wire Products*, Vol. 26, p. 228, pp. 257–259.

Hruska, F. H., 1951, "Calculations of Stresses in Wire Ropes," *Wire and Wire Products*, Vol. 26, pp. 766–767, pp. 799–801.

Hruska, F. H., 1952, "Radial Forces in Wire Ropes," *Wire and Wire Products*, Vol. 27, pp. 459–463.

Hruska, F. H., 1953, "Tangential Forces in Wire Ropes," *Wire and Wire Products*, Vol. 28, pp. 455–460.

Huang, N. C., 1975, "On the Extension of an Elastic Strand with a Central Core," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 821–824.

Huang, N. C., 1978, "Finite Extension of an Elastic Strand with a Central Core," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 852–857.

Kumar, K., and Cochran, J. E., Jr., 1987, "Analytical Solutions for Static Elastic Deformations of Wire Ropes," AIAA Paper No. 87-0720, *Proceedings of the 28th AIAA/ASME/ASCE/AHS Structures, Structural Dynamics and Materials Conference*, Part 1, April 6–8, Monterey, CA, pp. 88–92.

Leissa, A. W., 1959, "Contact Stresses in Wire Ropes," *Wire and Wire Products*, Vol. 34, pp. 307–316, pp. 372–373.

Love, A. E. H., 1944, *A Treatise on the Theory of Elasticity*, Dover Publications, New York, pp. 251–276.

Machida, S., and Durelli, A. J., 1973, "Response of a Strand to Axial and Torsional Displacements," *Journal of Mechanical Engineering Science*, Vol. 15, pp. 241–251.

Phillips, J. W., and Costello, G. A., 1973, "Contact Stresses in Twisted Wire Cables," *ASCE Journal of the Engineering Mechanics Division*, Vol. 99, pp. 331–341.

Phillips, J. W., and Costello, G. A., 1985, "Analysis of Wire Ropes with Internal-Wire-Rope Cores," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 510–516.

Scoble, W. A., 1920–35, "First Report of the Wire Rope Research Committee," *Proc. Institution of Mechanical Engineers*, Vol. 115, pp. 835–868; Vol. 119, pp. 1193–1290; Third Report, Vol. 123, pp. 353–404; Fourth Report, Vol. 130, pp. 373–478.

Starkey, W. L., and Cress, H. A., 1959, "An Analysis of Critical Stresses and Mode of Failure of a Wire Rope," ASME *Journal of Engineering for Industry*, Vol. 81, pp. 307–316.

Velinsky, S. A., Anderson, G. L., and Costello, G. A., 1984, "Wire Rope with Complex Cross Sections," *ASCE Journal of the Engineering Mechanics Division*, Vol. 110, pp. 380–391.

Wire Rope Users Manual, 1979, Washington, D. C., American Iron and Steel Institute.

A Method to Improve the Modal Convergence for Structures With External Forcing[1]

Keqin Gu
Graduate Student,
Student Mem. ASME

Benson H. Tongue
Assistant Professor,
Assoc. Mem. ASME

School of Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332

The traditional approach of using free vibration modes in the assumed mode method often leads to an extremely slow convergence rate, especially when discete interactive forces are involved. By introducing a number of forced modes, significant improvements can be achieved. These forced modes are intrinsic to the structure and the spatial distribution of forces. The motion of the structure can be described exactly by these forced modes and a few free vibration modes provided that certain conditions are satisfied. The forced modes can be viewed as an extension of static modes. The development of a forced mode formulation is outlined and a numerical example is presented.

1 Introduction

A great deal of work is currently underway with regard to the analysis and control of flexible structures. Such work often involves discrete interactive forces between flexible components and/or discrete constraints of flexible components to a base. Two examples are flexible robot arms and serially connected flexible space structures (Book, 1984; Baker, 1984; Nurre et al., 1984). The motions of such structures are often expressed in terms of some assumed modes of their individual flexible components. Even though the configuration space of a flexible structure is infinite dimensional, numerical considerations dictate that only a finite number of assumed modes be used to express the flexible motion. The errors of such an approximation can lead to an inaccurate estimation of the actual motion; indeed with regard to flexible structure control, such errors lead to "spill-over," which can deteriorate the stabilitiy of the entire system. A question therefore arises as to the assumed mode set that produces minimal errors with the least computational effort. This demand is especially keen when we are confronted with a real time control problem, in which both stability and speed are crucial.

Free vibration modes are the most frequently used assumed modes. A relatively simple analytical formulation can be achieved in most cases, due to the orthogonality of the modes. However, using only free vibration modes can often lead to intolerably slow convergence, as will be shown. This is especially true when the discrete forces and/or moments become large.

The static modes of a system, which are proportional to the displacements due to the external forces, are often used (Craig, 1985) as a means of accelerating the computational convergence rate. Another technique to increase the convergence speed is the modal acceleration method (Bisplinghoff and Ashley, 1962). The forced modes to be discussed in this paper can be considered as an extension of the above mentioned methods.

2 Forced Modes

Consider a uniform bar fixed at one end and subject to a tension at the other end, as shown in Fig. 1. The equation of motion is

$$\frac{\partial^2}{\partial x^2} u(x,t) = a^2 \frac{\partial^2}{\partial t^2} u(x,t) \tag{1}$$

with boundary conditions

$$u(0,t) = 0, \quad \frac{\partial}{\partial x} u(l,t) = f(t) \tag{2a}$$

where

$$a^2 = \frac{m}{EA}, \quad f(t) = \frac{F(t)}{EA}$$

and the initial conditions are

$$u(x,0) = \phi(x) \tag{2b}$$

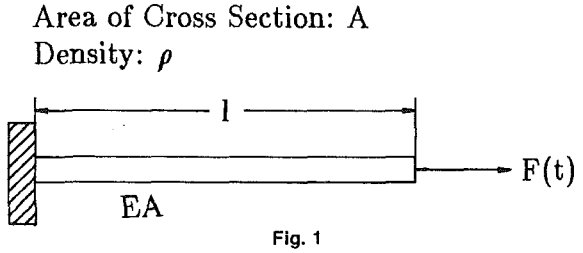


Fig. 1

[1]This work was supported by the National Science Foundation, Grant No. MSM-8451186.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February 23, 1987; final revision May 27, 1987.

Paper No. 88-WA/APM-1.

The following transformation is utilized:

$$u(x,t) = u_1(x,t) + f(t)\,Y_1(x) \tag{3}$$

where

$$Y_1(x) = x \tag{4}$$

in this case. By so doing, we are using the modal acceleration method (Bisplinghoff, 1962). Equations (1) and (2) are transformed to

$$\frac{\partial^2}{\partial x^2} u_1(x,t) = a^2 \frac{\partial^2}{\partial t^2} u_1(x,t) + a^2 Y_1(x) \frac{d^2 f(t)}{dt^2} \tag{5}$$

$$u_1(0,t) = 0, \quad \frac{\partial}{\partial x} u_1(l,t) = 0 \tag{6a}$$

$$u_1(x,0) = \phi(x) - f(0)\,Y_1(x) \tag{6b}$$

The system has thus been converted into a distributed force system. Since the function $Y_1(x) = x$ is more readily expressed by a linear combination of free vibration modes, a modal decomposition of the solution of the system involves fewer modes of high mode number. Notice also that the free vibration modes are *comparison functions* now instead of general *admissible functions* (Meirovitch, 1975).

The preceding procedure can also be carried out on equations (5)–(6) to further accelerate the convergence rate. Let

$$u_1(x,t) = u_2(x,t) + a^2 \frac{d^2 f(t)}{dt^2} Y_2(x)$$

where

$$Y_2(x) = \frac{x^3}{6} - \frac{l^2 x}{2}$$

The equation for u_2 is thus transformed to

$$\frac{\partial^2}{\partial x^2} u_2(x,t) = a^2 \frac{\partial^2}{\partial t^2} u_2(x,t) + a^4 \frac{d^4 f(t)}{dt^4} Y_2(x)$$

with

$$u_2(0,t) = 0, \quad \frac{\partial}{\partial x} u_2(l,t) = 0$$

$$u_2(x,0) = \phi(x) - f(0)\,Y_1(x) - a^2 \frac{d^2 f(0)}{dt^2} Y_2(x)$$

This process can be continued to obtain

$$u(x,t) = u_n(x,t) + f(t)\,Y_1(x) + a^2 \frac{d^2 f(t)}{dt^2} Y_2(x)$$

$$+ \cdots + a^{2n-2} \frac{d^{2n-2} f(t)}{dt^{2n-2}} Y_n(x) \tag{7}$$

where u_n satisfies

$$\frac{\partial^2}{\partial x^2} u_n(x,t) = a^2 \frac{\partial^2}{\partial t^2} u_n(x,t) + a^{2n} \frac{d^{2n} f(t)}{dt^{2n}} Y_n(x) \tag{8}$$

$$u_n(0,t) = 0, \quad \frac{\partial}{\partial x} u_n(l,t) = 0 \tag{9}$$

and the iteration formula for Y_k is

$$\frac{\partial^2}{\partial x^2}[a^{2k} Y_{k+1}(x)] = a^{2k} Y_k(x)$$

$$Y_{k+1}(0) = 0, \quad \frac{\partial}{\partial x} Y_{k+1}(l) = 0$$

The corresponding initial condition is

$$u_n(x,0) = \phi(x) - f(0)\,Y_1(x) - \cdots - a^{2n-2} \frac{d^{2n-2} f(0)}{dt^{2n-2}} Y_n(x) \tag{10}$$

Suppose the solution of u_n can be expressed approximately by a linear combination of m free vibrational modes of the system, (equations (1) and (2) with $f(t) = 0$),

$$u_n(x,t) = \sum_{j=1}^{m} \eta_j(t)\,U_j(x) \tag{11}$$

Then

$$u(x,t) = \sum_{j=1}^{m} \eta_j(t)\,U_j(x) + \sum_{i=1}^{n} \xi_i(t)\,Y_i(x) \tag{12}$$

where

$$\xi_i(t) = a^{2(i-1)} \frac{\partial^{2(i-1)} f(t)}{\partial t^{2(i-1)}} \tag{13}$$

From equation (12), $u(x,t)$ is expressed as a linear combination of two sets of space functions: the *free vibration modes*, $U_i(x)$, and the *forced modes*, $Y_j(x)$.

The forced modes are intrinsic to the structure and the distribution of the external forces and are independent of the time characteristics of the forcing function. The solution $u(x, t)$ can be expressed as a linear combination of a few forced modes and a few free vibrtation modes provided that the forcing function $f(t)$ is a polynomial function and the initial conditions are such that the $u_n(x, 0)$ are a linear combination of the free vibration modes. In general, it is expected that using forced modes can greatly accelerate the convergence rate, thus requiring fewer free vibration modes. The forced modes converge to the first free vibration mode very quickly (see Appendix for a heuristic proof). Therefore, *it is generally sufficient to take only a few forced modes.*

In many real problems, the structure is often a component of a complex system, and $f(t)$ is an unknown interactive force. It is also possible that $f(t)$ is too complex to differentiate explicitly. In spite of this, since the linear combination form of $u(x, t)$ is known, as in equation (12), Lagrange's equations can be applied to derive the equations of motion. The forced modes can also be normalized for numerical consideration.

Forced modes do not satisfy orthogonality in general. The set of forced modes and free vibration modes can be orthogonalized with respect to either the mass or stiffness matrix, but not both. An orthogonalized assumed mode set (either with respect to the mass or stiffness matrix) is recommended in applications, since it avoids the numerical errors caused by a linear combination of very similar mode shapes.

3 General Case

We now state the results for a general flexible structure. The equations of motion are

$$\mathbf{L}_{\mathbf{x}}[\mathbf{u}] = m(\mathbf{x}) \frac{\partial^2 \mathbf{u}}{\partial t^2} + f(t)\,\mathbf{Y_0}(\mathbf{x}), \quad x \in \mathbf{D} \tag{14}$$

$$\hat{\mathbf{L}}_{\mathbf{x}}[\mathbf{u}] = f(t)\,\hat{\mathbf{Y}}_{\mathbf{0}}(\mathbf{x}), \quad \mathbf{x} \in \partial\mathbf{D}_1 \tag{15a}$$

$$\mathbf{u}(x,t) = \mathbf{0}, \quad \mathbf{x} \in \partial\mathbf{D}_2 \tag{15b}$$

$$\mathbf{u}(\mathbf{x},\mathbf{0}) = \phi(\mathbf{x}) \tag{15c}$$

where $\mathbf{x}$ is n-dimensional position vector, $\mathbf{u}$ is n-dimensional displacement vector, $\mathbf{L}$ and $\hat{\mathbf{L}}$ are time independent linear differential operators, $\mathbf{D}$ is the domain of the structure, $\partial\mathbf{D}_1$ is the portion of boundary with a stress boundary condition, and $\partial\mathbf{D}_2$ is the portion of the boundary with a displacement boundary condition. Both discrete and distributed external forces can be applied as long as they depend on a single time function, $f(t)$. n is the dimension of $\mathbf{D}$. $n = 3$ in three-dimensional problems, 2 in plane problems, and 1 in beam or other one-dimensional problems. The general case of a nonuniform mass distribution is considered.

Area of Cross Section: A
Density: ρ

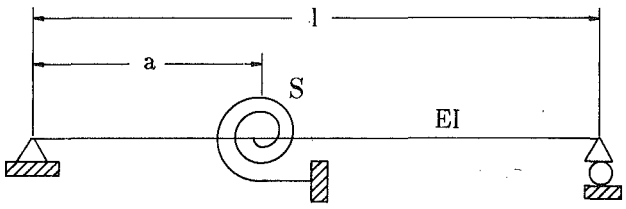


Fig. 2

In an analogous manner to the previous simple case, the forced modes, $\mathbf{Y}_i$, are found from the following equations:

$$\mathbf{L}_x[\bar{\mathbf{Y}}_1(\mathbf{x})] = \mathbf{Y}_0(\mathbf{x}), \quad \mathbf{x}\in\mathbf{D} \tag{16}$$

$$\hat{\mathbf{L}}_x[\bar{\mathbf{Y}}_1(\mathbf{x})] = \hat{\mathbf{Y}}_0(\mathbf{x}), \quad \mathbf{x}\in\partial\mathbf{D}_1 \tag{16a}$$

$$\bar{\mathbf{Y}}_1(\mathbf{x}) = \mathbf{0}, \quad \mathbf{x}\in\partial\mathbf{D}_2 \tag{16b}$$

$$\mathbf{Y}_1 = \bar{\mathbf{Y}}_1/\|\bar{\mathbf{Y}}_1\| \tag{16c}$$

$$\mathbf{L}_x[\bar{\mathbf{Y}}_{i+1}(\mathbf{x})] = m(\mathbf{x})\mathbf{Y}_i(\mathbf{x}), \quad \mathbf{x}\in\mathbf{D} \tag{17}$$

$$\hat{\mathbf{L}}_x[\bar{\mathbf{Y}}_{i+1}(\mathbf{x})] = \mathbf{0}, \quad \mathbf{x}\in\partial\mathbf{D}_1 \tag{17a}$$

$$\bar{\mathbf{Y}}_{i+1}(\mathbf{x}) = \mathbf{0}, \quad \mathbf{x}\in\partial\mathbf{D}_2 \tag{17b}$$

$$\mathbf{Y}_{i+1} = \bar{\mathbf{Y}}_{i+1}/\|\bar{\mathbf{Y}}_{i+1}\| \tag{17c}$$

$$\mathbf{i} = 1,2,\cdots n-1$$

where $\|\bullet\|$ is some norm defined on the function space of all the possible displacements which do not violate the geometric boundary condition. The function $\mathbf{u}$ is expanded as

$$\mathbf{u}(\mathbf{x},t) = \mathbf{u}_n(\mathbf{x},t) + c_1 f(t)\mathbf{Y}_1(\mathbf{x}) + \ldots + c_n \frac{d^{2n-2}f(t)}{dt^{2n-2}}\mathbf{Y}_n(\mathbf{x}) \tag{18}$$

where the c_i are constants introduced to account for the fact that the forced modes are normalized, $c_{i+1} = c_i\|\bar{\mathbf{Y}}_{i+1}\|$, $c_1 = \|\bar{\mathbf{Y}}_1\|$.

The equation for $\mathbf{u}_n(\mathbf{x}, \mathbf{t})$ is

$$\mathbf{L}_x[\mathbf{u}_n] = m(\mathbf{x})\frac{\partial^2\mathbf{u}_n}{\partial t^2} + c_n\frac{d^{2n-2}f(t)}{dt^{2n-2}}\mathbf{Y}_n(\mathbf{x}), \quad \mathbf{x}\in\mathbf{D} \tag{19}$$

with

$$\hat{\mathbf{L}}_x[\mathbf{u}_n] = \mathbf{0}, \quad \mathbf{x}\in\partial\mathbf{D}_1 \tag{20a}$$

$$\mathbf{u}_n(\mathbf{x},t) = \mathbf{0}, \quad \mathbf{x}\in\partial\mathbf{D}_2 \tag{20b}$$

$$\mathbf{u}_n(\mathbf{x},0) = \phi(\mathbf{x}) - c_1 f(0)\mathbf{Y}_1(\mathbf{x}) - \cdots - c_n\frac{d^{2n-2}f(0)}{dt^{2n-2}}\mathbf{Y}_n(\mathbf{x}) \tag{20c}$$

If $\mathbf{u}_n$ can be approximately expressed in the form of equation (11), then $\mathbf{u}$ can be expanded in the form of equation (12), with

$$\xi_i(t) = c_i\frac{d^{2i-2}}{dt^{2i-2}}f(t) \tag{21}$$

The remarks of the last three paragraphs of the last section are also true for the general case.

It should be noted that the forced modes in the general case may not have a closed form expression and that a numerical solution will have to be employed. The discretization of equations (16) and (17) results in a large set of linear equations. Several methods are available that efficiently address this problem (Lanczos, 1950; Wilson and Itoh, 1983; Parlett, 1980; Nour-Omid et al., 1983).

4 Numerical Example

The system to be considered is shown in Fig. 2. The pinned-pinned beam represents the structure, while the torsional spring represents the interactive force. This system was previously studied (Pierre, 1985) for *separation phenomena*. To keep the formulation simple, the position constraint in the midspan of the beam is removed. The natural frequencies and the natural modes of the whole system are computed by means of a Rayleigh-Ritz formulation.

The deflection of the beam is expressed as

$$w(x,t) = \sum_{i=1}^{N_1} q_i(t)\phi_i(x) + \sum_{i=1}^{N_2} p_i(t)\psi_i(x) \tag{22}$$

where

$$\phi_i(t) = \sin\frac{i\pi x}{l}, \quad i = 1,2,\ldots,N_1 \tag{23}$$

are the free vibration modes of the pinned-pinned beam. The $q_i(t)$'s are the corresponding modal coordinates. The $\psi_i(x)$'s are the forced modes due to a force couple exerted on the beam at $x = a$. For example,

$$\psi_1(x) = \begin{cases} \bar{x}(1-3\bar{b}^2-\bar{x}^2), & \text{for } x<a \\ -\bar{x}^3+3(\bar{x}-\bar{a})^2+(1-3\bar{b}^2)\bar{x}, & \text{for } x\geq a \end{cases} \tag{24}$$

is the shape of deflection due to a couple applied at $x = a$. In equation (24) and the following, $b = l - a$, and the variables with a bar represent the corresponding variables divided by l.

The deflection due to the distributed forces of $\psi_1(x)$ applied to the pinned-pinned beam is given by:

$$\psi_2(x) = \begin{cases} \bar{x}\left(-\frac{4}{315}+\frac{\bar{a}^2}{15}-\frac{\bar{a}^4}{12}+\frac{\bar{a}^5}{20}-\frac{\bar{a}^6}{120}\right)+\bar{x}^3\left(\frac{1}{45}-\frac{\bar{a}^2}{6}+\frac{\bar{a}^3}{6}-\frac{\bar{a}^4}{24}\right) \\ \quad +\bar{x}^5\left(-\frac{1}{60}+\frac{\bar{a}}{20}-\frac{\bar{a}^2}{40}\right)-\frac{\bar{x}}{840}, & \text{for } x<a \\ \frac{\bar{a}^6}{120}+\bar{x}\left(-\frac{4}{315}+\frac{\bar{a}^2}{15}-\frac{\bar{a}^4}{12}-\frac{\bar{a}^6}{120}\right)+\frac{\bar{a}^4\bar{x}^2}{8}+\bar{x}^3\left(\frac{1}{45}-\frac{\bar{a}^2}{6}-\frac{\bar{a}^4}{24}\right) \\ \quad +\frac{\bar{a}^2\bar{x}^4}{8}+\bar{x}^5\left(-\frac{1}{60}-\frac{\bar{a}}{40}\right)+\frac{\bar{x}^6}{120}-\frac{\bar{x}^7}{840}, & \text{for } x\geq a \end{cases} \tag{25}$$

The kinetic energy of the system is

$$T = \frac{1}{2}\int_0^l \rho A\left(\frac{\partial w}{\partial t}\right)^2 dx$$

$$= \frac{1}{2}\left[\sum_{i=1}^{N_1} m_i^0\dot{q}_i^2(t) + \sum_{i=1}^{N_2}\sum_{j=1}^{N_2} m_{ij}^2\dot{p}_i(t)\dot{p}_j(t) + 2\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} m_{ij}^3\dot{q}_i(t)\dot{p}_j(t)\right] \tag{26}$$

where

$$\begin{cases} m_i^0 = \rho A \int_0^l \phi_i^2(x)\,dx \\ m_{ij}^2 = m_{ji}^2 = \rho A \int_0^l \psi_i(x)\psi_j(x)\,dx \\ m_{ij}^3 = \rho A \int_0^l \phi_i(x)\psi_j(x)\,dx \end{cases} \tag{27}$$

The potential energy is

$$P = \frac{1}{2}\int_0^l EI\left(\frac{\partial^2 w}{\partial x^2}\right)^2 dx + \frac{1}{2} S[w'(a)]^2$$

$$= \frac{1}{2}\left[\sum_{i=1}^{N_1} k_i^0 q_i^2(t) + \sum_{i=1}^{N_2}\sum_{j=1}^{N_2} k_{ij}^2 p_i(t)p_j(t) + 2\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} k_{ij}^3 q_i(t)p_j(t)\right] + \frac{1}{2} S[w'(a)]^2 \tag{28}$$

where

$$\begin{cases} k_i^0 = EI\int_0^l [\psi_i''(x)]^2 dx \\ k_{ij}^2 = k_{ji}^2 = EI\int_0^l \psi_i''(x)\psi_j''(x)\,dx \\ k_{ij}^3 = EI\int_0^l \phi_i''(x)\psi_j''(x)\,dx \end{cases} \tag{29}$$

A constraint

$$f = w'(a) - \sum_{i=1}^{N_1} q_i(t)\phi_i'(a) - \sum_{i=1}^{N_2} p_i(t)\psi_i'(a) = 0 \tag{30}$$

is imposed to account for the spring interaction and is introduced to the Lagrangian by means of the Lagrange Multiplier μ:

$$L = T - P - \mu f \tag{31}$$

Applying Lagrange's equation yields:

$$m_i^0\ddot{q}_i + \sum_{j=1}^{N_2} m_{ij}^3\ddot{p}_j + k_i^0 q_i + \sum_{j=1}^{N_2} k_{ij}^3 p_j + \mu\phi_i'(a) = 0,$$

$$i = 1,2,\ldots,N_1 \tag{32}$$

$$\sum_{j=1}^{N_2} m_{ij}^2\ddot{p}_j + \sum_{j=1}^{N_1} m_{ji}^3\ddot{q}_j + \sum_{j=1}^{N_2} k_{ij}^2 p_j + \sum_{j=1}^{N_1} k_{ji}^3 q_j + \mu\psi_i'(a) = 0,$$

$$i = 1,2,\ldots,N_2 \tag{33}$$

$$\mu - Sw'(a) = 0 \tag{34}$$

$$w'(a) = \sum_{i=1}^{N_1} q_i(t)\phi_i'(a) + \sum_{i=1}^{N_2} p_i(t)\psi_i'(a) \tag{35}$$

Let

$q_i = a_i e^{j\omega t}$, $p_i = b_i e^{j\omega t}$, $w'(a) = ce^{j\omega t}$, $\mu = de^{j\omega t}$

After cancelling the factor $e^{k\omega t}$, equation (32)–(35) become

$$(k_i^0 - m_i^0\omega^2)a_i + \sum_{j=1}^{N_2}(k_{ij}^3 - m_{ij}^3\omega^2)b_j + \phi_i'(a)d = 0 \tag{36}$$

$$\sum_{j=1}^{N_2}(k_{ij}^2 - m_{ij}^2\omega^2)b_j + \sum_{j=1}^{N_1}(k_{ji}^3 - m_{ji}^3\omega^2)a_j + \psi_i'(a)d = 0 \tag{37}$$

$$d - Sc = 0 \tag{38}$$

$$c = \sum_{i=1}^{N_1} a_i\phi_i'(a) + \sum_{i=1}^{N_2} b_i\psi_i'(a) \tag{39}$$

From equation (38),

$$d = Sc \tag{40}$$

Solving equation (36) for a_i, and considering equation (40), gives

$$a_i = -\frac{1}{k_i^0 - m_i^0\omega^2}\left[\sum_{j=1}^{N_2}(k_{ij}^3 - m_{ij}^3\omega^2)b_j + S\phi_i'(a)c\right] \tag{41}$$

Substituting equations (40) and (41) in equations (37) and (39) gives:

$$\sum_{j=1}^{N_2}\left[(k_{kj}^2 - m_{kj}^2\omega^2) - \sum_{i=1}^{N_1}\frac{1}{k_i^0 - m_i^0\omega^2}(k_{ik}^3 - m_{ik}^3\omega^2)(k_{ij}^3 - m_{ij}^3\omega^2)\right]b_j$$

$$+\left[S\psi_k'(a) - \sum_{i=1}^{N_1}\frac{(k_{ij}^3 - m_{ik}^3\omega^2)S\phi_i'(a)}{k_i^0 - m_i^0\omega^2}\right]c = 0, \quad k = 1,2,\ldots,N_2 \tag{42}$$

$$\sum_{i=1}^{N_2}\left[S\psi_j'(a) - \sum_{i=1}^{N_1}\frac{S\phi_i'(a)(k_{ij}^3 - m_{ij}^3\omega^2)}{k_i^0 - m_i^0\omega^2}\right]b_j - \left[S + \sum_{i=1}^{N_1}\frac{S^2[\phi_i'(a)]^2}{k_i^0 - m_i^0\omega^2}\right]c = 0 \tag{43}$$

Equations (42) and (43) comprise a set of homogeneous linear equations in the variables b_j, $j = 1, 2, \ldots, N_2$ and c. Setting the determinant of the coefficients of these equations to zero gives the natural frequencies of vibration. A corresponding nontrivial solution can be substituted into equation (41) to find the corresponding a_i's. This allows a mode shape to be computed from the relation

$$W(x) = \lambda\left[\sum_{i=1}^{N_1} a_i\phi_i + \sum_{j=1}^{N_2} b_j\psi_j\right] \tag{44}$$

Three kinds of assumed mode sets are used in the following computations: (1) Free vibration modes of the pinned-pinned beam only; (2) one forced mode and several free vibration modes; (3) two forced modes and several free vibration modes. The numerical values of the parameters used and the results of the numerical computation are summarized in Table 1. In order to compare the shapes of the natural modes, 19 evenly distributed points are selected such that

$$x_i = \frac{i}{20}\,l, \quad i = 1,2,\ldots,19$$

λ, in equation (44), is chosen such that

$$\sqrt{\frac{1}{19}\sum_{i=1}^{19} W^2(x_i)} = 1$$

The odd order frequencies contain smaller errors, due to the fact that the slopes of these modes at $x = a = 0.49$ are very small, and therefore the interactive force has relatively little effect on the modes. It is seen that using only vibration modes give relatively poor numerical results. To bring the second frequency to within 1 percent error requires more than 50 modes; to bring the first two modes shapes to within 1 percent error requires more than 100 modes. Even 200 assumed modes cannot bring the fourth mode shape to within 1 percent error. Significant improvements are achieved by using one forced mode: one forced mode plus 5 free vibration modes give the first 4 frequencies to within 0.5 percent error. To find the first

Table 1 Results of computation †

A. Frequency Errors

Number of Modes Used $N_1 + N_2$*	First Frequency Error	Second Frequency Error	Third Frequency Error	Fourth Frequency Error
10 + 0	0.00129	0.56524	0.00773	1.9861
50 + 0	0.00003	0.10383	0.00274	0.3340
100 + 0	0.00001	0.05027	0.00130	0.1598
200 + 0	0.00000	0.02499	0.00064	0.0790
1000 + 0	0.00000	0.00498	0.00012	0.0157
5 + 1	0.00000	0.00018	0.00001	0.0889
6 + 1	0.00000	0.00001	0.00000	0.0044
7 + 1	0.00000	0.00001	0.00000	0.0037
4 + 2	0.00000	0.00000	0.00000	0.0086
5 + 2	0.00000	0.00000	0.00000	0.0005
True Values	1.00115	6.22834	9.03366	20.1359

B. Mode Shape Errors**

Number of Modes Used $N_1 + N_2$*	First Mode Shape Error	Second Mode Shape Error	Third Mode Shape Error	Fourth Mode Shape Error
10 + 0	1.5×10^{-3}	1.7×10^{-1}	7.4×10^{-2}	3.31
50 + 0	3.2×10^{-4}	2.6×10^{-2}	1.1×10^{-2}	6.3×10^{-2}
100 + 0	1.5×10^{-4}	1.2×10^{-2}	5.3×10^{-3}	2.9×10^{-2}
200 + 0	7.8×10^{-5}	6.4×10^{-3}	2.6×10^{-3}	1.4×10^{-2}
1000 + 0	1.6×10^{-5}	1.2×10^{-3}	5.1×10^{-4}	2.8×10^{-3}
5 + 1	$< 10^{-5}$	1.4×10^{-3}	6.5×10^{-4}	8.2×10^{-2}
6 + 1	$< 10^{-5}$	3.0×10^{-4}	1.2×10^{-4}	1.1×10^{-2}
7 + 1	$< 10^{-5}$	2.0×10^{-4}	8.5×10^{-5}	7.6×10^{-3}
4 + 2	$< 10^{-5}$	4.4×10^{-5}	4.2×10^{-5}	6.0×10^{-2}
5 + 2	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$	3.9×10^{-3}

*N_1 is the number of free modes, N_2 is the number of forced modes.
†The numerical values of the parameters used in the computation are:
$l = 1, \quad EI = 1, \quad S = 500\pi^4, \quad \rho A = \pi^4, \quad a = 0.49l$
**Error $\equiv \max_{1 \le i \le 19} |W(x_i) - W_{true}(x_i)|$

4 mode shapes to within 1 percent error only requires 1 forced mode plus 7 free modes. Further improvements are observed by using 2 forced modes: two forced modes plus 5 free modes yield estimates of the first four modes with only half the error obtained by using one forced mode and 7 free modes.

5 Extension and Conclusion

In the above, only one independent interactive force was discussed. However, an extension to two or more independent interactive forces is straightforward due to the superposition principle. If one has N_0 free modes and N_k forced modes for the kth independent interactive force, the displacement can be expressed as

$$w(x,t) = \sum_{k=0}^{M} \sum_{i=1}^{N_k} \eta_i^k(t)\phi_i^k(x)$$

where M is the number of independent interactive forces, $\phi_i^0(x)$, $i = 1, 2, \ldots, N_0$ are N_0 free vibration assumed modes, and $\phi_i^k(x)$, $i = 1, 2, \ldots, N_k$ is the set of forced modes due to the kth independent interactive force.

The following conclusions can be drawn:

1. The convergence rate of an assumed mode analysis, where only a few independent interactive forces are involved, can be greatly accelerated by including a few forced modes.
2. The forced modes are intrinsic to the structure and the spatial distribution of the forces. An exact description can be achieved by use of these modes under appropriate conditions.
3. It is usually adequate to use a few forced modes, since the forced mode shapes usually approach the first free vibration mode shape very rapidly with increasing mode number.

References

Baker, M., 1984, "Component Mode Synthesis Methods for Test-Based, Rigidly Connected, Flexible Components," a Collection of Technical Papers, *AIAA/ASME/ASCE/AHS 25th Structures, Structural Dynamics and Materials Conference and AIAA Dynamics Specialists Conference*, Part 2, Palm Springs, May, pp. 153-163.

Bisplinghoff, R. L., and Ashley, H., 1962, *Principles of Aeroelasticity*, Wiley, New York

Book, W. J., 1984, "Recursive Lagrangian Dynamics of Flexible Manipulator Arms," *The International Journal of Robotics Research*, Vol. 3, No. 3, Fall, pp. 87-101.

Craig, R. R., Jr., 1985, "A Review of Time-Domain and Frequency-Domain Component Mode Synthesis Methods," paper presented at Joint ASME/ASCE Applied Mechanics Conference, Albuquerque, N.M.

Lanczos, C., 1950, "An Interation Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators," *Journal of Research of the National Bureau of Standards*, Vol. 45, No. 4, Oct., pp. 255-282.

Meirovitch, L., 1975, *Elements of Vibration Analysis*, McGraw Hill.

Nour-Omid, Bahram, et al., 1983, "A Newton-Lanczos Method for Solution of Nonlinear Finite Element Equations," *Computers and Structures*, Vol. 16, No. 1-4, pp. 241-252.

Nurre, G. S., Ryan, R. S., Scofield, H. N., and Sims, J. L., 1984, "Dynamics and Control of Large Space Structures," *Journal of Guidance, Control and Dynamics*, Vol. 7, No. 5, pp. 514-526.

Parlett, B. N., 1980, "A New Look at the Lanczos Algorithm for Solving Symmetric Systems of Linear Equations," *Linear Algebra and Its Applications*, Vol. 29, Feb., pp. 323-346.

Pierre, C., 1985, "Analysis of Structural Systems with Parameter Uncertainties," Ph.D. Dissertation, Duke University, Department of Mechanical Engineering and Materials Science.

Wilson, E. L., and Itoh, T., 1983, "An Eigensolution Strategy for Large Systems," *Computers and Structures*, Vol. 16, No. 1-4, pp. 259-265.

APPENDIX

Equations (17) can be discretized by the finite element method to yield the equation

$$\bar{\mathbf{K}}\bar{\mathbf{Y}}'_{i+1} = \mathbf{P}_i \qquad (A1)$$

where $\mathbf{P}_i$ is the force vector obtained by considering $m(\mathbf{x})\mathbf{Y}_i(\mathbf{x})$ as the distributed force. In accord with the virtual work principle, it is required that

$$(\delta^*)^T\mathbf{P}_i = \int_D \left(\sum_{j=1}^{N} N_j(\mathbf{x})\delta_j^*\right)^T m(\mathbf{x})\mathbf{Y}_i(\mathbf{x})d\mathbf{x}$$

$$= \sum_{j=1}^{N} \delta_j^{*T} \int_D N_j(\mathbf{x})m(\mathbf{x})\mathbf{Y}_i(\mathbf{x})d\mathbf{x}$$

where $N_j(\mathbf{x})$ is the global shape function for the system. This is the sum of all the elemental shape functions and is nonzero only in the neighboring elements of the node j. δ^* is the virtual displacement. Since δ^* is arbitrary, the jth nodal value of the vector $\mathbf{P}_i$ is expressed as

$$\mathbf{P}_{ij} = \int_D N_j(\mathbf{x})m(\mathbf{x})\mathbf{Y}_i(\mathbf{x})d\mathbf{x}$$

$\mathbf{Y}_i(\mathbf{x})$ is a continuous function which can be approximated by its values at all the nodes multiplied by the shape functions:

$$\mathbf{Y}_i(\mathbf{x}) = \sum_{k=1}^{N} N_k(\mathbf{x})\mathbf{Y}'_{ik}$$

It follows that

$$\mathbf{P}_{ij} = \int_D N_j(\mathbf{x})m(\mathbf{x}) \sum_{k=1}^{N} N_k(\mathbf{x})\mathbf{Y}'_{ik}d\mathbf{x}$$

$$= \sum_{k=1}^{N} \left(\int_D m(\mathbf{x})N_j(\mathbf{x})N_k(\mathbf{x})d\mathbf{x}\mathbf{I}\right)\mathbf{Y}'_{ik}$$

The matrix in the parentheses is recognized as the jkth submatrix of the mass matrix. Therefore:

$$\mathbf{P}_{ij} = \sum_{k=1}^{N} \bar{\mathbf{M}}_{jk} \mathbf{Y}'_{ik}$$

$$\mathbf{P}_i = \bar{\mathbf{M}} \mathbf{Y}'_i$$

Equation ($A1$) thus becomes

$$\bar{\mathbf{K}} \bar{\mathbf{Y}}'_{i+1} = \bar{\mathbf{M}} \mathbf{Y}'_i$$

After substituting the displacement boundary conditions ($17b$), several rows and columns can be cancelled from the singular stiffness matrix $\bar{\mathbf{K}}$ and mass matrix $\bar{\mathbf{M}}$ to yield

$$\mathbf{K} \bar{\mathbf{Y}}_{i+1} = \mathbf{M} \mathbf{Y}_i \tag{A2}$$

where $\mathbf{K}$ and $\mathbf{M}$ are nonsingular. Therefore

$$\mathbf{Y}_n = c \mathbf{D}^{n-1} \mathbf{Y}_1 \tag{A3}$$

where

$$\mathbf{D} = \mathbf{K}^{-1} \mathbf{M} \tag{A4}$$

Equation ($A3$) is recognized as a power method, which tends to sort out the eigenvector corresponding to the dominant eigenvalue of the matrix $\mathbf{D}$, provided that the expansion of the original vector $\mathbf{Y}_1$ in terms of an eigengvector set contains this eigenvector. This eigenvector corresponds to the least eigenvalue (low frequency) of the system. When a sufficiently fine discretization is used, this approaches the first mode of the system (14)–(15).

General Dynamic Equations of Helical Springs With Static Solution and Experimental Verification

Yuyi Lin
Research Assistant.

Albert P. Pisano
Assistant Professor.
Assoc. Mem. ASME

Mechanical Engineering Department,
University of California,
Berkeley, CA 94720

The general dynamic equations of helical compression springs with circular wire cross section, variable pitch angle, and variable helix radius are derived. The equations are formulated by Hamilton's principle and a variational method. In contrast to previous studies, the effects of coil flexure bending, variable pitch angle and variable helix radius are taken into account. The general equations are shown to agree with dynamic equations found in literature when the general equations are reduced to simplified forms. For a specific helical spring and static loading, the equations are solved with both the predicted radial expansion and the predicted longitudinal spring compression force in excellent agreement with experimental data.

Introduction

The properties of helical springs have interested many researchers, dating back to early experimental work (Donkin, 1929; Jehle, 1929; Lehr, 1933) conducted with internal-combustion engine valve springs. To improve the performance of such springs, many phenomena related to the dynamics of the helical spring, such as resonance, radial expansion, and transient stress have been studied.

Modern machinery, especially automotive valve trains and vehicle suspension systems, has relied on, to an ever-greater degree, helical spring designs that actively exploit variable pitch angle and variable helical radius. The current industrial practice is to use springs of variable pitch angle to thwart spring resonance and to provide progressive spring rates; the radial expansion of the spring helix is being occasionally used to provide coil vibration damping through coil contact with external, "cup," damping elements. This paper addresses the need to generalize the analysis of the dynamics of helical springs to include, simultaneously, both of these effects.

In the classical works of Thomson (1883) and Love (1927), the geometrical relationships among spring parameters, such as pitch angle, static load, coil curvature, and radius of spring helix were analytically derived. Many investigators have applied and modified Love's equations, and experimental results in the literature appear to confirm the theory. Phillips (1972) and Stokes (1974), for example, used the geometrical relations among wire curvature, helix radius, and pitch angle in their studies of spring response to impact. Chen (1983) rederived and used these relationships in his work on eigenvalue analysis and experiment.

The combined effects of pitch angle and coil curvature were studied by Ancker and Goodier (1958), who found that pitch angle did contribute significantly to strain and stress. This conclusion supported the early investigation of Sayre (1936), which was purely a static analysis. Some of the assumptions made in the Ancker and Goodier analysis (1958) limit the theory to infinitesimally small pitch angle (for example, the cross section of coil cut by a vertical plane was assumed to always remain circular), despite the authors' intent to analyze the large pitch angle case. Other assumptions (for example, uniform force and moment assumed for all cross sections) limit the analysis to static, and not dynamic, stress determination.

Wahl (1963) summarized spring research up to the year 1963, noting that radial expansion could be analyzed for large static displacements, but only for the case of small and constant pitch angles in helical springs. More recently, Wittrick (1966) extended Love's analysis to the large pitch angle case. In the analysis, however, the deformation due to tension and shearing forces in the coil, as well as the effect of curvature of the coil on the torsional and flexural rigidities, was neglected.

Phillips and Costello (1972) modeled the impact response of helical springs, admitting large deflections into their analysis. A radial expansion variable was included in the two, coupled, partial differential equations, with coefficients linearized for small strain and small constant pitch angle. The effects of the varying pitch angle at the ends of the spring, and at the boundary condition of the partial differential equations, were not explicitly stated in the research paper (Phillips, 1972). For a specific spring, the experimental results shown agreed well with the solution. In a later paper, Costello (1975) predicted a

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 23, 1987; final revision May 19, 1987.

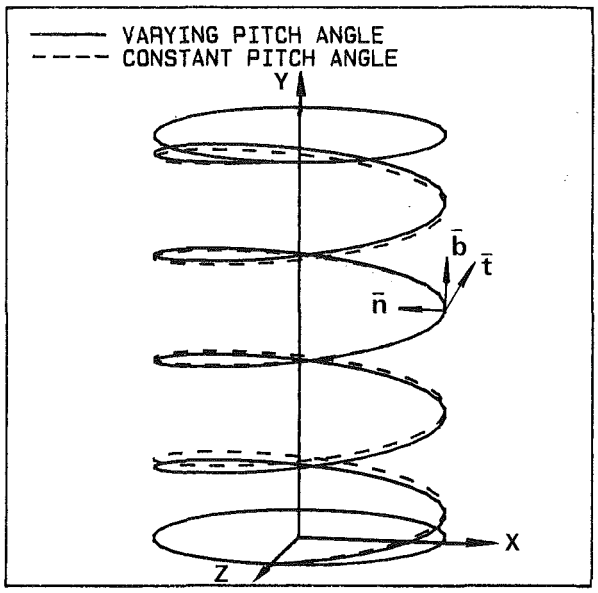


Fig. 1 Helical spring with end effects

much larger radial expansion than Love's theory. It was mentioned that the analytical result was tested in the previously reported dynamic experiment, but the spring analyzed was different from that previously reported in the experiment and the varying pitch angle was not considered in the analysis.

In Kato's work (1974), a varying pitch angle spring model was proposed and solved in the way similar to the Stefan problem in melting and freezing phenomena. However, the spring model (the partial differential equation describing the dynamics) is the simplest linear wave equation, which could be derived from just one term of strain energy and one term of kinetic energy, with linear viscous damping. Nonlinearity was introduced solely by the solution technique, not by inclusion of additional energy terms. For simplicity of solution, one end of the spring had been assumed having the smallest pitch angle and the other end the largest, which is generally not the case in industrial applications. Experimental results in the paper showed that the maximum discrepancy between theory and experiment could be well over ten percent.

In this paper, the full dynamic equations for helical springs are derived, including variable pitch angle and variable helix radius.

Geometrical Constraints Among Variables

From the Frenet formula of differential geometry (Do Carmo, 1976) it follows that curvature, $\kappa(s,t)$, and torsion, $\tau(s,t)$, (where s and t are independent variables for arc length and time, respectively,) can completely specify a curve in three-dimensional space. However, curvature and torsion are not convenient variables for engineering applications, and therefore other variables are defined in this paper for the formulation and solution of the dynamic equations of helical springs. The intrinsic coordinate system for some of these variables is shown in Fig. 1. The relationship between curvature and pitch angle can be derived as (Thomson, 1883):

$$\kappa(s,t)=\frac{\cos^2[p(s,t)]}{r(s,t)} \tag{1}$$

where $p(s,t)$ and $r(s,t)$ are pitch angle and radius of the helix, respectively. Torsion, the relative rotation between cross-sectional "slices" of spring wire, is defined by:

$$\tau(s,t)=\frac{\sin[p(s,t)]\cos[p(s,t)]}{r(s,t)} \tag{2}$$

It should be noted that the rotation angle, $\psi(s,t)$, of any cross section is the absolute rotation with respect to an inertial frame. This is not the same as torsion, $\tau(s,t)$, which has been conventionally used in place of rotation, $\psi(s,t)$. Here, $\psi(s,t)$ is defined as the integral of the rate of twist along the spring helix:

$$\psi(s,t)=\int_0^s\left(\frac{\sin[p(\zeta,t)]\cos[p(\zeta,t)]}{r(\zeta,t)}-\frac{\sin[p(\zeta,0)]\cos[p(\zeta,0)]}{r(\zeta,0)}\right)d\zeta \tag{3}$$

where the second term in the integrand is the initial torsion of the spring at the free length condition. By differentiating equation (3), and using the definition of $\tau(s,t)$ in equation (2), a relationship between "rate of twist" and rotation is obtained:

$$\frac{\partial\psi(s,t)}{\partial s}=\tau(s,t)-\tau(s,0) \tag{4}$$

The constraints in equation (1) through equation (4) can be used to define $\kappa(s,t)$, $\psi(s,t)$, and $\tau(s,t)$ in terms of $p_0(s)$, $r_0(s)$, $p(s,t)$, and $r(s,t)$. A constraint relating axial displacement, $y(s,t)$, and local pitch angle, $p(s,t)$ can be stated as an integral:

$$y(s,t)=\int_0^s\sin[p(\zeta,t)]d\zeta \quad 0\le s\le L \tag{5}$$

For convenience, the axial motion forced at one end of the spring helix (where $s = L$) can be defined as:

$$Y(t)=\int_0^L\sin[p(\zeta,t)]d\zeta \tag{5'}$$

where the interpretation of the above equation is that the total height of the spring equals the specified input, $Y(t)$. Equation (3) and equation (5) are nonholonomic, rheonomic constraints, to which the Lagrange multiplier technique (Rosenberg, 1977) is applied in the adjoining of these constraints to the dynamic equations. This formulation of the dynamic equations of the helical spring utilizes four generalized coordinates with two nonholonomic constraints to fully describe the spring coil motion.

Formulation of the Dynamic Equations

The dynamic equations describing the motion of the spring with varying pitch angle and varying helix radius can be deduced from Hamilton's principle:

$$\int_{t_1}^{t_2}\delta\left(\sum T_i-\sum U_i\right)dt=0 \tag{6}$$

where T_i is kinetic energy and U_i potential energy due to elastic strain in the helical spring. To incorporate damping effects, the extended Hamilton's Principle is utilized:

$$\int_{t_1}^{t_2}\delta\left(\sum T_i-\sum U_i\right)dt+\int_{t_1}^{t_2}\sum\bar{F}_i\delta\bar{r}_i dt=0 \tag{7}$$

The second integral is the virtual work done by all nonconservative, damping forces, $\bar{F}_i$, which may be functions of time, space, displacement, and time derivatives of displacement. Since damping effects can be handled separately, the dynamic equations are first derived without taking damping into account. It is worth noting that before adjoining the constraint force arising from nonholonomic constraints, the system can still be treated as holonomic (Whittaker, 1944), although the inclusion of such damping and constraint forces can make the formulation of the dynamic equations proceed in a manner not traditionally associated with the principles of the calculus of variations (Leitmann, 1963).

The total strain energy, ΣU_i, was taken to be the sum of strain energies from three specific phenomena. The first energy term is expressed as:

$$U_1 = \int_0^L \frac{1}{2} GJ\left(\frac{\partial \psi(s,t)}{\partial s}\right)^2 ds \tag{8}$$

where U_1 is the strain energy due to the torsional twist of the spring wire, G is shear modulus of the spring material, and J polar moment of inertia of the wire cross section. This energy term is expected to dominate the three strain energy terms, and is always included even in simplest linear spring models. A change in local wire curvature (measured on $\bar{t} - \bar{n}$ plane of the intrinsic coordinate frame) generates strain energy which is included in the second energy term:

$$U_2 = \int_0^L \frac{1}{2} EI[\kappa(s,t) - \kappa(s,0)]^2 ds \tag{9}$$

where E is elastic modulus of the spring material, and I is moment of inertia of the wire cross section. Finally, the strain energy due to the local change in pitch angle (corresponding to the curvature change on the $\bar{t} - \bar{b}$ plane of the intrinsic coordinate frame) can be written as:

$$U_3 = \int_0^L \frac{1}{2} EI[y''(s,t)]^2 ds \tag{10}$$

Here partial derivatives with respect to space variable, s, are denoted by primes. Note that equation (10) is based on the assumption that the normal to the spring wire, $\bar{n}$, is oriented along the helical radial direction (Fig. 1) from spring coil to helix centerline. This assumption is accurate for the case where the spring coil experiences small changes in local curvature.

Corresponding to the variables describing the strain energies, there can be considered three kinetic energy terms: T_1, T_2, and T_3. The rotation of the wire about its own geometric axis can store the kinetic energy defined as:

$$T_1 = \int_0^L \frac{1}{2} I_m\left(\frac{\partial \psi(s,t)}{\partial t}\right)^2 ds \tag{11}$$

where I_m is mass moment of inertia per unit length of spring wire. The radial motion of the wire has kinetic energy of:

$$T_2 = \int_0^L \frac{1}{2} m\left(\frac{\partial r(s,t)}{\partial t}\right)^2 ds \tag{12}$$

where m is mass per unit length of the spring wire, assumed constant. The kinetic energy of vertical motion, expected to be the dominant kinetic energy, can be expressed as:

$$T_3 = \int_0^L \frac{1}{2} m\left(\frac{\partial y(s,t)}{\partial t}\right)^2 ds \tag{13}$$

This expression accounts, approximately, for the kinetic energy resulting from translational motion along a spring axis, and is the only source of kinetic energy considered in simplified analyses (Love, 1927; Wahl, 1963; Kato, 1974) of the helical spring.

The application of Hamilton's principle and the variational method results in four equations of motion, which are fully derived in Appendix A. Utilizing the Lagrange multiplier technique, the nonholonomic constraints, equation (3) and equation (5) are adjoined into the equations. The four equations that result are:

$$\frac{md^2}{8}\frac{\partial^2 \psi(s,t)}{\partial t^2} - \frac{EI}{1+\nu}\frac{\partial^2 \psi(s,t)}{\partial s^2} - \lambda_2(s,t) = 0 \tag{14}$$

$$m\frac{\partial^2 y(s,t)}{\partial t^2} + EI\frac{\partial^4 y(s,t)}{\partial s^4} - \lambda_2(s,t) = 0 \tag{15}$$

$$m\frac{\partial^2 r(s,t)}{\partial t^2} - EI\left[\frac{\cos^2 p(s,t)}{r(s,t)} - \frac{\cos^2 p_0(s)}{r_0(s)}\right]\frac{\cos^2 p(s,t)}{r(s,t)^2}$$

$$-\lambda_1(s,t)\int_0^s \frac{\cos p(\zeta,t)\sin p(\zeta,t)}{r(s,t)^2} d\zeta = 0 \tag{16}$$

$$\frac{EI \sin 2p(s,t)}{r(s,t)}\left[\frac{\cos^2 p(s,t)}{r(s,t)} - \frac{\cos^2 p_0(s)}{r_0(s)}\right]$$

$$-\lambda_1(s,t)\int_0^s \frac{\cos 2p(\zeta,t)}{r(s,t)} d\zeta - \lambda_2(s,t)\int_0^s \cos p(\zeta,t) d\zeta = 0 \tag{17}$$

In above equations, ν is Poisson ratio of spring material and the $\lambda_i(s,t)$ are Lagrange multipliers. Since equation (17) is a natural result from the derivation, it must be satisfied by the solution that also satisfies equations (14)–(16). These equations, together with constraint equation (3) and equation (5), must be solved for unknowns $\lambda_1(s,t)$, $\lambda_2(s,t)$, $\psi(s,t)$, $y(s,t)$, $r(s,t)$, and $p(s,t)$. To facilitate a numerical solution, coefficients $\cos(p)$ and $\sin(p)$ in these equations can be approximated by:

$$\sin[p(s,t)] = p(s,t) \tag{18}$$

$$\cos[p(s,t)] = 1 - \frac{[p(s,t)]^2}{2} \tag{19}$$

although for this research, the above approximation was not used. Note that the size of the dynamic equations is considerably greater if second order terms in $p(s,t)$ are kept, but a solution containing at least the second order terms should be attempted, since dropping second order terms of $p(s,t)$ means ignoring energy terms T_1, U_1, and U_3.

Geometric Boundary Conditions

The geometric boundary conditions applied to the dynamic equations are determined intuitively, and require the specification of various helical spring design parameters. The boundary conditions concerning the local spring helix radius, $r(s,t)$, were chosen with the assumption that sufficient friction with the mounting surface forces the two ends of the helical spring to remain fixed at all times. At the two ends of the spring, the radius of the helix is given as:

$$r(0,t) = r_0(0) \tag{20}$$

$$r(L,t) = r_0(L) \tag{21}$$

The radius at the spring ends is considered constant in time, and thus:

$$\dot{r}(0,t) = \dot{r}(L,t) = 0 \tag{22}$$

Here the time derivative is denoted by the dot. At initial time, the local helix radius is a design parameter, and is expressed as a function of distance, s, along the helix:

$$r(s,0) = r_0(s). \tag{23}$$

According to conventional industrial practice, helical springs have a near-zero pitch angle at the ends (due to "closed end coils") and larger pitch angle at the middle of spring. This reasonable consideration is used in specifying the initial pitch angle distribution. Symbolically, then, the boundary conditions chosen for the pitch angle, $p(s,t)$, are:

$$p(0,t) = p(L,t) = 0 \tag{24}$$

Again, at initial time, pitch angle is a known function along the helix, specified by the spring designer:

$$p(s,0) = p_0(s) \tag{25}$$

The ends of the spring are assumed fixed against rotation, therefore, the boundary conditions on torsion and its time derivative can be set to zero at each end of the spring, as in:

$$\tau(0,t) = \tau(L,t) = 0 \tag{26}$$

$$\dot{\tau}(0,t) = \dot{\tau}(L,t) = 0 \tag{27}$$

Since the constraint among $\psi(s,t)$, $\tau(s,t)$, and $\tau(s,0)$ is known (Thomson, 1883), the required bounary conditions for $\psi(s,t)$ are found in equation (4)

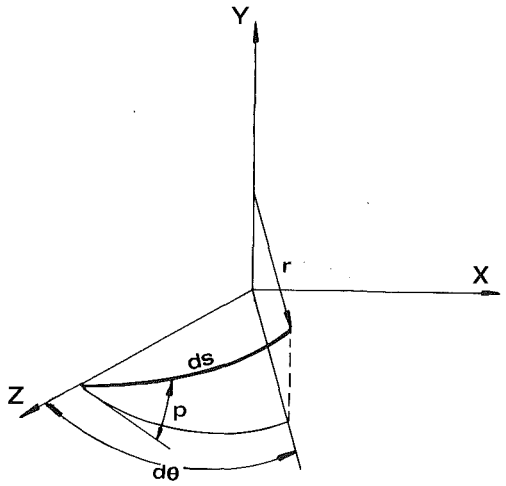


Fig. 2 Pitch angle-radius relationship

$$\frac{\partial \psi}{\partial s}(0,t) = \tau(0,t) = 0 \tag{28}$$

$$\frac{\partial \psi}{\partial s}(L,t) = \tau(L,t) = 0 \tag{29}$$

A time-varying boundary condition on the longitudinal displacement, $y(0,t)$, is the dynamic input to one end of the helical spring:

$$y(0,t) = Y(t) \tag{30}$$

$$\dot{y}(0,t) = \dot{Y}(t) \tag{31}$$

The other end, where $s = L$, is assumed fixed at a specific value y_0, which can be a displacment measured from the free length of the spring, because in cases where the spring is installed with a force preload, the maximum installed height is less than the free length. Thus, the displacement boundary conditions are:

$$y(L,0) = y_0 \tag{32}$$

$$\dot{y}(L,t) = 0 \tag{33}$$

Boundary conditions for y' and y'' are determined from constraint, equation (5), and the specifications of pitch angle boundary conditions:

$$y'(0,t = \sin p(0,t) \tag{34}$$

$$y'(L,t) = \sin p(L,t) \tag{35}$$

$$y''(0,t) = p'(0,t) \cos p(0,t) \tag{36}$$

$$y''(L,t) = p'(L,t) \cos p(L,t) \tag{37}$$

This set of boundary conditions allows the model to function in circumstances where the helical spring is in contact with two mounting surfaces perpendicular to spring axis, one fixed in space and one with a prescribed motion parallel to the axis of the spring helix. To free the spring from initial stress, no initial displacement, $y(L, 0)$ was assumed in this analysis. However, a precompression of the spring can be included in the solution of the dynamic equations by first solving a statical problem with input $\dot{Y}(t)$ zero and $Y(t)$ constant. Then the static solution, namely, $y(s,t_0)$, $\psi(s,t_0)$, and $p(s,t_0)$, can be used as the initial condition for the dynamic problem.

Comparison with Existing Dynamic Equations

As a check on the accuracy and generality of the derivation of the dynamic equations, comparisons were made with spring equations found in the literature. Since these dynamic equations are for cylindrical springs, i.e., for $r_0(s)$ = constant, a holonomic constraint relating radius, $r(s,t)$, and pitch angle, $p(s,t)$, can be derived (Appendix B) and used to simplify the general equations. This holonomic constraint is:

$$r(s,t) = \frac{r_0(s) \cos [p(s,t)]}{\cos [p_0(s)]} \tag{38}$$

and it can be interpreted physically as the local radial expansion due to the change of local pitch angle. In fact, it is a form of conservation of mass, provided inextensible spring wire is assumed. Embedding of constraint equation (38) in the formulation of dynamic equations, three nonlinear differential equations can be obtained:

$$\frac{md^2}{8}\frac{\partial^2 \psi(s,t)}{\partial t^2} - \frac{EI}{1+\nu}\frac{\partial^2 \psi(s,t)}{\partial s^2} - \lambda_1(s,t) = 0 \tag{39}$$

$$-m\left(\frac{r_0(s)}{\cos p_0(s)}\right)^2 \left[\cos p(s,t)\left(\dot{p}(s,t)\right)^2 + \sin p(s,t)\ddot{p}(s,t)\right] \sin p(s,t)$$
$$+EI\left(\frac{\cos p_0(s)}{r_0(s)}\right)^2 \left[\cos p(s,t) - \cos p_0(s)\right] \sin p(s,t)$$
$$-\lambda_1(s,t)\int_0^s \frac{\cos p(\zeta,t) \cos p_0(\zeta)}{r_0(\zeta)}\, d\zeta$$
$$-\lambda_2(s,t)\int_0^s \cos p(\zeta,t)\, d\zeta = 0 \tag{40}$$

$$m\ddot{y}(s,t) + EI\frac{\partial^4 y(s,t)}{\partial s^4} - \lambda_2(s,t) = 0 \tag{41}$$

Further simplifications can be made. For example, if the change of radius of the spring during the motion is to be ignored, then the radius function, $r(s,t)$ becomes a constant. If, in addition, the pitch angle is relatively small, then in the formulation of the dynamic equations, one can let:

$$\sin[p(s,t)] = p(s,t) \tag{42}$$

$$\cos[p(s,t)] = 1 \tag{43}$$

and neglect bending energies U_2 and U_3 (equation (9) and equation (10)), and the entire set of three nonlinear differential equations reduces to the single linear differential equation:

$$\frac{w}{g}(\pi r^2 d^2/4 + I_m)\frac{\partial^2 y}{\partial t^2} = GJ\frac{\partial^2 y}{\partial x^2} \tag{44}$$

where w is the weight of per volume spring material. This is the undamped wave equation for a helical spring appearing in many texts, as well as in Wahl (1963) and Pisano and Freudenstein (1983).

Another case to which comparison can be made is to Costello's equation (1972), which is stated verbatim:

$$\sin\alpha\left(1 - \frac{\nu}{1+\nu}\cos^2\alpha\right)\frac{\partial^2 \bar{u}}{\partial \bar{x}^2}$$
$$-\frac{\nu}{1+\nu}\sin^2\alpha \cos\alpha \frac{\partial^2 \bar{v}}{\partial \bar{x}^2} = \frac{\partial^2 \bar{u}}{\partial \bar{t}^2} \tag{45}$$

in which $\bar{v}$ (defined as normalized azimuthal motion, or the rotation of a spring element around the spring axis, $\bar{u}$) was normalized axial displacement, and α was small, constant pitch angle. Since $\bar{v} = 0$ at two ends, constant radius and constant pitch angle were assumed, $\bar{v}$ must be in higher order than $\bar{u}$. Also it can be seen in equation (45), using $\nu = 0.29$, $\alpha = 0.1$ as assumed in the Phillips paper, the coefficient of $\partial^2 \bar{v}/\partial \bar{x}^2$ is

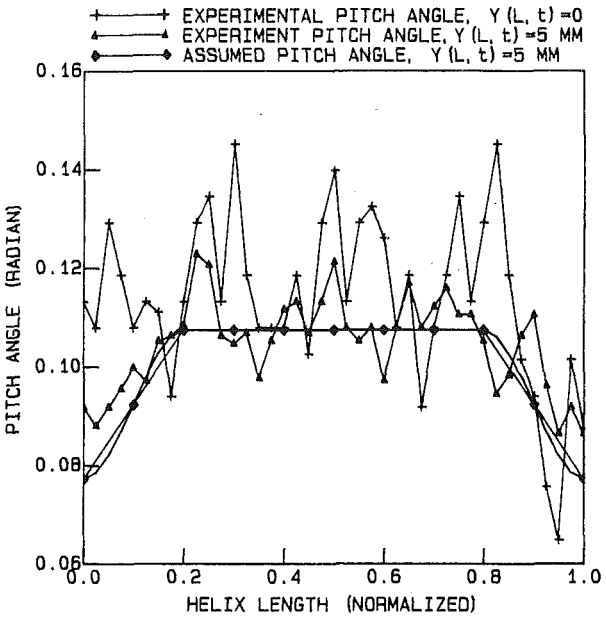


Fig. 3 Initial pitch angle distributions

about 35 times smaller than that of $\partial^2\bar{u}/\partial\bar{x}^2$ and 450 times smaller than that of $\partial^2\bar{u}/\partial\bar{t}^2$. Therefore, terms involving $\bar{\nu}$ are less important. If only two energy terms, U_1 and T_3 are used to formulate the equations developed in this paper, then the result is:

$$\frac{\sin p}{1+\nu}\frac{\partial^2 y}{\partial x^2}=\frac{Mr^2}{hEI}\frac{\partial^2 y}{\partial t^2} \tag{46}$$

where M is total mass of the spring. Equation (46) is identical to equation (45), Costello's equation, without the $\bar{\nu}$ terms.

Static Solution and Experimental Verification

It is clear that the set of nonlinear dynamic equations is difficult to solve analytically for the general case. An analytical solution for the cylindrical spring and the static case, however, can be expressed in a compact form. Since the loading is static, all terms involving time derivatives are set to zero. The first equation, equation (39), with constraint equation (3) yields:

$$\sin\,[p(s,t)]-\sin\,[p_0(s,0)]=\text{constant} \tag{47}$$

If the initial pitch angle distribution is known, this equation has a unique solution, subject to constraint, equation (5), and boundary conditions:

$$y(0,t)=0 \tag{48}$$

$$y(L,t_0)=Y(L,t_0) \tag{49}$$

where $Y(t_0)$ is static loading such that the compression of the spring is equal to the specified displacement. Since the solution is unique, it must also satisfy the other equations, if the system of equations has a solution at all. The initial pitch angle distribution, $p_0(s)$, must be found in order to complete the solution. This was done experimentally on one of the valve springs taken from a Pontiac, 1.8L, overhead cam, automotive engine. The resulting pitch angle distributions are shown in Fig. 3. Since the original design specification for the spring is not very restrictive, some other smoothed and approximated distribution curves besides the one shown here may also be used without violating the design specification. Some small variations from the distribution shown were tested. The changes in the predicted force and radial expansion are small.

To verify the correctness of the derived spring model and

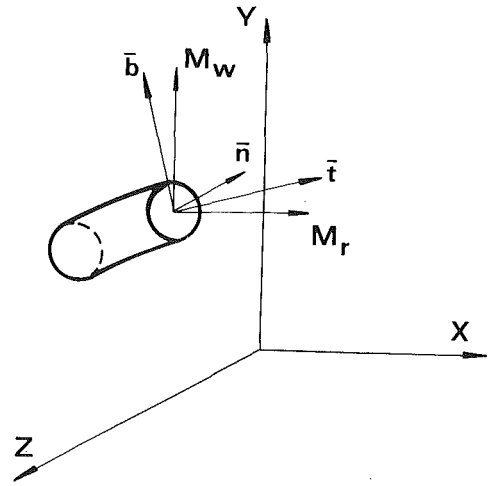


Fig. 4 Force and moment diagram

the solution under static loading, the applied force and moment produced by a given displacement were computed. In Fig. 4, a cross section of the spring wire in the middle portion of spring is shown. When a force, P, is applied, a torque, M_r, is produced at the cross section. This torque will be balanced by components of two moments resulting from strain energy U_1 and U_2 (U_3 is not counted here because the corresponding bending moment is approximately in the radial direction, and it does not contribute to either wind-up moment, M_w, or reaction force to P.) U_1 and U_2 both have components in y axis direction, which is the wind-up moment. From energy equation (9) and equation (10), one can write directly:

$$M_w=GJ\sin p\left(\frac{\sin p\cos p}{r}-\frac{\sin p_0\cos p_0}{r_0}\right)+$$

$$EI\cos p\left(\frac{\cos^2 p}{r}-\frac{\cos^2 p_0}{r_0}\right) \tag{50}$$

$$M_r=GJ\cos p\left(\frac{\sin p\cos p}{r}-\frac{\sin p_0\cos p_0}{r_0}\right)+$$

$$EI\sin p\left(\frac{\cos^2 p}{r}-\frac{\cos^2 p_0}{r_0}\right) \tag{51}$$

where p and r are functions of compression solvable from the static solution. Actual force displacement data were taken on a tensile testing machine fitted with a compression cage. The resolution of the force measurement apparatus is estimated to be 4.5 Newtons. The force-displacement data are shown in Fig. 5. To emphasize the differences between curves calculated from different theories, the linear solution has been subtracted from all data so that the nonlinear effects are made clear. It can be seen that the error introduced by linear spring theory can be as great as 39 Newtons. The actual force curve is well matched by the static solution, equation (47), with a maximum error less than 11 Newtons, which is close to the limit of precision of the measuring instrument.

It is shown in Fig. 5 that when end effects (varying pitch) are ignored, a large error in predicted force can result for the case of relatively large spring compressions. Figure 5 also shows the curve calculated by the Sayre-de Forest (1936) nonlinear formula using the same spring parameters. It can be seen that the nonlinear correction term has a very small effect on the result, and therefore the calculated force for Sayre-de Forest is basically linear but with a different slope than the linear theory. The resulting curve from the Phillips and Costello (1972) equations is not shown, since assumption of constant

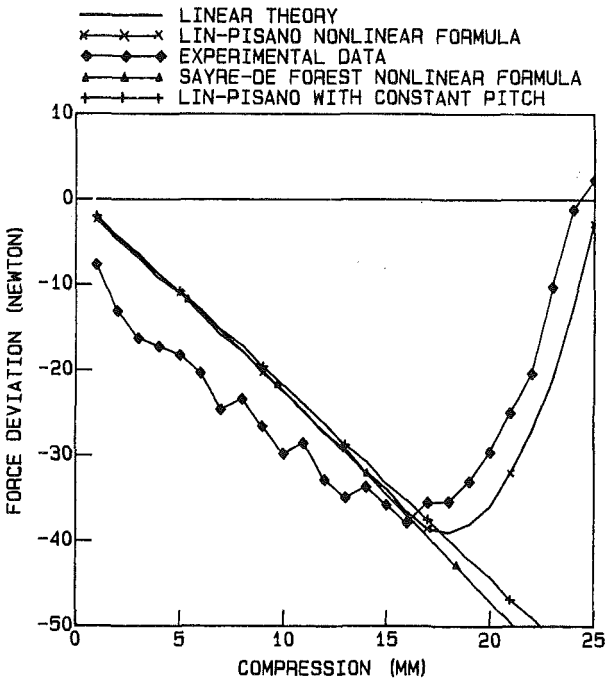


Fig. 5 Force-displacement curve comparison—deviation from linear theory

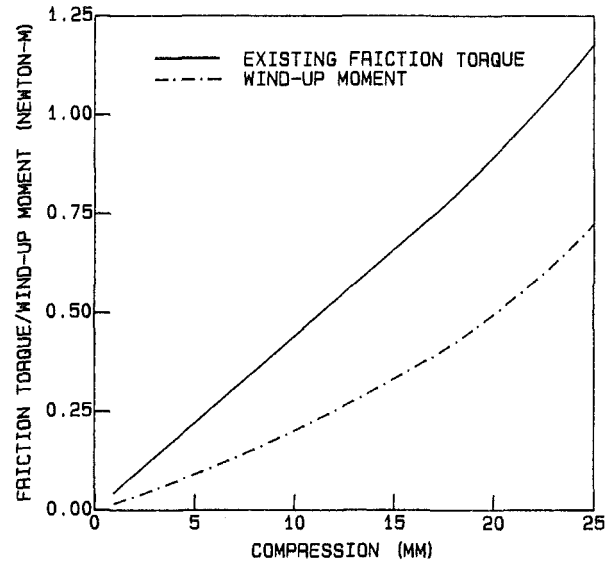


Fig. 6 Friction torque and wind-up moment versus compression

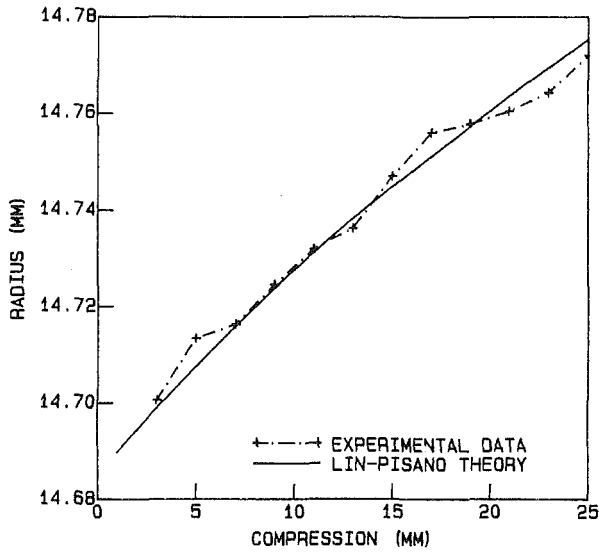


Fig. 7 Radial expansion verification

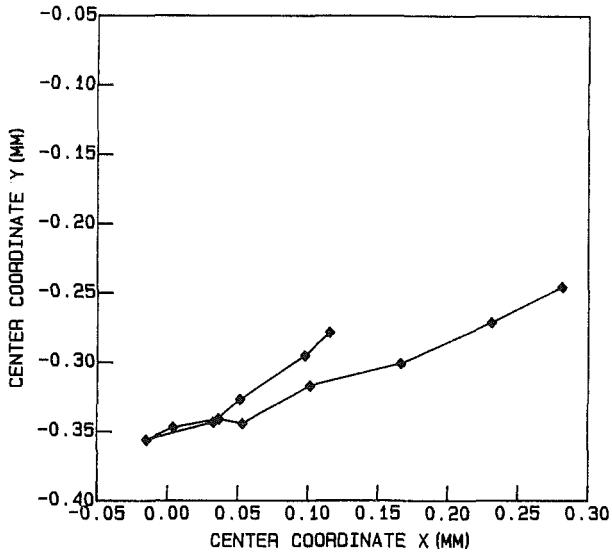


Fig. 8 Center locus during loading

pitch and constant radius by Phillips-Costello made the two partial differential equations completely linear.

The friction torque and wind-up moment produced by the loading is shown in Fig. 6. It is clearly seen that the friction torque is sufficient to stop the axial end rotation caused by wind-up moment for coefficients of coulomb friction as low as 0.1, and thereby justifies the assumption that the ends of spring are held fixed against rotation.

The radial-expansion experiment was performed on a milling machine equipped with index head and a specially-designed fixture. The radial position of eight points around one turn of the spring in the middle section were measured for every two millimeters of spring compression. The resulting data were processed by a nonlinear least-squares curve fit to determine the center coordinates and the radius of this coil. The results are shown in Fig. 7.

It was found that spring coils experienced significant lateral deflections in random directions even for static loads that do not instigate buckling. This phenomenon is expected to be accentuated during the dynamic process. Since the magnitude of the center movement is larger than the radial expansion, it will make more sense to study the center motion whenever the radial expansion needs be taken into account. Figure 8 is the locus of the center of the measured turn of spring taken from one experiment during static compression.

Conclusions

The general dynamic equations of the helical compression spring are derived, including the effects of coil flexure, variable pitch angle, and variable helix radius. Although the size of the general dynamic equations is formidable, they can be reduced to simpler forms in most application situations. The equations can also be simplified by replacing coefficients $\sin[p(s,t)]$ and $\cos[p(s,t)]$ in the equations by a Taylor series expansion in pitch angle $p(s,t)$. To observe nonlinear effects, at least second order terms must be retained. A static loading

experiment verifies the improvements over existing theories and formulas in predicting the nonlinear force-displacement relation and compression-radial expansion. Current research on the full dynamic solution of these equations is in progress.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grants No. MEA-8404822 and DMC-8451199 (a Presidential Young Investigator Award). Research support provided by 2.5 Liter Staff, Chevrolet-Pontiac-Canada Group, General Motor Corporation is gratefully acknowledged. The authors also wish to acknowledge the support of the Department of Mechanical Engineering, University of California, Berkeley, for providing computer and laboratory resources.

References

Ancker, C. J., Jr., and Goodier, J. N., 1958, "Pitch and Curvature Corrections for Helical Springs," ASME JOURNAL OF APPLIED MECHANICS, Vol. 25, pp. 466–495.

Chen, W. H., and Tsai, P., 1983, "On Static and Dynamic Finite Element Analysis of Helical Springs and Experimental Verifications," *Chung Kuo Kang Ch'eng Hsueh K'an*, Vol. 6, No. 1, in Chinese.

Do Carmo, M. P., 1976, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, pp. 19–22.

Donkin, W. T., and Clark, H. H., 1929, "The Electric Telemeter and Valve-Spring Surge," *Trans. SAE*, Vol. 24, pp. 185–196.

Costello, G. A., 1975, "Radial Expansion of Impacted Helical Springs," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 789–792.

Jehle, F., and Spiller, W. R., 1929, "Idiosyncrasies of Valve Mechanisms and Their Causes," *Trans. SAE*, Vol. 24, pp. 197–212.

Kato, Michio, 1974, "Study on Dynamic Properties of a Varying-Pitch Helical Compression Spring," *Bulletin of the JSME*, Vol. 17, No. 110, pp. 1015–1022.

Lehr, E., 1933, "Schwingungen in Ventilfedern," *VDI*, Vol. 77, No. 18, pp. 457–462.

Leitmann, G., 1963, "Some Remarks on Hamilton's Principle," ASME JOURNAL OF APPLIED MECHANICS, Vol. 30, pp. 623–625.

Love, A. E. H., 1927, *A Treaties on the Mathematical Theory of Elasticity*, 4th Ed., Dover, New York.

Phillips, J. W., and Costello, G. A., 1972, "Large Deflections of Impacted Helical Springs," *Journal of the Acoustical Society of America*, Vol. 51, No. 3, pp. 967–973.

Pisano, A. P., and Freudenstein, F., 1983, "An Experimental and Analytical Investigation of the Dynamic Response of a High-Speed Cam Follower System; Part 2: A Combined, Lumped/Distributed Parameter Dynamic Model," *ASME Journal of Mechanisms, Transmissions and Automation in Design*, Vol. 105, pp. 699–704.

Rosenberg, R. M., 1977, *Analytical Dynamics of Discrete Systems*, Plenum Press, pp. 239–247.

Sayre, M. F., and De Forest, A. V., 1936, "New Spring Formulas and New Materials for Precision Spring Scales," *Transactions of ASME*, Vol. 58, pp. 379–387.

Stokes, V. K., 1974, "On the Dynamic Radial Expansion of Helical Springs due to Longitudinal Impact," *Journal of Sound and Vibration*, Vol. 35, pp. 77–99.

Thomson, W., and Tait, P. G., 1883, *Treaties on Natural Philosophy*, 2nd Ed., Oxford, pp. 136–144.

Wahl, A. M., 1963, *Mechanical Springs*, Penton Publishing Co., Cleveland, Ohio.

Whittaker, E. T., 1944, *Analytical Dynamics of Particles and Rigid Bodies*, Dover Publications, New York, pp. 249–252.

Wittrick, W. H., 1966, "On Elastic Wave Propagation in Helical Springs," *International Journal of Mechanical Science*, Vol. 8, pp. 25–47.

APPENDIX A

Variational Formulation of General Dynamic Equations

To formulate the dynamic equations by Hamilton's principle, energy terms are processed one by one as follows:

$$\int_{t_1}^{t_2} \delta T_1 dt = \int_{t_1}^{t_2} \delta \int_0^L \frac{1}{2} I_m \dot{\psi}^2 ds dt$$

Exchanging integration order, and integrating by parts, one can obtain:

$$\int_{t_1}^{t_2} \delta T_1 = \int_0^L (I_m \dot{\psi} \delta\psi)_{t_1}^{t_2} - \int_{t_1}^{t_2} I_m \ddot{\psi} \delta\psi \, dt ds$$

$$= \int_{t_1}^{t_2} \int_0^L -\frac{md^2}{8} \ddot{\psi} \delta\psi \, ds dt \qquad (A1)$$

where $I_m = md^2/8$, $\delta\psi = 0$ at $t = t_1$ and $t = t_2$. Similarly, for the remainder of the kinetic and potential energy terms:

$$\int_{t_1}^{t_2} \delta T_2 dt = \int_{t_1}^{t_2} \delta \int_0^L \frac{1}{2} m \dot{r}^2 ds dt$$

$$= \int_{t_1}^{t_2} \int_0^L -m\ddot{r} \, \delta r \, ds \, dt \qquad (A2)$$

$$\int_{t_1}^{t_2} \delta T_3 dt = \int_{t_1}^{t_2} \int_0^L -m\ddot{y} \, \delta y \, ds \, dt \qquad (A3)$$

$$\int_{t_1}^{t_2} \delta U_1 dt = \int_{t_1}^{t_2} \delta \int_0^L \frac{1}{2} GJ \left(\frac{\partial^2 \psi}{\partial s^2}\right)^2 ds \, dt$$

$$= \frac{EI}{1+\nu} \int_{t_1}^{t_2} \left(\frac{\partial \psi}{\partial s} \delta\psi\right) \Big|_0^L - \int_0^L \frac{\partial^2 \psi}{\partial s^2} \delta\psi \, ds \, dt$$

$$= \int_{t_1}^{t_2} \int_0^L -\frac{EI}{1+\nu} \frac{\partial^2 \psi}{\partial s^2} \delta\psi \, ds \, dt \qquad (A4)$$

$$\int_{t_1}^{t_2} \delta U_2 dt = \int_{t_1}^{t_2} \delta \int_0^L \frac{1}{2} EI(\kappa - \kappa_0)^2 ds \, dt$$

$$= \int_{t_1}^{t_2} \int_0^L EI(\kappa - \kappa_0) \delta \left(\frac{\cos^2 p}{r} - \frac{\cos^2 p_0}{r_0}\right) ds \, dt$$

$$= \int_{t_1}^{t_2} \int_0^L EI(\kappa - \kappa_0) \left(\frac{-\sin 2p}{r} \delta p - \frac{\cos^2 p}{r^2} \delta r\right) ds \, dt \qquad (A5)$$

$$\int_{t_1}^{t_2} \delta U_3 dt = \int_{t_1}^{t_2} \delta \int_0^L \frac{1}{2} EI \left(\frac{\partial^2 y}{\partial s^2}\right)^2 ds \, dt$$

$$= EI \int_{t_1}^{t_2} \left(y'' \delta \frac{\partial y}{\partial s}\right) \Big|_0^L - \int_0^L y''' \delta \left(\frac{\partial y}{\partial s}\right) ds \, dt$$

$$= \int_{t_1}^{t_2} \int_0^L EI \frac{\partial^4 y}{\partial s^4} \delta y \, ds \, dt \qquad (A6)$$

The vanishing of first term in second step occurs because

$$\delta p = 0$$

$$\frac{\partial y}{\partial s} = \sin p, \quad \text{for } s = 0 \text{ and } s = L \qquad (A7)$$

Nonholonomic constraint equations (3) and (5) can be written for virtual displacements as:

$$\delta\psi + \int_0^s \frac{\sin p(\zeta,t)\ \cos p(\zeta,t)}{[r(\zeta,t)]^2}\, d\zeta\delta r - \int_0^s \frac{\cos 2p(\zeta,t)}{r(\zeta,t)}\, d\zeta\delta p = 0 \tag{A8}$$

$$\delta y - \int_0^s \cos[p(\zeta,t)]d\zeta\ \delta p = 0 \tag{A9}$$

Then equations (14)–(17) are obtained by combining terms involving δy, δp, δr, and $\delta\psi$. Coefficients of virtual displacements $\delta\psi$, δr, and δp in equation (A7) were multiplied by λ_1, δy, and δp in equation (A8) multiplied by λ_2 and adjointed to equations, just as if applying generalized forces to the equations.

APPENDIX B

Derivation of Constraints Between Radius and Pitch Angle

Referring to Fig. 2, a relationship between radius $r(s,t)$ and pitch angle $p(s,t)$ can be derived as follows. Since,

$$r_0(s)d\theta = \cos p_0(s)ds \tag{B1}$$

$$r(s,t)d\theta = \cos p(s,t)ds \tag{B2}$$

and assuming that $d\theta$ and ds are constants during motion, dividing equation (B1) from equation (B2) yields:

$$r(s,t) = \frac{r_0(s)\ \cos [p(s,t)]}{\cos [p_0(s)]} \tag{B3}$$

Experimental data resulted from applying this constraint to simplification of dynamic equations confirms that above assumptions are reasonable.

S. F. Masri

R. K. Miller

A. F. Saud

Department of Civil Engineering,
University of Southern California,
Los Angeles, CA 90089-0242

T. K. Caughey

Division of Engineering and Applied Science,
California Institute of Technology,
Pasadena, CA 91106

Identification of Nonlinear Vibrating Structures: Part I—Formulation

A self-starting multistage, time-domain procedure is presented for the identification of nonlinear, multi-degree-of-freedom systems undergoing free oscillations or subjected to arbitrary direct force excitations and/or nonuniform support motions. Recursive least-squares parameter estimation methods combined with nonparametric identification techniques are used to represent, with sufficient accuracy, the identified system in a form that allows the convenient prediction of its transient response under excitations that differ from the test signals. The utility of this procedure is demonstrated in a companion paper.

1 Introduction

1.1 Background. The identification and modeling of nonlinear multidegree-of-freedom (MDOF) dynamic systems through the use of experimental data is a problem of considerable importance in the applied mechanics area. Since the model structure in many practical dynamics problems is by no means clear, an increasing amount of attention has recently been devoted to nonparametric identification methods.

One rather general nonparametric nonlinear identification approach is based on the expansion of the nonlinear restoring force functions in a power series or generalized Fourier series involving orthogonal polynomial functions. In applications, it is generally assumed that such series are rapidly convergent so that only a few terms need be retained for identification purposes. In such an approach, the coefficients of the retained terms from the series become parameters of the system which may be identified by many well-known techniques, such as least-squares fit in the time domain.

The origins of this basic approach are very classical and diverse, with roots in the theory of analytic functions and in the theory of Fourier series, and with applications in many engineering disciplines as well as operations research, economics, and the physical sciences. With regard to the engineering literature, the basic approach is outlined in the book by Graupe (1976). Applications of the method in the mechanical sciences appear to have originated in the early 1950's in several NACA technical notes (Greenberg, 1951; Shinbrot, 1951; Shinbrot, 1952; Briggs and Jones, 1953; and Shinbrot, 1954) and in the papers by Klotter (1953) and Shinbrot (1957). In the following years, interest in similar time series methods for nonlinear system identification of structures expanded, as attested to by the representative publications of Kohr (1963), Hoberock and Kohr (1967), Sprague and Kohr (1969), Sehitoglu and Klein (1975), Masri et al. (1982), Natke (1982), Masri et al. (1984), Tomlinson (1985), and Hac and Spanos (1987).

Most of the research in this area has been concerned with SDOF systems with nonlinearities of varying complexity. The basic identification method becomes generally impractical for complex MDOF systems due to excessive computation and computer memory requirements caused by slow convergence of the series expansions. However, Masri et al. (1982) demonstrated by example that rapid series convergence (and hence practical identification results) may be obtained in at least some MDOF structural applications by basing the identification procedure on a set of generalized coordinates corresponding to the mode shapes of a comparison linear structural system.

In the paper by Masri et al. (1982), certain restrictions were made on the class of nonlinear structural systems to be identified. In particular, it was assumed that (1) the system mass matrix M is diagonal and known; (2) the equivalent linear system stiffness matrix K is symmetric and known; and (3) the excitation to the system is furnished through forces directly applied to the discrete mass locations. The requirement of knowing the linearized system parameters pertaining to M and K, as well as the exclusion of the class of problems involving support motion (such as in the case of earthquake ground motion), limited the utility of the approach in practical cases.

The present paper further extends the above-referenced work by generalizing the approach to handle, approximately, the case of arbitrary nonlinear MDOF dynamic systems with multiple inputs and outputs under the action of force excitations and/or nonuniform support motion. The method is based on the use of time-domain estimation techniques to identify the parameters of an equivalent linear model whose eigenvectors are then used to estimate the "modes" of the nonlinear system. Regression techniques involving the use of two-dimensional orthogonal functions are then employed to develop an approximate expression for the system generalized

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 13, 1987; final revision June 23, 1987.

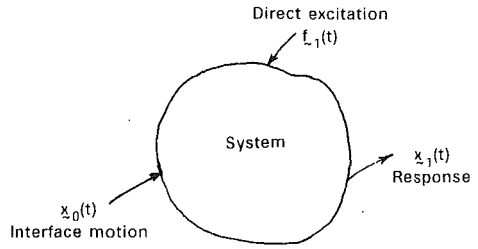


Fig. 1 Model of system

restoring forces in terms of the corresponding generalized system state variables.

Section 2 of this paper extends the work of previous investigators by presenting a unified approach for handling the time-domain identification of the system matrices associated with a variety of classes of linear problems arising in the field of structural dynamics. The formulation under discussion includes the cases of free vibrations as well as direct force and/or independent support motion.

Section 3 incorporates the results of Section 2 in the identification of nonlinear vibrating structures. The "calibration" of this approach is accomplished in the companion paper (Masri et al., 1987) by applying the method under discussion to a representative multi-input/multi-output nonlinear system incorporating polynomial as well as hysteretic nonlinearities.

1.2 Formulation of Time Domain Identification Procedure. Consider a discrete multi-degree-of-freedom (MDOF) system of the type shown in Fig. 1, which is subjected to directly applied excitation forces $\mathbf{f}_1(t)$ as well as prescribed support motions $\mathbf{x}_0(t)$. The motion of this multi-input/multi-output nonlinear system is governed by the set of equations

$$\mathbf{f}_T(\ddot{\mathbf{x}},\dot{\mathbf{x}},\mathbf{x}) = \mathbf{f}_1(t) \tag{1}$$

where:

$\mathbf{f}_T$ = an n_1 column vector representing the total sum of all the inertia and restoring forces acting on the system,

$\mathbf{f}_1(t)$ = an n_1 column vector of directly applied forces,

$\mathbf{x}(t)$ = $(\mathbf{x}_1(t), \mathbf{x}_0(t))^T$ = system displacement vector of order $(n_1 + n_0)$,

$\mathbf{x}_1(t)$ = active degree-of-freedom (DOF) displacement vector of order n_1,

$\mathbf{x}_0(t)$ = prescribed support displacement vector of order n_0.

Let $\mathbf{f}_T(t)$ be expressed as

$$\mathbf{f}_T(t) = {}^L\mathbf{f}_1(t) + {}^S\mathbf{f}_1(t) + {}^L\mathbf{f}_0(t) + {}^S\mathbf{f}_0(t) + \mathbf{f}_N(t), \tag{2}$$

where

$$ {}^L\mathbf{f}_1(t) = M_{11}\ddot{\mathbf{x}}_1(t) + C_{11}\dot{\mathbf{x}}_1(t) + K_{11}\mathbf{x}_1(t), \tag{3}$$

$$ {}^S\mathbf{f}_1(t) = M^S_{11}\ddot{\mathbf{x}}_1(t) + C^S_{11}\dot{\mathbf{x}}_1(t) + K^S_{11}\mathbf{x}_1(t), \tag{4}$$

$$ {}^L\mathbf{f}_0(t) = M_{10}\ddot{\mathbf{x}}_0(t) + C_{10}\dot{\mathbf{x}}_0(t) + K_{10}\mathbf{x}_0(t), \tag{5}$$

$$ {}^S\mathbf{f}_0(t) = M^S_{10}\ddot{\mathbf{x}}_0(t) + C^S_{10}\dot{\mathbf{x}}_0(t) + K^S_{10}\mathbf{x}_0(t), \tag{6}$$

M_{11}, C_{11}, K_{11} = constant matrices that characterize the inertia, damping, and stiffness forces associated with the unconstrained DOF of the system, each of order $n_1 \times n_1$,

$M^S_{11}, C^S_{11}, K^S_{11}$ = response-dependent matrices that characterize the inertia, damping, and stiffness forces associated with the unconstrained DOF of the system, each of order $n_1 \times n_1$,

M_{10}, C_{10}, K_{10} = constant matrices that characterize the inertia, damping, and stiffness forces associated with the support motions, each of order $n_1 \times n_0$,

$M^S_{10}, C^S_{10}, K^S_{10}$ = response-dependent matrices that characterize the inertia, damping, and stiffness forces associated with the support motions, each of order $n_1 \times n_0$,

${}^L\mathbf{f}_1(t)$ = an n_1 column vector of linear forces involving $\mathbf{x}_1(t)$,

${}^S\mathbf{f}_1(t)$ = an n_1 column vector of response-dependent forces involving $\mathbf{x}_1(t)$,

${}^L\mathbf{f}_0(t)$ = an n_1 column vector of linear forces involving $\mathbf{x}_0(t)$,

${}^S\mathbf{f}_0(t)$ = an n_1 column vector of response-dependent forces involving $\mathbf{x}_1(t)$ as well as $\mathbf{x}_0(t)$,

$\mathbf{f}_N(t)$ = an n_1 column vector of nonlinear nonconservative forces involving $\mathbf{x}_1(t)$ as well as $\mathbf{x}_0(t)$.

Making use of equation (2), the system equation of motion (1) can be expressed as

$$M^e_{11}\ddot{\mathbf{x}}_1(t) + C^e_{11}\dot{\mathbf{x}}_1(t) + K^e_{11}\mathbf{x}_1(t) + M^e_{10}\ddot{\mathbf{x}}_0(t) + C^e_{10}\dot{\mathbf{x}}_0(t) + K^e_{10}\mathbf{x}_0(t) + \mathbf{f}_N(t) = \mathbf{f}_1(t), \tag{7}$$

where:

$$\begin{aligned} M^e_{11} &= M_{11} + M^S_{11}, & M^e_{10} &= M_{10} + M^S_{10}, \\ C^e_{11} &= C_{11} + C^S_{11}, & C^e_{10} &= C_{10} + C^S_{10}, \\ K^e_{11} &= K_{11} + K^S_{11}, & K^e_{10} &= K_{10} + K^S_{10}, \end{aligned} \tag{8}$$

This study is concerned with a time-domain method for the identification of the system matrices appearing in equation (7) as well as the nonlinear forces acting on the system. The representation of the identifed system will be in a form that allows the prediction of its transient response under arbitrary excitations, by using conventional numerical techniques for initial-value problems in ordinary differential equations.

Note that equation (7) can be expressed as

$$\ddot{\mathbf{x}}_1(t) = [M^e_{11}]^{-1}(\mathbf{f}_1(t) - \mathbf{f}^e_L(t) - \mathbf{f}_N(t)), \tag{9}$$

where:

$$\mathbf{f}^e_L(t) = \mathbf{b}_1(t) + \mathbf{b}_0(t), \tag{10}$$

$$\mathbf{b}_1(t) = C^e_{11}\dot{\mathbf{x}}_1(t) + K^e_{11}\mathbf{x}_1(t), \tag{11}$$

$$\mathbf{b}_0(t) = M^e_{10}\ddot{\mathbf{x}}_0(t) + C^e_{10}\dot{\mathbf{x}}_0(t) + K^e_{10}\mathbf{x}_0(t). \tag{12}$$

Thus, by introducing the state vector $\mathbf{y}$ of order $2n_1$ where

$$y_{2i-1} = x_{1_i} \tag{13}$$

$$y_{2i} = \dot{x}_{1_i} \qquad i = 1, 2, \ldots, n_1 \tag{14}$$

standard time-marching techniques can be used to solve

$$\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y}, \mathbf{f}_1, \mathbf{x}_0). \tag{15}$$

2 Time-Domain Identification of Linear System Matrices

The use of least-squares methods to estimate unknown parameters is a well known and developed approach which occupies significant portions of numerous books devoted to the subject of parameter estimation, particularly in the field of electrical engineering control and system theory (Mendel, 1973; Graupe, 1976; Hsia, 1977; Sorenson, 1980). While this approach has also been frequently applied in the field of structural dynamics (Caravani and Thomson, 1974, 1977; Ibrahim and Mikulcik, 1973, 1976, 1977; Ibrahim and Pappa, 1982; Ibrahim, 1977, 1978, 1983; Junkins, 1978; Beck and Jennings, 1980; Yao, 1985; Torkamani and Hart, 1975; Shinozuka et al., 1982; Rajaram and Junkins, 1985; Hac and Spanos, 1987), there is a paucity of studies that are concerned with the problems encountered by this approach when applied to realistic problems arising in the vibration field. Consequently, the present section of this paper is devoted to presenting an in-depth, unified, and efficient approach for using least-squares

parameter estimation methods to identify the needed system matrices associated with a wide variety of realistic situations commonly encountered when dealing with experimental measurements of vibrating structures.

2.1 Formulation. Consider a linearized version of the system shown in Fig. 1, and assume it is governed by

$$M_{11}\ddot{\mathbf{x}}_1(t)+C_{11}\dot{\mathbf{x}}_1(t)+K_{11}\mathbf{x}_1(t)+ M_{10}\ddot{\mathbf{x}}_0(t)+C_{10}\dot{\mathbf{x}}_0(t)+K_{10}\mathbf{x}_0(t)=\mathbf{f}_1(t). \quad (16)$$

Let the response vector $\mathbf{r}(t)$ of order $3(n_1 + n_0)$ be defined as

$$\mathbf{r}(t)=(\ddot{\mathbf{x}}_1^T(t),\dot{\mathbf{x}}_1^T(t),\mathbf{x}_1^T(t),\ddot{\mathbf{x}}_0^T(t),\dot{\mathbf{x}}_0^T(t),\mathbf{x}_0^T(t))^T. \quad (17)$$

For clarity of presentation, let the six matrices appearing in equation (16) be denoted by ${}^1A, {}^2A, \ldots, {}^6A$, respectively.

Let $\langle {}^jA_i\rangle$ = ith row of a generic matrix ${}^jA$, and introduce the parameter vector $\boldsymbol{\alpha}_i$.

$$\boldsymbol{\alpha}_i=(\langle {}^1A_i\rangle,\langle {}^2A_i\rangle,\langle {}^3A_i\rangle,\langle {}^4A_i\rangle,\langle {}^5A_i\rangle,\langle {}^6A_i\rangle)^T. \quad (18)$$

Suppose that the excitation and the response of the system governed by equation (16) is measured at times $t_1, t_2, \ldots, t_N$. Then at every t_k,

$${}^1A\ddot{\mathbf{x}}_1(t_k)+{}^2A\dot{\mathbf{x}}_1(t_k)+{}^3A\mathbf{x}_1(t_k)+{}^4A\ddot{\mathbf{x}}_0(t_k)+ {}^5A\dot{\mathbf{x}}_0(t_k)+{}^6A\mathbf{x}_0(t_k)=\mathbf{f}_1(t_k);\quad k=1,2,\ldots,N. \quad (19)$$

Introducing matrix R

$$R=\begin{bmatrix}\mathbf{r}^T(t_1)\\ \mathbf{r}^T(t_2)\\ \vdots\\ \mathbf{r}^T(t_N)\end{bmatrix} \quad (20)$$

and using the notation above, the grouping of the measurements can be expressed concisely as

$$\hat{R}\hat{\boldsymbol{\alpha}}=\hat{\mathbf{b}} \quad (21)$$

where $\hat{R}$ is a block-diagonal matrix whose diagonal elements are equal to R, $\hat{\boldsymbol{\alpha}} = (\boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_2^T, \ldots, \boldsymbol{\alpha}_{n_1}^T)^T$, and $\hat{\mathbf{b}}$ is the corresponding vector of excitation measurements.

Keeping in mind that $\hat{R}$ is of order $m \times n$ where $m = Nn_1$, and $n = 3n_1\ (n_1 + n_0)$, then if a sufficient number of measurements is taken, this will result in $m > n$. Under these conditions, least-squares procedures can be used to solve for all the system parameters that constitute the entries in $\hat{\boldsymbol{\alpha}}$:

$$\hat{\boldsymbol{\alpha}}=\hat{R}^{\dagger}\hat{\mathbf{b}} \quad (22)$$

where $\hat{R}^{\dagger}$ is the pseudoinverse of $\hat{R}$ (Golub and Van Loan, 1983).

Using the weighted least-squares approximation to minimize the cost function, J, results in the approximate solution

$$\hat{\boldsymbol{\alpha}}=(\hat{R}^T W\hat{R})^{-1}\hat{R}^T W\hat{\mathbf{b}}, \quad (23)$$

where W is the error weighting matrix.

2.2 Computational Efficiency.

2.2.1 Decoupling. One way to reduce the order of equation (21) to a manageable level is by making use of the diagonal nature of partitioned matrix $\hat{R}$, thus resulting in a set of n_1 decoupled matrix equations each of the form

$$R\boldsymbol{\alpha}_i=\mathbf{b}_i,\quad i=1,2,\ldots,n_1. \quad (24)$$

Comparing the order of R in equation (24) with that of $\hat{R}$ in equation (21), shows that the order of R is smaller by a factor of n_1^2 compared to $\hat{R}$. Least-squares techniques can again be used to obtain the components of the n_1 parameter vectors $\boldsymbol{\alpha}_i$:

$$\hat{\boldsymbol{\alpha}}_i=R^{\dagger}\hat{\mathbf{b}}_i;\quad i=1,2,\ldots,n_1. \quad (25)$$

Note that $R^{\dagger}$ needs to be computed only *once*.

While the formulation in equation (24) is obviously superior to the corresponding formulation in equation (21), the former suffers from a significant (practical) limitation pertaining to the number of system DOFs simultaneously excited.

2.2.2 Recursive Solutions. Suppose that a set of m equations

$$\hat{R}_k\hat{\boldsymbol{\alpha}}=\hat{\mathbf{b}}^{(k)} \quad (26)$$

has been used to obtain a weighted least-squares estimate for $\hat{\boldsymbol{\alpha}}$, denoted by $\hat{\boldsymbol{\alpha}}^{(k)}$:

$$\hat{\boldsymbol{\alpha}}^{(k)}=(\hat{R}_k^T W_k\hat{R}_k)^{-1}\hat{R}_k^T W_k\hat{\mathbf{b}}^{(k)}. \quad (27)$$

Using an additional set of relations

$$\hat{R}_{(k+1)}\hat{\boldsymbol{\alpha}}=\hat{\mathbf{b}}^{(k+1)} \quad (28)$$

a new estimate of $\hat{\boldsymbol{\alpha}}$, denoted by $\hat{\boldsymbol{\alpha}}^{(k+1)}$, can be obtained without reprocessing the whole set of equations involving $(\hat{\mathbf{b}}^{(k)}, \hat{\mathbf{b}}^{(k+1)})$ (Brogan, 1985).

2.3 Special Cases. In the work of Masri et al. (1987a), special cases that influence the application of the method in practical situations are discussed and steps are provided for alleviating some of the problems appearing in realistic cases. Among these topics are the uniqueness issues, partial knowledge of system parameters, conditions under which the approach fails to yield desired results, symmetry assumptions, and recursive approaches to enhance computational efficiency.

3 Identification of Nonlinear Systems

Consider the nonlinear system governed by equation (1) and assume that the identification procedure discussed in Section 2 has yielded the system matrices needed to determine the equivalent linear internal force vector $\mathbf{f}_L^e(t)$ appearing in equation (9) and defined by equation (10).

3.1 Restoring Force Estimation. Solving equation (7) for the nonlinear force vector $\mathbf{f}_N(t)$ results in

$$\mathbf{f}_N(t)=\mathbf{f}_1(t)-(M_{11}^e\ddot{\mathbf{x}}_1(t)+\mathbf{f}_L^e(t)). \quad (29)$$

Since all the terms appearing on the right-hand side of equation (29) are available from measurements or have been previously identified, the time history of $\mathbf{f}_N$ can be determined. Note from equation (29) that $\mathbf{f}_N(t)$ can be interpreted as the residual force vector corresponding to the difference between the excitation vector $\mathbf{f}_1(t)$ and the equivalent linear force vector composed of the inertia, damping, and the stiffness terms.

An alternative form of equation (29) is

$$\mathbf{f}_R(t)\equiv\mathbf{f}_N(t)+\mathbf{f}_L^e(t)=\mathbf{f}_1(t)-M_{11}^e\ddot{\mathbf{x}}_1(t), \quad (30)$$

where $\mathbf{f}_R(t)$ represents the difference between the excitation and equivalent linear inertia forces associated with the active degrees of freedom. The force $\mathbf{f}_R$ can be thought of as the "restoring force" of the system.

Let $h_i(t)$ represent either the ith component of the nonlinear residual force vector $\mathbf{f}_N(t)$ as defined by equation (29) or the restoring force vector $\mathbf{f}_R(t)$ as defined by equation (30). In general, vector $\mathbf{h}$ depends simultaneously on all the components of the system acceleration, velocity, and displacement vectors associated with the n_1 active DOF as well as the n_0 support components:

$$\mathbf{h}(t)=\mathbf{h}(\mathbf{x},\dot{\mathbf{x}},\ddot{\mathbf{x}}). \quad (31)$$

The central idea of the present method is that, in the case of nonlinear dynamic systems commonly encountered in the applied mechanics field, a judicious assumption is that each component of $\mathbf{h}$ can be expressed in terms of a series of the form:

$$h_i(\mathbf{x},\dot{\mathbf{x}},\ddot{\mathbf{x}})\approx\sum_{j=1}^{J_{\max_i}}\hat{h}_i^{(j)}(v_{1_i}^{(j)},v_{2_i}^{(j)}) \quad (32)$$

where the v_1's and v_2's are suitable generalized coordinates which, in turn, are linear combinations of the physical displacements, velocities, and accelerations. The approximation indicated in equation (32) is that each component h_i of the nonlinear force vector $\mathbf{h}$ can be adequately estimated by a collection of terms $\hat{h}_i^{(j)}$, each one of which involves a pair of generalized coordinates. The particular choice of combinations and permutations of the generalized coordinates and the number of terms $J_{\max_i}$ needed for a given h_i depends on the nature and extent of the nonlinearity of the system and its effect on the specific DOF i.

3.2 Eigenvector Expansion. If $h_i(t)$ is chosen as the ith component of $\mathbf{f}_N(t)$, then the procedure expressed by equation (32) will directly estimate the corresponding component of the unknown nonlinear force. For certain structural configurations (e.g., localized nonlinearities) and/or relatively low-order systems, the choice of suitable generalized coordinates for the series in equation (32) is a relatively straightforward task. However, in many practical cases involving distributed nonlinearities coupled with a relatively high-order system, an improved rate of convergence of the series in equation (32) can be achieved by performing the least-squares fit of the nonlinear forces in the "modal" domain as outlined below.

Using the identification results for the linear system discussed in Section 2, the eigenvalue problem associated with $M_{11}^{-1}K_{11}$ is solved resulting in the eigenvector or modal matrix Φ and the corresponding vector of generalized coordinates $\mathbf{u}$:

$$\mathbf{h}_1(\mathbf{u}, \dot{\mathbf{u}}) = \Phi^T \mathbf{f}_N(t) \tag{33}$$

with

$$\mathbf{u}(t) = \Phi^{-1}\mathbf{x}(t) \tag{34}$$

With this formulation in mind, equation (32) can be viewed as allowing for "modal" interaction between all generalized coordinates, taken two at a time. Note that the formulation in equation (32) allows for "modal" interaction between all "modal" displacements, velocities, and accelerations.

3.3 Series Expansion. The individual terms appearing in the series expansion of equation (32) may be evaluated by using the least-squares approach to determine the optimum fit for the time history of each h_i. Thus $\hat{h}_i^{(1)}$ may be expressed as a double series involving a suitable choice of generalized coordinates:

$$\hat{h}_i^{(1)}(v_{1_i}^{(1)}, v_{2_i}^{(1)}) \equiv \sum_k \sum_{\ell}{}^{(1)} C_{k\ell}^{(i)} T_k(v_{1_i}^{(1)}) T_\ell(v_{2_i}^{(1)}) \tag{35}$$

where the C_{kl}'s are a set of undetermined constants and $T_k(.)$ are suitable basis functions, such as orthogonal polynomials. Let $h_i^{(2)}$, the deviation (residual) error between h_i and its first estimate $\hat{h}_i^{(1)}$, be given by

$$h_i^{(2)}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}) = h_i(\mathbf{x}_1, \dot{\mathbf{x}}_1, \ddot{\mathbf{x}}_1) - \hat{h}_i^{(1)}(v_{1_i}^{(1)}, v_{2_i}^{(1)}). \tag{36}$$

Equation (32) accounts for the contribution to the nonlinear force h_i of generalized coordinates $v_{1_i}^{(1)}$ and $v_{2_i}^{(1)}$ appearing in the form $(v_{1_i}^{(1)})^k (v_{2_i}^{(1)})^\ell$. Consequently, the residual error as defined by equation (32) can be further reduced by fitting $h_i^{(2)}$ by a similar double series involving variables $v_{1_i}^{(2)}$ and $v_{2_i}^{(2)}$:

$$\mathbf{h}_i^{(2)}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}) \approx \hat{h}_i^{(2)}(v_{1_i}^{(2)}, v_{2_i}^{(2)}) \tag{37}$$

where

$$\hat{h}_i^{(2)}(v_{1_i}^{(2)}, v_{2_i}^{(2)}) \equiv \sum_k \sum_{\ell}{}^{(2)} C_{k\ell}^{(i)} T_k(v_{1_i}^{(2)}) T_\ell(v_{2_i}^{(2)}). \tag{38}$$

By extending this procedure to account for all DOFs that have significant interaction with DOF i, equation (32) is obtained with

$$h_i^{(j+1)}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}) = h_i^{(j)}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}) - \hat{h}_i^{(j)}(v_{1_i}^{(j)}, v_{2_i}^{(j)}); \quad j = 1, 2, \ldots, J_{\max_i} \tag{39}$$

where

$$h_i^{(1)}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}) \equiv h_i(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}), \tag{40}$$

and

$$\hat{h}_i^{(j)}(v_{1_i}^{(j)}, v_{2_i}^{(j)}) \equiv \sum_k \sum_{\ell}{}^{(j)} C_{k\ell}^{(i)} T_k(v_{1_i}^{(j)}) T_\ell(v_{2_i}^{(j)}). \tag{41}$$

Note that, in general, the range of the summation indices k and ℓ appearing in equation (41) may vary with the series index j and DOF index i. Similarly, $J_{\max_i}$, the total number of series terms needed to achieve a given level of accuracy in fitting the nonlinear force time history, depends on the DOF index i.

3.4 Least Squares Fit for Nonlinear Forces. Using two-dimensional orthogonal polynomials $T_k(.)$ to estimate each $h_i(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}})$ by a series of approximating functions $\hat{h}_i^{(j)}$ of the form indicated in equation (41), then the numerical value of the $C_{k\ell}$ coefficients can be determined by invoking the applicable orthogonality conditions for the chosen polynomials. While there is a wide choice of suitable basis functions for least-squares application, the orthogonal nature of the Chebyshev polynomials and their "equal-ripple" characteristics make them convenient to use in the present work.

Let each generalized coordinate v appearing in equation (32) be normalized to lie in the range -1 to 1:

$$v' = [v - (v_{\max} + v_{\min})/2]/[(v_{\max} - v_{\min})/2]. \tag{42}$$

If the Chebyshev polynomials, given by

$$T_n(\xi) = \cos(n \cos^{-1}\xi), \quad -1 < \xi < 1 \tag{43}$$

and satisfying the weighted orthogonality property

$$\int_{-1}^{1} w(x) T_n(\xi) T_m(\xi) d\xi = \begin{cases} 0, & n \neq m \\ \pi/2, & n = m \neq 0 \\ \pi, & n = m = 0 \end{cases} \tag{44}$$

where $w(x) = (1-x^2)^{-1/2}$ is the weighting function, are used, then the $C_{k\ell}$ coefficients would be given by

$$C_{k\ell} = \begin{cases} (2/\pi)^2 S_{k\ell}, & k \text{ and } \ell \neq 0 \\ (2/\pi^2) S_{k\ell}, & k \text{ or } \ell = 0 \\ (1/\pi^2) S_{k\ell}, & k \text{ and } \ell = 0 \end{cases} \tag{45}$$

where

$$S_{k\ell} = \int_0^\pi \int_0^\pi h(\cos^{-1} v_1, \cos^{-1} v_2) T_k(\theta_1) T_\ell(\theta_2) d\theta_1 d\theta_2 \tag{46}$$

and

$$v_i = \cos\theta_i, \quad i = 1, 2. \tag{47}$$

Note that in the special case when no cross-product terms are involved in any of the series terms, functions h can be expressed as the sum of two one-dimensional orthogonal polynomial series instead of a single two-dimensional series of the type under discussion.

4 Summary and Conclusions

An approximate time-domain method is presented for the identification of nonlinear multi-degree-of-freedom systems subjected to arbitrary direct force excitations and/or not-necessarily-identical support motions. This self-starting method uses recursive least-squares parameter estimation methods, combined wtih nonparametric identification techniques, to generate a reduced-order nonlinear mathematical model suitable for use in subsequent studies to predict, with

good fidelity, the response of the test article under arbitrary dynamic excitations. The utility of this procedure is demonstrated in a companion paper.

Acknowledgment

This study was supported in part by a grant from the National Science Foundation and the National Institute of Health Biomedical Simulations Resource at the University of Southern California.

References

Beck, J. L., and Jennings, P. C., 1980, "Structural Identification Using Linear Models and Earthquake Records," *International Journal of Earthquake Engineering and Structural Dynamics*, Vol. 8, April, pp. 145–160.

Briggs, B. R., and Jones, A. L., 1953, "Techniques for Calculating Parameters of Nonlinear Dynamic Systems from Response Data," NACA TN 2977.

Brogan, W. L., 1985, *Modern Control Theory*, Second Ed., Quantum Publishers Inc., New York.

Caravani, P., and Thomson, W. T., 1974, "Identification of Damping Coefficients in Multidimensional Linear System" ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, No. 2, June, pp. 379–382.

Caravani, P., Watson, M. L. and Thomson, W. T., 1977, "Recursive Least-Squares Time Domain Identification of Structural Parameters," ASME JOURNAL OF APPLIED MECHANICS, Vol. 99, Series E, Mar., pp. 135–140.

Golub, G. H., and Van Loan, C., 1983, *Matrix Computations*, The John Hopkins University Press, Baltimore, MD.

Graupe, D., 1976, *Identification of Systems*, Robert E. Krieger Publishing Co., Huntington, NY.

Greenberg, H., 1951, "A Survey of Methods for Determining Stability Parameters of an Airplane from Dynamic Flight Measurements," NACA TN 2340.

Hac, A., and Spanos, P. D., 1987, "Time Domain Method for Parameter Identification of Structures," *Engineering Mechanics—6th Conference Abstracts*, Proc. ASCE, 20–22 May, p. 323.

Hoberock, L. L., Kohr, R. H., 1967, "An Experimental Determination of Differential Equations to Describe Simple Nonlinear Systems," ASME *Journal of Basic Engineering*, June, pp. 393–398.

Hsia, T. C., 1977, *System Identification: Least Squares Methods*, D. C. Heath and Co., Lexington, MA.

Ibrahim, S. R., and Mikulcik, E. C., 1973, "A Time Domain Modal Vibration Test Technique," *The Shock and Vibration Bulletin*, Bulletin 43, June, pp. 21–37.

Ibrahim, S. R., and Mikulcik, E. C., 1976, "The Experimental Determination of Vibration Parameters from Time Responses," *The Shock and Vibration Bulletin*, Bulletin 46, Aug., pp. 187–196.

Ibrahim, S. R., and Mikulcik, E. C., 1977, "A Method for the Direct Identification of Vibration Parameters from the Free Response," *The Shock and Vibration Bulletin*, Bulletin 47, Sept., pp. 183–198.

Ibrahim, S. R., 1977, "Random Decrement Technique for Modal Identification of Structures," *Journal of Spacecrafts and Rockets*, AIAA, Vol. 14, No. 11, Nov., pp. 696–700.

Ibrahim, S. R., 1978, "Modal Confidence Factor in Vibration Testing," *Journal of Spacecrafts and Rockets*, AIAA, Vol. 15, No. 11, Sept.–Oct., pp. 313–316.

Ibrahim, S. R., and Pappa, R. S., 1982, "Large Modal Survey Testing Using the Ibrahim Time Domain Identification Technique," *Journal of Spacecrafts and Rockets*, AIAA, Vol. 19, No. 5, Sept.–Oct., pp. 459–465.

Ibrahim, S. R., 1983, "Time Domain Parameter Identification and Modeling of Structures," *Proc. of The American Control Conference*, San Francisco, CA, June, pp. 989–996.

Junkins, J. L., 1978, *Optimal Estimation of Dynamical Systems*, Sijthoff-Noordhoff, Leyden, The Netherlands.

Klotter, K., 1953, "The Attenuation of Damped Free Vibrations and the Derivation of the Damping Law From Recorded Data," Stanford University Division of Engineering Mechanics, Contract N6-ONR-251, Technical Report 23, Nov. 1.

Kohr, R. H., 1963, "A Method for Determination of a Differential Equation Model for Simple Nonlinear Systems," *IEEE Transactions on Electronic Computers*, Vol. EC-12, No. 4, pp. 394–400.

Masri, S. F., Sassi, H., and Caughey, T. K., 1982, "Nonparametric Identification of Nearly Arbitrary Nonlinear Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, Sept., pp. 619–628.

Masri, S. F., Miller, R. K., Sassi, H., and Caughey, T. K., 1984, "A Method for Reducing the Order of Nonlinear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, USC Report USC-CE-8307, June, pp. 391–398.

Masri, S. F., Miller, R. K., Saud, A. F., and Caughey, T. K., 1987, "Identification of Nonlinear Vibrating Structures," University of Southern California Report No. USC-CE 8710, June.

Masri, S. F., Miller, R. K., Saud, A. F., and Caughey, T. K., 1987, "Identification of Nonlinear Vibrating Structures; Part II: Applications, ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, Dec.

Mendel, J. M., 1973, *Discrete Techniques of Parameter Estimation*, Marcel Dekker Inc., New York.

Natke, H. G., 1982, *Identification of Vibrating Structures*, Springer-Verlag, Wien, New York.

Rajaram, S., and Junkins, J. L., 1985, "Identification of Vibrating Flexible Structures," *AIAA Journal of Guidance Control and Dynamics*, Vol. 8, No. 4, July, pp. 463–470.

Sehitoglu, H., and Klein, R. E., 1957, "Identification of Multivalued and Memory Nonlinearities in Dynamic Process," *Simulation*, Sept., pp. 86–92.

Shinbrot, M., 1951, "A Least Squares Curve Fitting Method with Applications to the Calculation of Stability Coefficients from Transient Response Data," NACA TN 2341.

Shinbrot, M., 1952, "A Description and a Comparison of Certain Nonlinear Curve Fitting Techniques, With Applications to the Analysis of Transient-Response Data," NACA TN 2622.

Shinbrot, M., 1954, "On the Analysis of Linear and Nonlinear Dynamical Systems From Transient-Response Data," NACA TN 3288.

Shinbrot, M., 1957, "On the Analysis of Linear and Nonlinear Systems," Trans. ASME, Vol. 79, April, pp. 547–552.

Shinozuka, M., Yun, C.-B., and Imai, H., 1982, "Identification of Linear Structural Dynamic Systems," ASCE *Journal of Engineering Mechanics Division*, Vol. 108, No. 1, EMD, pp. 1371–1390.

Sorenson, H. W., 1980, *Parameter Estimation: Principles and Problems*, Marcel Dekker Inc., New York.

Sprague, C. H., and Kohr, R. H., 1969, "The Use of Piecewise Continuous Expansions in the Identification of Nonlinear Systems," *Journal of Basic Engineering*, ASME Transactions Series D, Vol. 91, pp. 173–178.

Tomlinson, G. R., 1985, "Dynamics of Nonlinear Systems," Report, Dept. of Mechanical Engineering, Heriot-Watt University, Edinburgh, U.K.

Torkamani, M., and Hart, G., 1975, "Building System Identification Using Earthquake Data," UCLA-ENG 7507, Jan.

Yao, J. T. P., 1985, *Safety and Reliability of Existing Structures*, Pitman Publishing Co., Boston, MA.

S. F. Masri

R. K. Miller

A. F. Saud

Department of Civil Engineering,
University of Southern California,
Los Angeles, CA 90089-0242

T. K. Caughey

Division of Engineering and Applied Science,
California Institute of Technology,
Pasadena, CA 91106

Identification of Nonlinear Vibrating Structures: Part II—Applications

A time-domain procedure for the identification of nonlinear vibrating structures, presented in a companion paper, is applied to a "calibration" problem which incorporates realistic test situations and nonlinear structural characteristics widely encountered in the applied mechanics field. The "data" set is analyzed to develop suitable, approximate nonlinear system representations. Subsequently, a "validation" test is conducted to demonstrate the range of validity of the method under discussion. It is shown that the procedure furnishes a convenient means for constructing reduced-order nonlinear nonparametric mathematical models of reasonably high fidelity in regard to reproducing the response of the test article under dynamic loads that differ from the identification test loads.

1 Introduction

1.1 Background. In the study by Masri et al. (1987b), henceforth referred to as the "companion paper," the authors presented the formulation of a time-domain method for the identification of arbitrarily nonlinear multi-degree-of-freedom (MDOF) vibrating systems undergoing free vibrations or subjected to direct force excitations and/or support motion that is not necessarily uniform. This paper applies the identification procedure in the cited reference to a "calibration" problem which incorporates realistic test situations and nonlinear characteristics. Subsequently, a "validation" problem is considered to invesitgate the range of validity of the identification/prediction procedure.

1.2 Scope. Section 2 of this paper defines the model configuration, the nonlinear (polynomial and hysteretic) element characteristics, and the exact system parameters corresponding to the "small oscillations" range.

Section 3 discusses a synthetic "experiment" meant to simulate a conventional "hammer-blow" test that is routinely used in contemporary experimental modal analysis procedures. After describing the probing signal characteristics and input/output measurements, the data processing approach under consideration is used to extract the linearized system inertia, damping and stiffness matrices.

Section 4 is concerned with a simulated forced vibration test wherein the excitation (stationary, wide-band random) is not directly applied to the system but rather to its moving support points. By using the parametric identification results for the linearized system parameters obtained in Section 3, the time histories of the nonlinear forces involving all system degrees of freedom are obtained. Using the eigenvectors associated with the linearized system as basis functions to transform the "measured" nonlinear forces, the generalized nonlinear forces and corresponding generalized state variables are obtained. Applying the nonparametric identification procedure under discussion, and approximating analytical function involving a series expansion in terms of a set of orthogonal polynomials is obtained and shown to yield a good estimate of the presumably unknown nonlinear restoring forces of the system.

Section 5 is concerned with the "validation" of the present identification procedure by using the identification results obtained in Section 4 to predict (estimate) the response of the "exact" nonlinear system, when the location of the disturbance as well as its form is different from what was used for the probing signal in the identification phase discussed in Section 4.

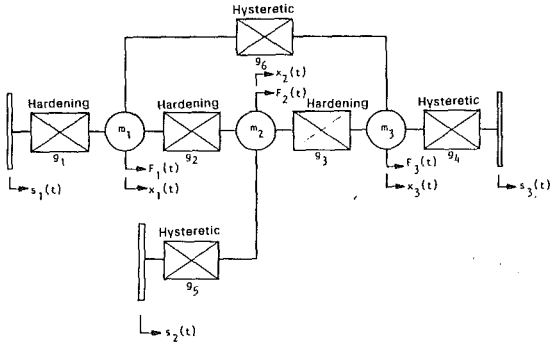
2 Model Characteristics

2.1 Example System Characteristics. To illustrate the application of the method under discussion, consider the hypothetical finite element model shown in Fig. 1. This one-dimensional (rectilinear horizontal motion) structure consists of three nearly equal masses m_i, $i=1,2,3$ that are interconnected by means of six truss elements anchored to an interface at three locations, thus resulting in a redundant system with three degrees of freedom.

The absolute displacement of each m_i is designated by x_i, while the prescribed motion of the three support points are designated by $s_i(t)$, $i=1,2,3$. The three excitation forces that directly act on the system components are denoted by $F_i(t)$, $i=1,2,3$. Thus, in terms of the notation introduced in the companion paper,

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 13, 1987; final revision June 23, 1987.

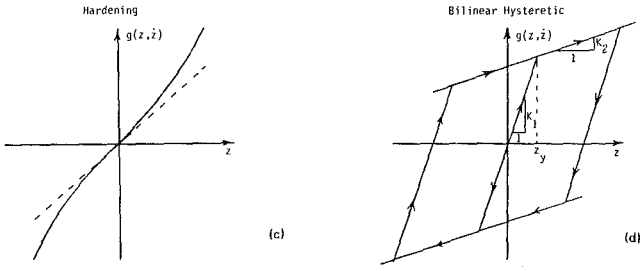


(a)

$m_1 = 0.8, \quad m_2 = 2.0, \quad m_3 = 1.2$

Element (i)	Type	$p_1^{(i)}$	$p_2^{(i)}$	$p_3^{(i)}$	$p_4^{(i)}$	$p_5^{(i)}$
1	Hardening	2.0	0.2	1.0		
2	Hardening	1.0	0.1	20.0		
3	Hardening	1.0	0.1	20.0		
4	Hysteretic	2.0	0.2	1.0	0.0	0.4
5	Hysteretic	2.0	0.2	1.0	0.0	0.4
6	Hysteretic	1.0	0.1	0.5	0.0	0.4

(b)



(c) (d)

Fig. 1 Model of example nonlinear 3DOF system: (a) configuration; (b) element characteristics; (c) generic element with polynomial nonlinearity; (d) generic element with bilinear hysteretic properties. $x_i(t)$ designates the absolute displacement of m_i, and $s_j(t)$ designates the absolute displacement time history associated with the support DOF$_j$.

(a)

$$M = \begin{pmatrix} 0.8 & 0.0 & 0.0 \\ 0.0 & 2.0 & 0.0 \\ 0.0 & 0.0 & 1.2 \end{pmatrix} \quad C = \begin{pmatrix} 0.4 & -0.1 & -0.1 \\ -0.1 & 0.4 & -0.1 \\ -0.1 & -0.1 & 0.4 \end{pmatrix} \quad K = \begin{pmatrix} 4.0 & -1.0 & -1.0 \\ -1.0 & 4.0 & -1.0 \\ -1.0 & -1.0 & 4.0 \end{pmatrix}$$

(b)

$$\phi = \begin{pmatrix} 1.000 & 1.000 & 1.000 \\ 1.735 & -1.242 & -0.093 \\ 1.165 & 2.511 & -0.342 \end{pmatrix} \quad \omega = \begin{pmatrix} 1.173 \\ 1.848 \\ 2.354 \end{pmatrix} \quad \zeta = \begin{pmatrix} 0.0587 \\ 0.0924 \\ 0.1177 \end{pmatrix}$$

Fig. 2 (a) Exact values of linearized system matrices corresponding to the small oscillations (infinitesimal) motion range. (b) Modal characteristics involving the matrices M, C, and K corresponding to fixed-base motion.

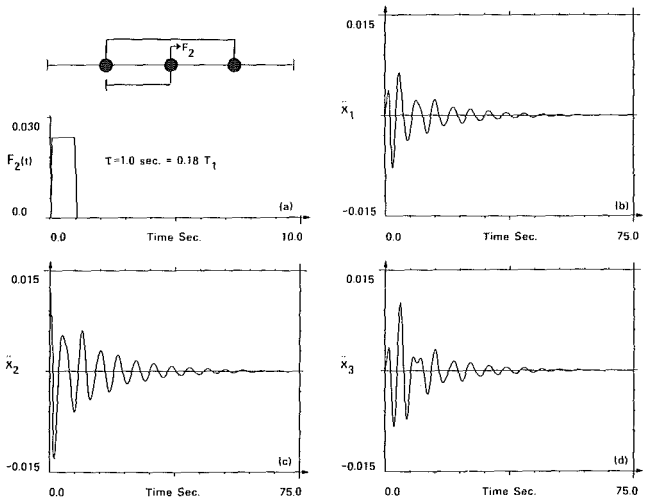


Fig. 3 Acceleration time history of the three masses in the nonlinear system under "hammer-blow" test applied to mass m_2. The duration of the impulsive excitation is approximately 0.18 of the system's fundamental period. (a) F(t); (b) $\ddot{x}_1(t)$; (c) $\ddot{x}_2(t)$; (d) $\ddot{x}_3(t)$. The same amplitude and time scale is used for all plots. Time span shown covers approximately 14 fundamental periods.

$$\mathbf{x}_1(t) = (x_1, x_2, x_3)^T, \tag{1}$$

$$\mathbf{x}_0(t) = (s_1, s_2, s_3)^T, \tag{2}$$

$$\mathbf{f}_1(t) = (F_1, F_2, F_3)^T, \tag{3}$$

$n_1 = 3$ and $n_0 = 3$.

The arbitrary nonlinear elements, denoted by g_i, that are interposed between the masses and between the support points are dependent on the relative displacement z and the velocity $\dot{z}$ across the terminals of each element. In the case of polynomial nonlinearities, the elements assume that form,

$$g_i(z, \dot{z}) = p_1^{(i)} z + p_2^{(i)} \dot{z} + p_3^{(i)} z^3, \tag{4}$$

where $p_1^{(i)}$ is the linear stiffness component, $p_2^{(i)}$ is the linear viscous damping term, and $p_3^{(i)}$ corresponds to the coefficient of the nonlinear (cubic) displacement term. Thus, depending on the sign of $p_3^{(i)}$, the form of g_i in equation (4) can be made to represent restoring forces with hardening or softening nonlinearities—a commonly encounterd type of nonlinearity in physical systems.

The form of the nonlinearity expressed by equation (4) is a polynomial-type without cross-product terms. To illustrate the wide applicability of the present method, a hysteretic-type restoring force will be considered. Such a nonlinearity not only involves cross-product terms of displacement and velocity, but is of course not even expressible in polynomial form. Hysteretic systems, widely encountered in all areas of applied mechanics (e.g., building and equipment systems, as well as aerospace structures containing collapsible or retractable elements), are among the more difficult types of nonlinear properties to investigate and identify (Caughey, 1960, 1975; Iwan, 1965, 1966; Iwan and Lutes, 1968; Jennings, 1964; Lutes and Takemiya, 1974; Andronikou and Bekey, 1984).

In the example structure under discussion, three elements (g_1, g_2, and g_3) have hardening-type polynomial nonlinear properties, and the remaining three elements (g_4, g_5, and g_6) have bilinear hysteretic properties characterized by the following parameters:

$p_1^{(i)} = k_1$ = stiffness in the elastic range,
$p_2^{(i)} = c_1$ = viscous damping term in the linear range,
$p_3^{(i)} = k_2$ = stiffness in the nonlinear range,
$p_4^{(i)} = c_2$ = viscous damping term in the nonlinear range,
$p_5^{(i)} = z_y$ = yield displacement level.

The magnitudes of the system masses as well as the material properties of the nonlinear model elements are tabulated in Fig. 1(*b*).

Notice that the structure of the model is not chain-like, consequently the linearized system stiffness matrix is not banded. The exact values of the system mass, damping, and stiffness matrices corresponding the an infintesimal ("small oscillations") range of the motion in the neighborhood of the position of static equilibrium are shown in Fig. 2 together with the associated mode shapes, natural frequencies, and modal damping values corresponding to a fixed-base configuration of the model.

$$M = \begin{pmatrix} 0.804 & 0.000 & 0.000 \\ 0.000 & 2.000 & 0.000 \\ 0.000 & 0.000 & 1.203 \end{pmatrix} \quad C = \begin{pmatrix} 0.403 & -0.100 & -0.102 \\ -0.100 & 0.400 & -0.100 \\ -0.102 & -0.100 & 0.401 \end{pmatrix} \quad K = \begin{pmatrix} 4.03 & -1.00 & -1.01 \\ -1.00 & 4.00 & -1.00 \\ -1.01 & -1.00 & 4.01 \end{pmatrix}$$

$$\phi = \begin{pmatrix} 1.000 & 1.000 & 1.000 \\ 1.734 & -1.242 & -0.093 \\ 1.164 & 2.512 & -0.341 \end{pmatrix} \quad \omega = \begin{pmatrix} 1.173 \\ 1.848 \\ 2.355 \end{pmatrix} \quad \zeta = \begin{pmatrix} 0.0586 \\ 0.0923 \\ 0.1177 \end{pmatrix}$$

$$\frac{\omega - \omega^{(i)}}{\omega^{(i)}} = \begin{pmatrix} 0.00\ \% \\ 0.00\ \% \\ 0.04\ \% \end{pmatrix} \quad \frac{\zeta - \zeta^{(i)}}{\zeta^{(i)}} = \begin{pmatrix} -0.17\ \% \\ -0.11\ \% \\ 0.00\ \% \end{pmatrix}$$

Fig. 4 Parametric identification under impulsive direct force excitation. Formulation is for a full-order system with symmetric matrices under forced vibrations. Time segment used for identification is about two fundamental periods.

3 Impulsive Excitations and Response Measurement

3.1 Probing Signal. The method under consideration imposes no restrictions on the nature of the excitation source to be used as a probing signal. It will be assumed in the present example that an impulsive excitation (resembling a "hammer blow" disturbance of the type widely used in modal analysis techniques) is applied to mass m_2.

When the above-mentioned excitation is applied to the example structure, segments of the resulting acceleration time histories of the three mass locations would be as shown in Fig. 3. The same scale is used for all locations to make relative magnitude comparison easier. The time duration shown corresponds to about 15 system fundamental periods T_1.

3.2 Data Processing. By integrating the measured acceleration time histories shown in Fig. 3, the time histories of the corresponding velocities and displacments are obtained. From that, inter-element deformations $z_i(t)$ and velocities $\dot{z}_i(t)$ can be determined.

3.3 Parametric Identification. In what follows, the task of identifying the system matrices (determining the linearized system influence coefficients) will be referred to as the "parametric" identification phase of the procedure. Conversely, the task of developing an approximating analytical representation for the nonlinear forces involved in the system motion will be referred to as the "nonparametric" identification phase of the current procedure. For convenience, a prefix *A* will henceforth be used to indicate that a referenced section, equation, or figure is in the companion paper mentioned above.

With reference to the notation introduced earlier in the companion paper, the general parametric identification procedure can be applied to the present case by noting that the problem is one in which the number of degrees of freedom is $n_1 = 3$, the number of support degrees of freedom is $n_0 = 0$ (i.e., no support motion), and the number of nonzero excitation force components is $n_f = 1$ (since only $F_2(t) \neq 0$)

By using the recursive weighted least-squares approach discussed in Section *A*2.2, the symmetric system matrices M_{11}, C_{11}, and K_{11} are identified and shown in Fig. 4. Comparing the elements of matrices M_{11}, C_{11}, and K_{11} shown in Figs. 2 and 4 shows that, if all the response measures are used to identify the dynamic system, then the identified results are accurate to within a few percent for all system parameters, insofar as the infintesimal range of motion is concerned. The small discrepancies are attributable to slight changes in the nonlinear elements. Further details regarding the application of the parametric identification procedure, under a variety of test situations, are available in the work by Masri et al. (1987a).

4 Random Base Excitation Test

4.1 Probing Signals and Response. The motion of the

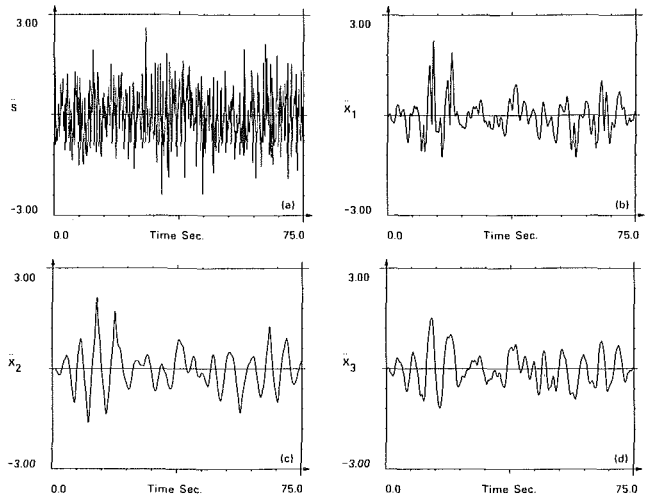


Fig. 5 Acceleration time history of the base motion $\ddot{s}(t)$ and the response $\ddot{x}_i(t)$ of the three masses in the nonlinear system shown in Fig. 1 under uniform base excitation applied to the three supports. Identical amplitude and time scales are used for all the plots. Time segment shown corresponds to approximately 14 fundamental periods of the linearized system. Input is stationary, wide-band random excitation with a flat power-spectral-density.

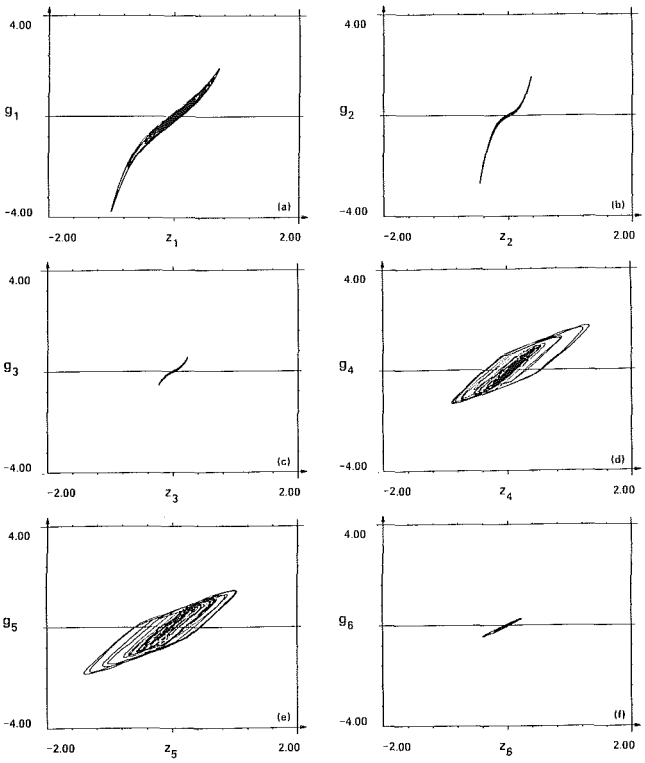


Fig. 6 Phase-plane plots of the six elements in the form of force-deformation characteristics g_i and z_i involving the nonlinear system finite elements when subjected to the uniform stationary random base excitation shown in Fig. 5

stucture discussed in Section 3 consisted of essentially small oscillations. In the present "test," the structure is assumed to be subjected to uniform, wide-band random support accelerations $\ddot{s}_1(t)$, $\ddot{s}_2(t)$, and $\ddot{s}_3(t)$. This particular choice of excitation is intended to mimic a situation in which a structure with multiple load paths is subjected to a random support motion of the type furnished by shaking tables.

4.2 Data Processing. For more realistic simulation, it will be further assumed that the only available

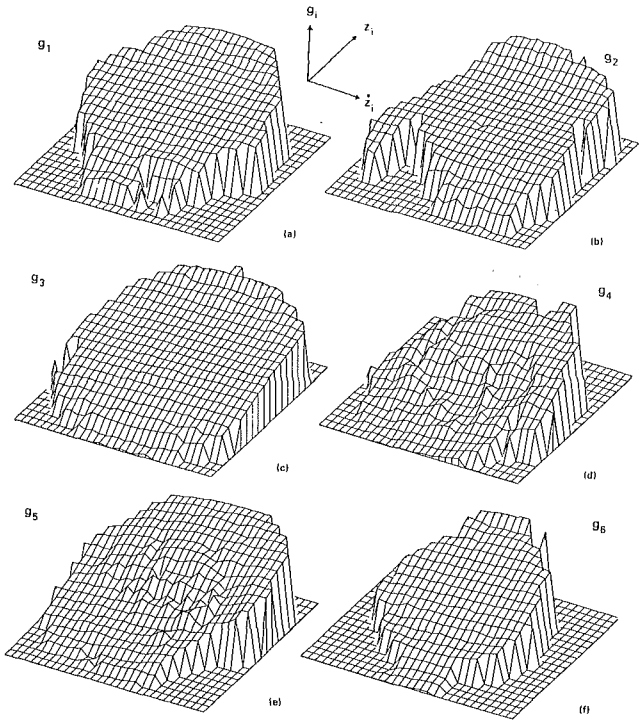


Fig. 7 Three-dimensional representation of the force deformation characteristics of the six finite elements of the system when subjected to wide-band random base acceleration

"measurements" are those of the acceleration of the three supports, ($\ddot{s}_i(t)$, $i=1,2,3$), and the acceleration of the three masses ($\ddot{x}_i(t)$, $i=1,2,3$): None of the system velocities or displacements are directly measured. A representative segment of each of the excitations (chosen, for simplicity, to be the same) and responses is shown in Fig. 5. To facilitate comparisons, the same amplitude scale is used for like response quantites at different locations. By processing the measured accelerations, the system velocities and displacements are obtained.

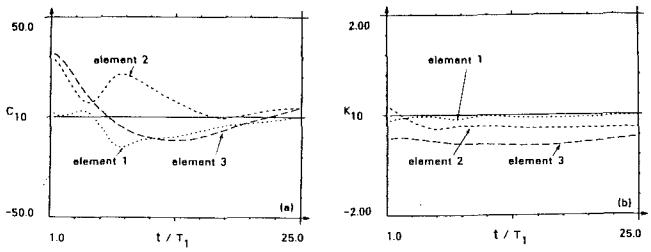
Phase-plane plots of all the elements' force-deformation curves are shown in Fig. 6. Unlike the results obtained under impulsive excitation, it is clear from the inspection of Fig. 6 that elements g_1, g_2, and g_3 have a hardening-spring characteristic, while elements g_4 and g_5 are undergoing hysteretic behavior. A three-dimensional representation of the resistance characteristics of each of the elements is shown in Fig. 7.

4.3 Parametric Identification Procedure. In the most general case for the example under discussion, the parametric identification procedure can be used to determine the elements of the six matrices M_{11}, C_{11}, and K_{11}, each of order $n_1 \times n_1$, and matrices M_{10}, C_{10}, and K_{10}, each of order $n_1 \times n_0$, where $n_1=3$ and $n_0=3$ and $n_f=0$ (since no direct excitation is applied). However, to further demonstrate the flexibility of the present method, the previously determined system matrices in Section 3, based on the "small oscillations" response, will be used as is (i.e., not recomputed). Thus, the remaining system matrices to be identified are M_{10}, C_{10}, and K_{10}. Furthermore, for simplicity, it will be assumed that the system mass matrix is diagonal. This implies that M_{10} is a null matrix.

With the above assumptions in mind, the parametric identification procedure can be expressed as:

$$C_{10}^{(e)}\dot{\mathbf{x}}_0 + K_{10}^{(e)}\mathbf{x}_0 = -(M_{11}^{(i)}\ddot{\mathbf{x}}_1 + C_{11}^{(i)}\dot{\mathbf{x}}_1 + K_{11}^{(i)}\mathbf{x}_1). \quad (5)$$

The superscripts (i) and (e) attached to the definition of the



$$C_{10} = \begin{pmatrix} -0.196 \\ -0.183 \\ -0.176 \end{pmatrix} \qquad K_{10} = \begin{pmatrix} -2.00 \\ -2.00 \\ -2.00 \end{pmatrix}$$

Fig. 8 The evolution of the element values of the identified matrices as a percentage of the exact value of the corresponding "infinitesimal motion range"

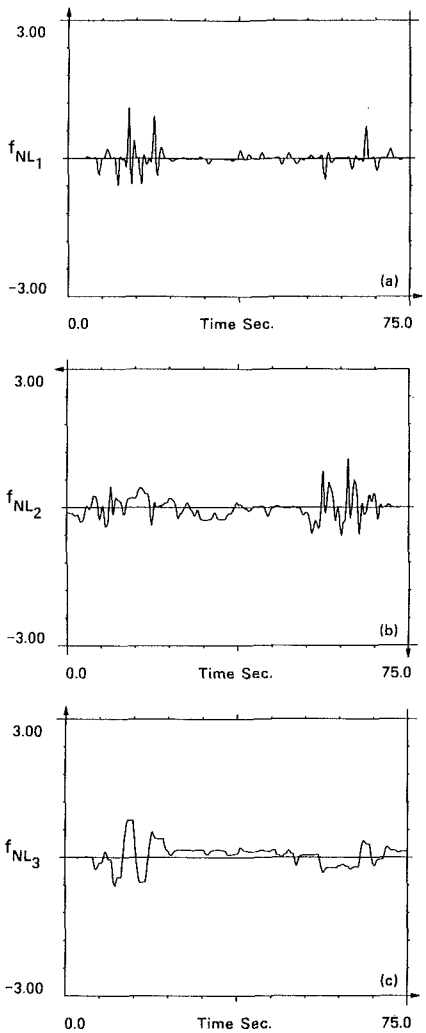


Fig. 9 Time history of the three components of the nonlinear residual force vector $f_N(t)$ obtained from the equation of motion by subtracting the contribution of the identified linear inertia, damping, and stiffness forces. The same amplitude and time scale is used for all plots. The amplitude scale matches the corresponding scale used in Fig. 5 to plot the time history of the system accelerations. The time scale used covers approximately 14 fundamental periods T_1.

system matrices appearing above designate "infinitesimal-motion range" and "equivalent linear," respectively.

Notice that, in the present case, the number of independent support motions is equal to unity since the motion of the three support points is uniform. Consequently, $n_1=3$, $n_0=1$, $n_f=0$, $\eta_{\alpha 1}=2n_0=2$, and the total number of unknown

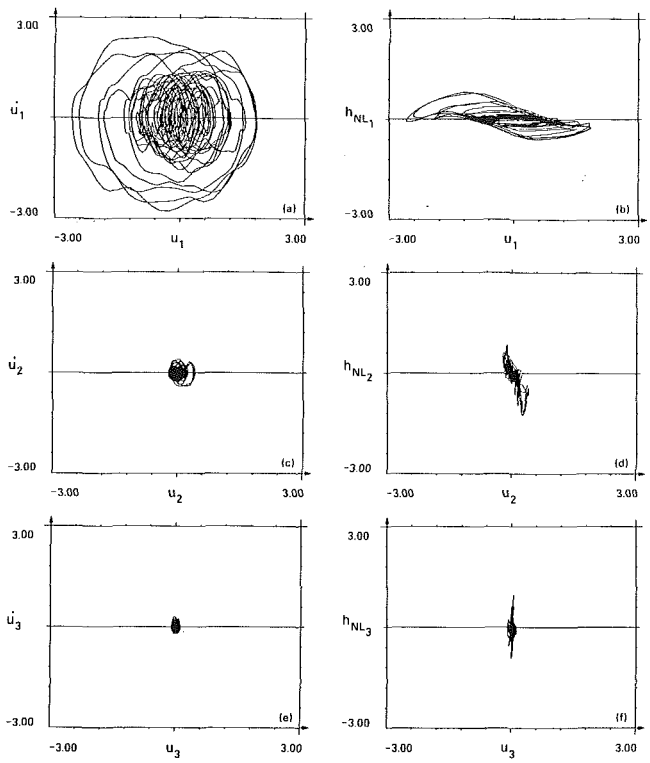


Fig. 10 Phase-plane plots of the "modal" state variables u_i and $\dot{u}_i$ and the three generalized nonlinear residual forces h_i with respect to the corresponding generalized displacement u_i. Identical horizontal and vertical scales are used for all plots. Note from plot (b) the clear evidence of the hysteretic behavior involved with the first mode.

parameters to be identified at this stage is $\eta_\alpha = n_1 \eta_{\alpha 1} = 6$. The results of the procedure are shown in Fig. 8.

The evolution of the magnitude of the identified system matrices $C_{10}^{(e)}$ and $K_{10}^{(e)}$ with time, obtained through the recursive algorithm discussed in the companion paper, is also shown in Fig. 8. For convenience, the magnitude of the ordinate of the plotted parameters have been normalized by the "exact" value of the corresponding "infinitesimal-motion range" parameters. The abscissa of the curves appearing in Fig. 8 have been normalized by T_1, the system fundamental period in the "small oscillations" range. Consequently, the ordinate of Fig. 8(*a*) covers a range ±50 percent relative to the unity value of the ratio $c_{0_j}^* = (c_{0_j}^{(e)}/c_{0_j}^{(i)})$, $j = 1,2,3$ where $c_{0_j}^{(e)}$ and $c_{0_j}^{(i)}$ is the single element in row j of the column matrix $C_{10}^{(e)}$ and $c_{10}^{(i)}$, respectively. Similarly, the ordinate of Fig. 8(*b*) covers a range ±2.0 percent relative to the unity value of the ratio $k_{0_j}^* = (k_{0_j}^{(e)}/k_{0_j}^{(i)})$, $j = 1,2,3$ where $k_{0_j}^{(e)}$ and $k_{0_j}^{(i)}$, is the element in row j of column matrix $K_{10}^{(e)}$ and $K_{10}^{(i)}$, respectively. The abscissa of the plots in both parts (*a*) and (*b*) in Fig. 8 covers a range of about $25T_1$.

It is worth noting from Fig. 8 that the spread of the results (i.e., dimensionless ordinate scales of the two plots) pertaining to the damping and stiffness influence coefficients differ by more than an order of magnitude (a factor of about 50). This behavior is consistent with the fact that, in the example under discussion, the relative contribution of damping-related forces and stiffenss-related forces is nearly inversely proportional to the above-mentioned spread.

4.4 Determination of Nonlinear Forces. Using the available measurements and the previously identified system matrices, the nonlinear system forces can now be computed from

$$\mathbf{f}_N(t) = \mathbf{f}_1(t) - (M_{11}^{(i)}\ddot{\mathbf{x}}_1 + C_{11}^{(i)}\dot{\mathbf{x}}_1 + K_{11}^{(i)}\mathbf{x}_1 + C_{10}^{(e)}\dot{\mathbf{x}}_0 + K_{10}^{(e)}\mathbf{x}_0 \tag{6}$$

(a) h_1 coefficients

$v_1 = u_1$ $\quad v_{1_{min}} = -2.5801$ $\quad v_{1_{max}} = 1.8444$

$v_2 = \dot{u}_1$ $\quad v_{2_{min}} = -2.7806$ $\quad v_{2_{max}} = 2.4147$

	$T_0(v_2)$	$T_1(v_2)$	$T_2(v_2)$	$T_3(v_2)$	$T_4(v_2)$	$T_5(v_2)$
$T_0(v_1)$	0.0430	0.4553	0.0978	0.0808	0.0330	-0.0071
$T_1(v_1)$	-0.1634	-0.0093	0.0245	0.0097	0.0786	0.0101
$T_2(v_1)$	-0.0577	0.0274	0.0351	-0.0199	-0.0050	-0.0118
$T_3(v_1)$	0.0736	-0.0003	0.0037	-0.0023	-0.1118	-0.0098
$T_4(v_1)$	0.0074	0.0228	0.0176	-0.0215	-0.0141	-0.0271
$T_5(v_1)$	-0.0081	0.0204	-0.0414	-0.0038	0.0171	-0.0120

(b) h_2 coefficients

$v_1 = u_2$ $\quad v_{1_{min}} = -0.2250$ $\quad v_{1_{max}} = 0.4094$

$v_2 = \dot{u}_2$ $\quad v_{2_{min}} = -0.4007$ $\quad v_{2_{max}} = 0.4080$

	$T_0(v_2)$	$T_1(v_2)$	$T_2(v_2)$	$T_3(v_2)$	$T_4(v_2)$	$T_5(v_2)$
$T_0(v_1)$	-0.1024	0.1227	-0.0011	0.1102	0.0190	0.0120
$T_1(v_1)$	-0.4054	-0.0229	0.0339	0.0295	0.0832	0.0622
$T_2(v_1)$	0.0154	0.0162	0.0007	-0.0254	-0.0292	-0.0083
$T_3(v_1)$	-0.0398	0.0564	-0.0105	-0.1108	-0.0794	-0.0304
$T_4(v_1)$	0.0837	0.0335	0.0296	-0.0705	-0.0621	-0.0084
$T_5(v_1)$	-0.0389	-0.0057	-0.0429	0.0415	-0.0387	0.0017

(c) h_3 coefficients

$v_1 = u_3$ $\quad v_{1_{min}} = -0.1047$ $\quad v_{1_{max}} = 0.1100$

$v_2 = \dot{u}_3$ $\quad v_{2_{min}} = -0.2076$ $\quad v_{2_{max}} = 0.3060$

	$T_0(v_2)$	$T_1(v_2)$	$T_2(v_2)$	$T_3(v_2)$	$T_4(v_2)$	$T_5(v_2)$
$T_0(v_1)$	-0.0180	-0.0912	0.0805	0.0009	0.0609	-0.0351
$T_1(v_1)$	-0.0331	0.0361	-0.0117	-0.0406	0.0298	0.0158
$T_2(v_1)$	-0.0139	-0.0314	0.0492	-0.0105	0.0287	-0.0288
$T_3(v_1)$	-0.0267	-0.0201	0.0104	-0.0145	0.0009	0.0119
$T_4(v_1)$	-0.0071	0.0200	-0.0121	0.0048	-0.0146	0.0133
$T_5(v_1)$	0.0066	-0.0193	-0.0237	0.0474	-0.0047	0.0210

Fig. 11 Nonparametric identification results. Note that the indicated coefficients correspond to the normalized Chebyshev polynomials.

With that, the time history of the nonlinear force vector $\mathbf{f}_N(t)$ components can be determined and are shown in Fig. 9. For convenience, identical scales are used for the three plots.

At this stage of the identification procedure, the "best" (in least-squares sense) equivalent linear model has been determined in the form of the identified matrices. Thus, if for the purposes of a particular application the norm of the residual error, $\|\mathbf{f}_N(t)\|$, as computed from equation (6) is sufficiently small, then the identification task can be terminated. For more demanding situations, additional processing is required to more accurately identify the residual forces that have been determined.

As pointed out earlier, if there is a need to augment the parametric identification results with additional results from the nonparametric phase of the data processing, one can proceed directly to develop approximating analytical representations, for as many of the components of $\mathbf{f}_N(t)$ as warranted, in terms of a series expansion involving suitable generalized coordinates. However, when the order of the dynamic system is relatively large, dealing with a transformed set of nonlinear forces may lead to a faster rate of convergence of the applicable series.

A convenient and natural transformation to use with realistic dynamic systems is the one expressed by equation (*A*33):

$$\mathbf{h}(\mathbf{u}, \dot{\mathbf{u}}) = \Phi^T \mathbf{f}_N(t), \tag{7}$$

$$\mathbf{u}(t) = \Phi^{-1}\mathbf{y}_1(t), \tag{8}$$

where

$$y_{1_j} = x_{1_j} - x_{0_j} \qquad j = 1,2,3, \tag{9}$$

and Φ is the modal matrix associated with $M_{11}^{-1}K_{11}$. Although the linear modal transformation of equation (8) does *not* lead

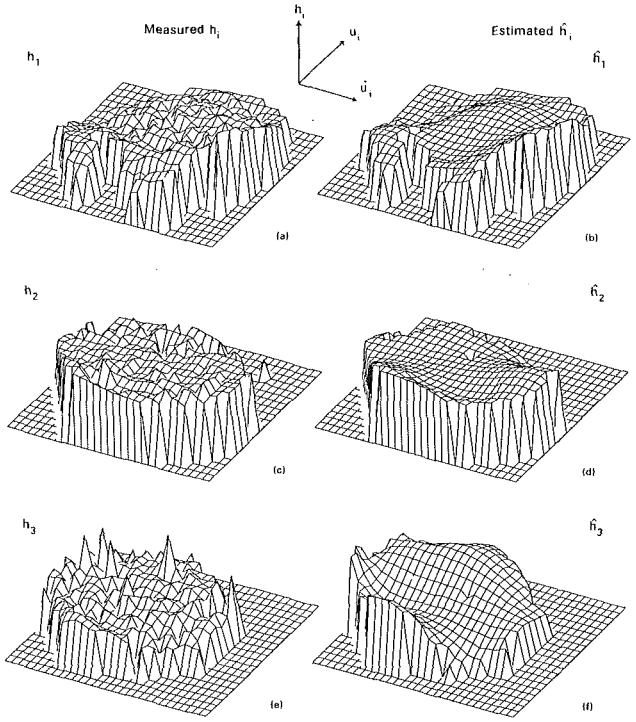


Fig. 12 (a), (c), and (e): Three-dimensional representation of the variation of each of the three generalized nonlinear forces h_i with the corresponding state variables u_j and $\dot{u}_j$; (b), (d), and (f): three-dimensional plots of the estimated (identified) generalized nonlinear forces $\hat{h}_i$ as a function of the corresponding state variables. For better resolution, different scales are used for each of the three plots.

to a decoupled set of equations in this nonlinear case, it has been found in many examples to lead to an increased rate of convergence of the series representation of the nonlinear force vector $\mathbf{f}_N(t)$. The time history of the "modal" state variables $\mathbf{u}(t)$ and $\dot{\mathbf{u}}(t)$ as well as the "modal" nonlinear forces computed in accordance with the above equations, are shown in the form of phase plots in Fig. 10.

The plots of the estimated modal restoring forces versus their corresponding modal displacement in Fig. 10 clearly indicate the presence of hysteretic components in the system.

4.5 Nonparametric Identification Procedure. Using Chebyshev polynomials in accordance with equation (*A*41) to obtain two-dimensional fits for the surfaces of the modal restoring functions will yield the typical identification results tabulated in Fig. 11. Three-dimensional representation of the transformed nonlinear forces in terms of their corresponding state variables are shown in Fig. 12 together with the approximating functions $\hat{h}_{N_j}, j=1,2,3$.

It is clear from the Tables of Fig. 11 that determining the optimum least-squares fit for the data associated with the hysteretic system does involve many cross-product terms in displacement and velocity. It also requires a relatively larger number of terms in the series (six used in the present example) for a good estimate.

The preceding is a good illustration of the need to use two-dimensional surface fits rather than the uncoupled one-dimensional series to estimate the system properties. Whether cross-coupling is significant or not is a decision that need not be made a priori when following the method presented here—the system will effectively "decide" by its own response (signature) the extent and relative dominance or contribution arising from various powers of $T_i(u)T_j(\dot{u})$.

Examination of the projections of h_i on u_j indicates negligible modal coupling in the present example. Such may not be the case in other applications. However, the presented method

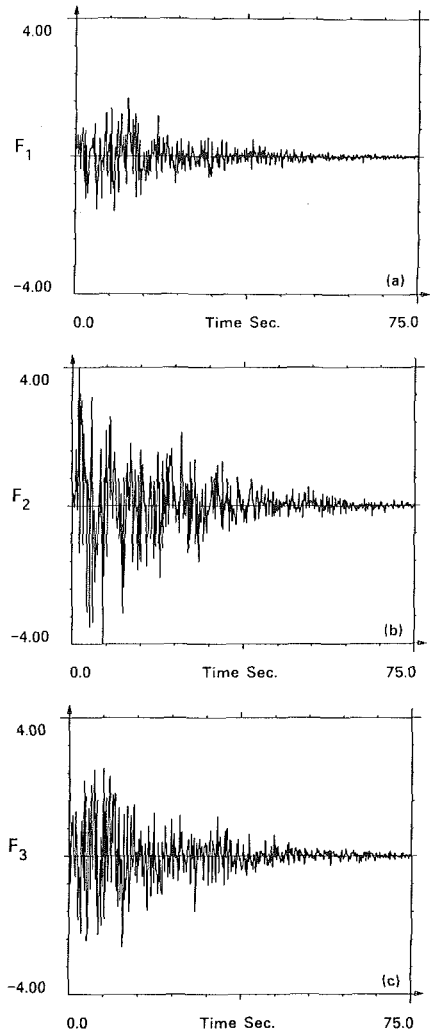


Fig. 13 Nonstationary excitation used in the validation test

can cope with nonlinear modal coupling by simply adding as many cross-coupling terms as necessary (see equation (*A*41)).

5 Response Prediction Under Different Excitation

In order to demonstrate the validity of the present identification approach, the model representation expressed by the C_{kl} coefficients shown in Fig. 11, which were extracted from the original ("exact") modal response under a probing signal consisting of stationary broad-band excitation, supplied through support motion, will now be used to predict the response of the original model when subjected to directly-applied (to mass m_2) nonstationary random excitation consisting of modulated white noise of the form

$$p_o(t) = e(t)n(t), \tag{10}$$

where $e(t)$ is a deterministic envelope function

$$e(t) = a_1 e^{a_2 t} + a_3 e^{a_4 t}, \tag{11}$$

with the a's being deterministic constants, and $n(t)$ is the output of a simulated Gaussian white noise process. The excitation time histories are shown in Fig. 13.

By following the steps indicated in equation (*A*15), the identification results can be used to predict the response time history. The adequacy of the approximate (identified) nonlinear model to predict the response of the exact (hysteretic) nonlinear system under arbitrary excitation is clearly illustrated by the results shown in Fig. 14 in which the "exact" relative displacement of each mass location of the

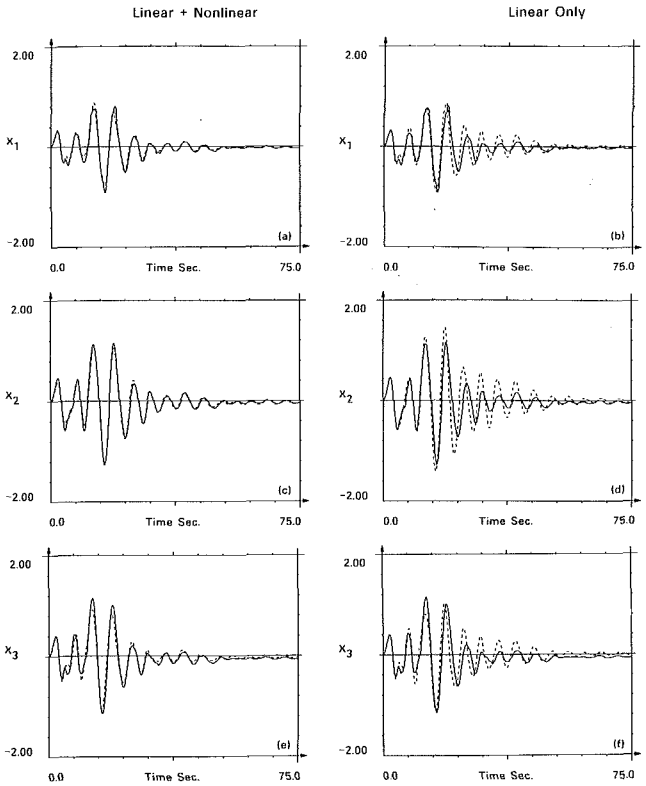


Fig. 14 Comparison between the measured and predicted response time history (*a*), (*c*), and (*e*) when both the linear and nonlinear terms are used; (*b*), (*d*) and (*f*) when only linear terms are used

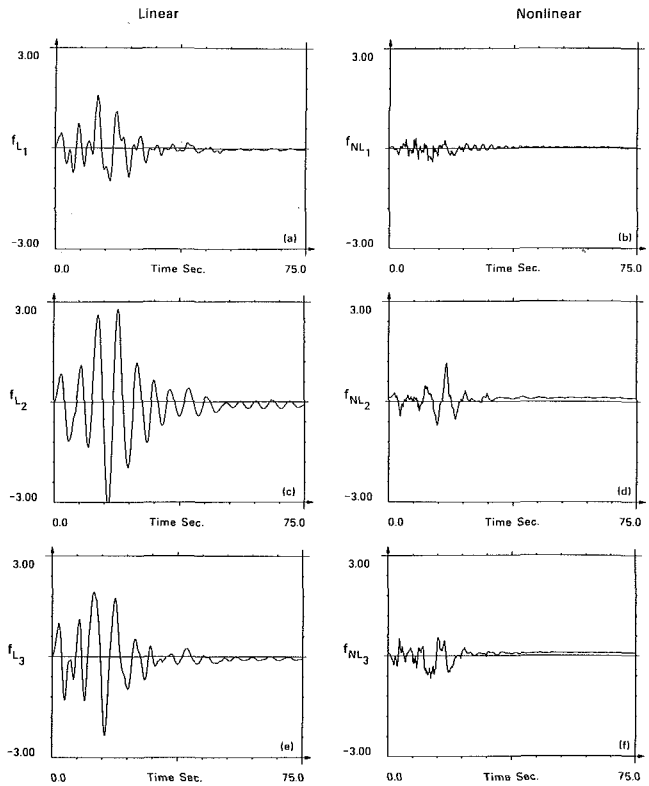


Fig. 15 Comparison between the contribution of the linear and nonlinear terms of the internal forces associated with the system degrees of freedom

nonlinear system is compared to its corresponding value as computed on the basis of the approximate nonlinear model.

The plots on the right-hand-side column of Fig. 14 show a comparison between the time history of the measured and predicted elements' deformation when only linear terms are used to compute the estimated response. Note that the exclusion of the nonlinear terms from the indentification results leads to a deterioration in the accuracy (of the amplitude as well as the phase) of the predicted response. The contribution of the nonlinear terms to the internal forces associated with the system degrees of freedom is compared to the corresponding linear terms in Fig. 15. Note that, as one would expect, the magnitude of these forces is correlated with the large amplitude range of motion.

6 Summary and Conclusions

Application of a time-domain procedure for the identification of nonlinear vibrating structures, presented in a companion paper, to a multi-degree-of-freedom nonlinear system incorporating hysteretic and polynomial-type nonlinearities, demonstrates the utility of the method under discussion. It is shown that an optimum (in the least-squares sense) reduced-order nonlinear mathematical model can be developed to match, with reasonable accuracy, all of the response time histories measured by the available sensors. Furthermore, the mathematical representation allows convenient separation of the contribution of the linear and nonlinear internal forces developed in the structure.

The illustrative examples indicate the considerable flexibility inherent in the procedure to cope with a variety of data processing and test performance situations.

Acknowledgment

This study was supported in part by a grant from the National Science Foundation and the National Institute of Health Biomedical Simulations Resource at the University of Southern California.

References

Andronikou, A. M., and Becky, G. A., 1979, "Identification of Hysteretic Systems," *Proc. of the 18th IEEE Conf. on Decision and Control*, Dec., pp. 1072–1073.

Caughey, T. K., 1960, "Random Excitation of a System with Bilinear Hysteresis," ASME Journal of Applied Mechanics, Vol. 27, No. 4, Dec., pp. 649–652.

Caughey, T. K., 1975, "Nonlinear Analysis, Synthesis and Identification Theory," *Proc. Symposium on Testing and Identification of Nonlinear Systems*, California Institute of Technology, Mar., pp. 1–14.

Distefano, N., and Todeschini, R., 1974, "Modeling, Identification and Prediction of a Class of Nonlinear Viscoelastic Materials," *International Journal of Solids and Structures*, Part I, Vol. 9, pp. 805–818.

Hudson, D. E., "Equivalent Viscous Friction for Hysteretic Systems with Earthquake-Like Excitation," *Proc. 3WCEE*, New Zealand.

Iwan, W. D., 1965, "The Steady-State Response of the Double Bilinear Hysteretic System," ASME Journal of Applied Mechanics, Vol. 32, No. 4, Dec., pp. 921–925.

Iwan, W. D., 1966, "A Distributed-Element Model for Hysteresis and Its Steady-State Dynamic Response," ASME Journal of Applied Mechanics, Vol. 33, No. 4, Dec., pp. 893–900.

Iwan, W. D., and Lutes, L. D., 1968, "Response of the Bilinear Hysteretic System to Stationary Random Excitation," *Journal of the Acoustical Society of America*, Vol. 43, No. 3, Mar., pp. 545–552.

Jennings, P. C., 1964, "Periodic Response of a General Yielding Structure," *ASCE Journal of Engineering Mechanics Division*, Vol. 90, No. EM2, pp. 131–166.

Jennings, P. C., 1968, "Equivalent Viscous Damping for Yielding Structures," ASCE *Journal of Engineering Mechanics*, Feb.

Lutes, L. D., and Takemiya, H., 1974, "Random Vibration Damping of Yielding Oscillator," *ASCE Journal of the Engineering Mechanics Division*, Vol. 100, No. EM2, pp. 343–358.

Masri, S. F., Miller, R. K., Saud, A. F., and Caughey, T. K., 1987, "Identification of Nonlinear Vibrating Structures," University of Southern California Report No. USC-CE 8710, June.

Masri, S. F., Miller, R. K., Saud, A. F., and Caughey, T. K., 1987, "Identification of Nonlinear Vibrating Structures; Part I: Formulation," ASME Journal of Applied Mechanics, Vol. 54, No. 4, Dec.

A. M. Whitman[1]
Department of Mechanical Engineering,
Villanova University,
Villanova, PA

J. E. Molyneux
Department of Mechanical and Civil Engineering,
Widener University,
Chester, PA

Limit Cycle Behavior of a Flexible Truck

We calculate the variation in critical speed of a flexible truck as a function of limit cycle amplitude and truck parameters (i.e., shear and bending stiffnesses, and truck geometry), by means of a perturbation procedure. We find that the creep force nonlinearity is dominant, and that it can cause the nonlinear critical speed to be either lower or higher than the linear critical speed, depending on the values of the two stiffnesses.

Introduction

In a previous paper (Whitman, 1983) an analytical formula for the linear critical speed of an idealized railway vehicle (Wickens, 1975; Scheffel, 1979) was obtained by doing a "short wheelbase" expansion of the system secular equation. It is tempting to use this formula (or some extension thereof (Whitman and Khaskia, 1984)), to predict the "hunting speed" of real railway vehicles. Such an identification relies on the assumption that the dominant nonlinear effects result in stable limit cycles only for speeds higher than this linear critical speed. However, such an assumption is untenable both on the basis of experience and experiment (Sweet, 1980). In the present article we identify the dominant nonlinear effect as creep force saturation, and on this basis we perform a perturbation analysis, in limit cycle amplitude, using our short wheelbase approximation as the bifurcation solution (Huilgol, 1979). The first order results indicate that limit cycles can occur either above or below the linear critical speed, depending on the values of the two suspension stiffness parameters of the model.

Formulation

We consider an idealized vehicle, Fig. 1, which is composed of two wheelsets connected by a massless frame. This is a model of a flexible truck. In terms of sum and difference coordinates

$$x_S = x_F + x_R \qquad x_D = x_F - x_R$$
$$\psi_S = \psi_F + \psi_R \qquad \psi_D = \psi_F - \psi_R \tag{1}$$

the equations for the lateral motion are

$$\begin{aligned} m\ddot{x}_S &= F_{TRF} + F_{TLF} + F_{TRR} + F_{TLR} \\ m\ddot{x}_D &= F_{TRF} + F_{TLF} + F_{TRR} + F_{TLR} - k_S\,(x_D - a\psi_S) \\ ml^2\ddot{\psi}_D &= l\,(F_{LRF} - F_{LLF} + F_{LRR}) - k_S a\,(a\psi_S - x_D) \\ ml^2\ddot{\psi}_D &= l\,(F_{LRF} - F_{LLF} - F_{LLR}) - k_b\psi_D \end{aligned} \tag{2}$$

Equations (2) contain nonlinear terms that we have not explicitly written. We will show, after normalization, that all of these terms are negligible in the bifurcation problem. We have also written, for the sake of simplicity, the quantity ml^2 for the wheelset inertia. For the configuration of Fig. 1 the stiffnesses k_S and k_b are given by (Wickens, 1975)

$$k_s = k_x k_y b^2/(k_x a^2 + k_y b^2) \tag{3}$$
$$k_b = k_y b^2$$

Wickens (1975) has shown that all forms of interconnection result in these, shear and bending, stiffnesses (although equation (3) applies only to the conventional truck sketched in Fig. 1).

The components of the contact forces between wheel and rail have been denoted in these equations by $F_{T/L}$ (transverse or longitudinal). The additional subscripts indicate the wheel

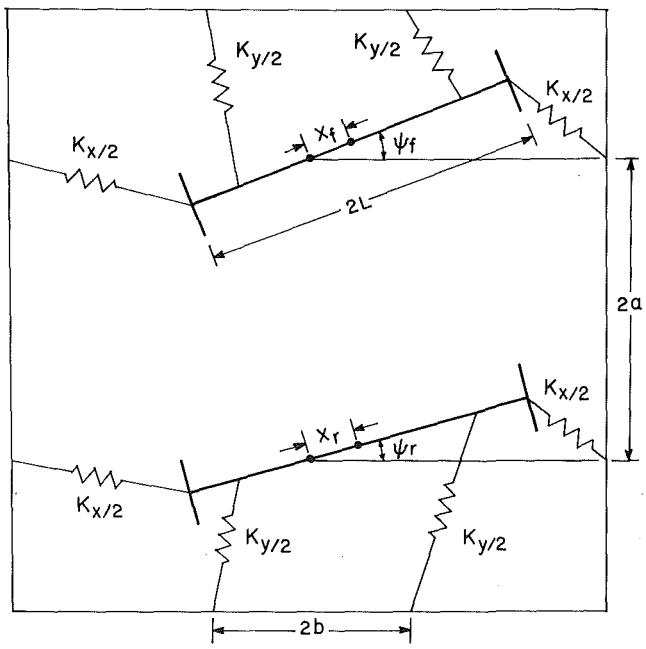


Fig. 1 Idealized model of a flexible truck adopted from Wickens (1975)

[1]Permanent Address: School of Engineering, Tel-Aviv University, Tel-Aviv, Israel.

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 5, 1985; final revision March 15, 1987.

(right or left, front or rear). For these (creep) force components, we use a nonlinear law

$$F_T = -\xi_T F/\xi \qquad F_L = -\xi_L F/\xi \tag{4a}$$

$$F/\mu_f N = \tilde{f}\xi - \frac{1}{3}(\tilde{f}\xi)^2 + \frac{1}{27}(\tilde{f}\xi)^3;\ \tilde{f}\xi \le 3$$
$$= 1;\ \tilde{f}\xi \ge 3 \tag{4b}$$

$$\xi = \sqrt{(\xi_T^2 + \xi_L^2)} \tag{4c}$$

In equations (4)

$$\tilde{f} = G\pi abC/\mu_f N \tag{5}$$

a and b are the semi-major radii of the contact ellipse, C is a number that depends on a and b, and ξ_T and ξ_L are the transverse and longitudinal contact point velocities made dimensionless by the forward speed (creepages). In addition, G is the shear modulus, μ_f is the coefficient of friction, and N is the normal force. Equation (4) is a simplified version of a force law originally developed by Vermeulin and Johnson (1964). A more complete version has recently been shown to agree rather well with numerical solutions of the governing equations (Shen, 1984).

We render equations (2) dimensionless in terms of the amplitude of the limit cycle motion, x_0, which we write as $\hat{\epsilon}$ ($\hat{\epsilon} = x_0/l$), and the kinematic frequency $\omega = \sqrt{\alpha} v/l$ (here $\alpha = \lambda l/r\ (1-\gamma)$ with r the wheel radius, λ the contact plane slope, and γ the ratio of wheel to rail curvature). Thus, in terms of

$$x_1 = x_s/\hat{\epsilon} l \qquad x_2 = \psi_s/\hat{\epsilon}\sqrt{a}$$
$$x_3 = x_D/\hat{\epsilon} a\sqrt{a} \qquad x_4 = l\psi_D/\hat{\epsilon}\alpha a \tag{6}$$

the equations of motion are

$$z^2\ddot{x}_1 = (F_{TRF} + F_{TLR} + F_{TRR} + F_{TLF})/2f\hat{\epsilon}\sqrt{a}$$
$$z^2\ddot{x}_2 - k_1\Delta(x_3 - x_2) = (F_{LRF} - F_{LLF} + F_{LRR} + F_{LLR})/2f\hat{\epsilon} a \tag{7}$$
$$z^2\ddot{x}_3 - k_1(1-\Delta)(x_2 - x_3) = (F_{TRF} + F_{LLF} + F_{TRR} - F_{LLR})l/2f\hat{\epsilon}\alpha a$$
$$z^2\ddot{x}_4 + k_2 x_4 =$$
$$(F_{LRF} - F_{LLF} - F_{LRR} + F_{LLR})l/2f\hat{\epsilon}\alpha a$$
$$\dot{\bullet} = d/d(\omega t)$$

In equations (7) we have used the quantities

$$f = \tilde{f}\mu_f N \qquad z^2 = \sqrt{a} mv^2/2f$$
$$k_1 = k_s(\alpha^2 + l^2)/2fl\sqrt{a} \qquad k_2 = k_b/2fl\sqrt{a} \tag{8}$$
$$\Delta = \alpha^2/(\alpha^2 + l^2)$$

The nonlinear terms in equations (7) that have not been written are at most of order $\hat{\epsilon}$ (Burton and Whitman, 1980). Further, the creepages are (Burton and Whitman, 1978)

$$\xi_{TF} = \hat{\epsilon}\sqrt{a}(\dot{x}_{F0} + \psi_{F0}) + O(\hat{\epsilon}^2) \qquad \xi_{LF} = \hat{\epsilon} a(\dot{\psi}_{F0} - x_{F0}) + O(\hat{\epsilon}^2)$$

where $x_{F0} = x_F/\hat{\epsilon} l$, $\psi_{F0} = \psi_F/\hat{\epsilon}\sqrt{\alpha}$. Since $\alpha << 1$, it is clear from these that we can approximate the total creepage, equation (4c), by

$$\xi_F \approx |\xi_{TF}| \qquad \xi_R \approx |\xi_{TR}|$$

On inverting equations (1), substituting the results into these approximate expressions for the total creepage of the front and rear wheelsets, using equations (6), and neglecting terms of order $\Delta\sqrt{\alpha}$, we find that

$$\xi_F \approx \xi_R \approx (\hat{\epsilon}\sqrt{\alpha}/2)|\dot{x}_1 + x_2|$$

Using this we can evaluate the creep force terms that appear in equations (7). For example

$$(F_{TRF} + F_{TLF} + F_{TRR} + F_{TLR})/2f\hat{\epsilon}\sqrt{\alpha} =$$
$$-(\dot{x}_1 + x_2) + (\hat{\epsilon}\tilde{f}\alpha/3)|\dot{x}_1 + x_2|(x_1 + x_2) + O(\tilde{f}^2\hat{\epsilon}^2\alpha) + O(\hat{\epsilon}) \tag{9}$$

The first term here is the usual linear creep force. The second term corresponds to the second term of equation (4b). This is the dominant creep force saturation behavior, and is a small correction to the linear law as long as

$$\epsilon = \hat{\epsilon} f\sqrt{\alpha}/3 << 1 \tag{10}$$

The third term corresponds to the third term of equation (4b), and is a higher order correction to the linear behavior when equation (10) is valid. The last term arises from the dominant nonlinear geometrical influences in the creepages, equation (8), acting through the linear term of the creep force law, equation (4b).

Now in the typical railway vehicle the parameter f is very large, on the order of 5000, and α lies in the range from 0.04 to 0.05. Thus, we find from equation (10) that $\hat{\epsilon}/\epsilon$ is on the order of 0.0025. Accordingly, for oscillator amplitudes for which the creep force is significantly nonlinear, say $\epsilon = 0.2$, the parameter $\hat{\epsilon} = 0.0005$; consequently all geometrical nonlinearities and dynamical nonlinearities are negligible at this level, since they are all $0(\hat{\epsilon})$. As a result of this argument we can infer that the variation of the critical speed with amplitude is governed, initially at least, by the nonlinearity given by the second term in equation (4b). The equations of motion are, therefore, correct to terms linear in ϵ,

$$z^2\ddot{x}_1 + (\dot{x}_1 + x_2) = \epsilon|\dot{x}_1 - x_2|(\dot{x}_1 + x_2)$$
$$z^2\ddot{x}_2 + (\dot{x}_2 - x_1) - k_1\Delta(x_3 - x_2) = \epsilon|\dot{x}_1 - x_2|(\dot{x}_2 - x_1)$$
$$z^2\ddot{x}_3 + (\dot{x}_3 + x_4) - k_1(1-\Delta)(x_2 - x_3) = \epsilon|\dot{x}_1 + x_2|(\dot{x}_3 + x_4) \tag{11}$$
$$z^2\ddot{x}_4 + (\dot{x}_4 - x_3) + k_2 x_4 = \epsilon|\dot{x}_1 + x_2|(\dot{x}_4 - x_3)$$

Analysis

The equations of motion may be conveniently written as a vector system

$$z^2\ddot{\boldsymbol{x}} + \dot{\boldsymbol{x}} + \boldsymbol{A}\boldsymbol{x} = \epsilon \boldsymbol{F}(\boldsymbol{x}, \dot{\boldsymbol{x}}) \tag{12}$$

where $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^T$, $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}) = |y_1 + x_2|\ (\boldsymbol{M}\boldsymbol{x} + \boldsymbol{y})$, and the 4×4 matrices $\boldsymbol{A}$, and $\boldsymbol{M}$ are given by

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & k_1\Delta & -k_1\Delta & 0 \\ 0 & -k_1(1-\Delta) & k_1(1-\Delta) & 1 \\ 0 & 0 & -1 & k_2 \end{pmatrix} \tag{13}$$

$$\boldsymbol{M} = diag\left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}$$

The behavior of the linear system, $\epsilon = 0$, in equation (12), has been analyzed in Whitman (1983). In this case, it is clear that for $\boldsymbol{x} = \boldsymbol{a}_\lambda \exp(\sigma t)$ to be a solution, $\boldsymbol{a}_\lambda$ must be an eigenvector of $\boldsymbol{A}$ belonging to the eigenvalue λ and σ must satisfy $z^2\sigma^2 + \sigma + \lambda = 0$. In our analysis of the nonlinear system, we will need the following results from Whitman (1983): (i) The eigenvalues are analytic functions of Δ in a neighborhood of $\Delta = 0$ provided $|k_1 - k_2| \neq 2$. In fact, $f(\lambda) = det(\boldsymbol{A} - \lambda \boldsymbol{I}) = f_1(\lambda) - k_1 k_2 \Delta$ with $f_1(\lambda) = (\lambda^2 + 1)(\lambda^2 - (k_1 + k_2)\lambda + 1 + k_1 k_2)$, so that if $\lambda(\Delta)$ denotes any eigenvalue, then $\lambda(\Delta) = \lambda(0) + \Delta k_1 k_2/f'(\lambda(0)) + O(\Delta^2)$ and $\lambda(0)$ is one of the roots of f_1. Using this expression we obtain

$$\lambda_1 = \lambda_2^* = i + \Delta k_0(1 + ik_0)/(2(1 + k_0^2)) + O(\Delta^2) \tag{14}$$
$$\lambda_3 = \lambda_4^* = \lambda_+ + \Delta k_1 k_2/((\lambda_+ - \lambda_-)(1 + \lambda_+^2)) + O(\Delta^2)$$

where

$$k_0 = k_1 k_2/(k_1 + k_2)$$
$$\lambda_\pm = (k_1 + k_2)/2 \pm i\{((k_1 - k_2)/2)^2 - 1\}^{1/2}$$

It is easy to see using the Routh test that if $\Delta > 0$, then $Re(\lambda_i) > 0$, $j = 1, \ldots, 4$. Eight characteristic exponents $\sigma_i(\lambda_j, x^2)$, $i = 1,2$, $j = 1, \ldots, 4$, are determined by the quadratic $z^2\sigma^2 + \sigma + \lambda_j = 0$. The minimum critical speed $z^2 = z^2(k_1, k_2, \Delta)$, as defined in Whitman (1983), is obtained by requiring

$Re(\sigma_i(\lambda_2, z^2)) = 0$ for, say $i = 1$. This value which we will denote by z_0^2 is given by

$$z_0^2 = Re(\lambda_2)/(Im(\lambda_2))^2 = \Delta k_0/2(1+k_0^2) + O(\Delta^2) \tag{15}$$

and the corresponding frequency of oscillation is

$$\omega_0 = Im(\sigma_1(\lambda_2, z^2)) = \Delta k_0^2/2(1+k_0^2) + O(\Delta^2) \tag{16}$$

For our analysis of equation (12) we will need, in addition to the results summarized above, the eigenvectors α_{λ_1} of the complexification of A corresponding to eigenvalues appearing in equation (14). Computation of these vectors as a power series in Δ is a straightforward, although tedious, task. The results are:

$$\alpha_{\lambda_1} = (1, \lambda_1, a_3, a_4)^T \tag{17}$$

where

$$a_3 = ik_0(k_2 - i)/\{k_2(k_0 - i)\} + O(\Delta)$$

$$a_4 = ik_0/\{k_2(k_0 - i)\} + O(\Delta)$$

and

$$\alpha_{\lambda_3} = (0, 0, 1, (k_2 - k_+)^{-1}) + O(\Delta)$$

Of course, $\alpha_{\lambda_{l+1}} = \alpha_{\lambda_l}^*$, $l = 1, 3$, and we have chosen a normalization which makes the vectors α_{λ_l} regular as $\Delta \to 0$. If we define the real vectors $\alpha_l = Im(\alpha_{\lambda_l})$, $\alpha_{l+1} = Re(\alpha_{\lambda_l})$, $l = 1, 3$, and if z^2 has the value z_0^2, then equation (12) with $\epsilon = 0$ has the real periodic solutions

$$p_1(\omega_0 t) = \sin(\omega_0 t)\alpha_1 + \cos(\omega_0 t)\alpha_2 \tag{18}$$

$$p_2(\omega_0 t) = \cos(\omega_0 t)\alpha_1 - \sin(\omega_0 t)\alpha_2 \tag{19}$$

and any solution with period $T_0 = 2\pi/\omega_0$ is a linear combination of $\mathbf{p}_1$ and $\mathbf{p}_2$.

Now let us investigate the existence of perodic solutions of equation (12) for $\epsilon > 0$. We will base our analysis on a formal perturbation method (Nayfeh, 1973), but will also discuss briefly at the conclusion of this section an alternative procedure by which our main results can be derived. In equation (6) we have adopted the normalization $\max(x_1(t,\epsilon)) = 1$, and we have seen in equation (18) that for $\epsilon = 0$, $p_1(\omega_0 t)$ is a $(2\pi/\omega_0)$-periodic solution of equation (12) (with $z^2 = z_0^2$) which satisfies this condition. For $\epsilon > 0$, we assume that the critical speed takes the form $z^2 = \tilde{z}^2(\epsilon) = z_0^2 + \epsilon z_1^2 + o(\epsilon)$ and that there exist $(2\pi/\omega)$-periodic solutions of equation (12), with

$$\omega = \hat{\omega}(\epsilon) = \omega_0 + \epsilon\omega_1 + o(\epsilon)$$

We suppose that any such solution can be written as $x(t,\epsilon) = x_0(s) + \epsilon x_1(s) + 0(\epsilon)$, where $s = \omega t$. Substituting these expressions into equation (12) and equating the coefficients of ϵ^0 and ϵ^1 equal to zero, we obtain (20)

$$z_0^2\omega_0^2 \ddot{x}_0 + \omega_0 \dot{x}_0 + Ax_0 = L(x_0) = 0 \tag{20}$$

$$\begin{aligned} L(x_1) &= -\{(z_1^2\omega_0^2 + 2z_0^2\omega_0\omega_1)\ddot{x}_0 + \omega_1 \dot{x}_0\} + F(x_0, \omega_0 \dot{x}_0) \\ &= G\{x_0\}(s) \end{aligned} \tag{21}$$

where $\dot{} = d/ds$. We see immediately that the 2π-periodic solution of equation (20) which has the required normalization is $x_0(s) = p_1(s)$. The right-hand side of equation (21), $G(p_1)(s)$ is, therefore, a 2π-periodic function and has the Fourier series expansion

$$G\{p_1\}(s) = \sum_{n=-\infty}^{\infty} G_n \exp(ins) \quad G_n = G_{-n}^0 \tag{22}$$

Since

$$p_1(s) = Re\{\alpha_{\lambda_2} \exp(is)\}$$

we see that

$$G_1 = (1/2)\{(z_1^2\omega_0^2 + 2z_0^2\omega_0\omega_1) - i\omega_1\}\alpha_{\lambda_2} + F_1 \tag{23}$$

$$G_n = F_n, \; n \geq 0, \; n \neq 1$$

where the $F_n = F_{-n}^*$ are the Fourier coefficients of

$$F(p_1, \omega_0 p_2) = z_0^2 |\cos(s)|\{M\rho_1 + \omega_0 \rho_2\}$$

$$= \sum_{n=-\infty}^{\infty} F_n \exp(ins) \tag{24}$$

It is now clear that any particular solution of equation (21) will contain secular terms increasing linearly with s unless G_1 lies in the subspace of C^4 spanned by the vectors α_{1_j}, $j \neq 2$. Hence, if $x_1(s)$ is to be periodic, we must require that

$$F_{\lambda_2} + (1/2)(z_1^2\omega_0^2 + 2z_0^2\omega_0\omega_1) - (1/2)i\omega_1 = 0 \tag{25}$$

where

$$F_1 = \sum_{j=1}^{4} F_{\lambda_j}\alpha_{\lambda_j}$$

This (complex) equation determines the (real) first order corrections, ω_1 and z_1^2, to the circular frequency and the critical speed. We see from equation (18) that

$$Mp_1 + \omega_0 p_2 = Re\{v_{\lambda_2} \exp(is)\}, \; v_{\lambda_2} = M\alpha_{\lambda_2} + i\omega_0\alpha_{\lambda_2}$$

Thus, equation (24) gives

$$F_1 = z_0^2\{(1/\pi)v_{\lambda_2} + (1/3\pi)v_{\lambda_2}^*\}$$

and by expressing v_{λ_2} as a linear combination of the eigenvectors,

$$v_{\lambda_j} = \sum_{j=1}^{4} \beta_j \alpha_{\lambda_j}$$

we find that $F_{\lambda_2} = z_0^2(\beta_2/\pi + \beta_1^*/3\pi)$. The actual computation of β_1 and β_2 using equations (13), (17) and the definition of v_{λ_2} is a straightforward but laborious process. We find

$$F_{\lambda_2} = k_0^2\{7k_0^2 + 1 - 2k_0(k_0^2+1)/k_2 - 6ik_0\}\{12\pi(k_0^2+1)^3\}^{-3}$$

The above expression and the values of z_0^2 and ω_0 given in equations (15) and (16) may now be substituted into equation (25) and the desired corrections to the critical speed and frequency may be computed. We find

$$\begin{aligned} z^2 = \{k_0\Delta/2(1+k_0^2)\}\{1 + (\epsilon\Delta/3\pi)k_1k_2[(k_1+k_2)^2 + k_1^2k_2^2]^{-2}(26) \\ \times [(k_1+k_2)^2(k_1-k_2) - k_1^2k_2^2(5k_1+7k_2)]\} + O(\epsilon, \Delta^2) \end{aligned}$$

$$\omega = 1 - \{k_0^2\Delta/2(1+k_0^2)\{1 + (\epsilon\Delta/\pi)2k_0(1+k_0^2)^{-2}\} + O(\epsilon, \Delta^2) \tag{27}$$

The dependence of z^2 on the stiffness parameters k_1 and k_2 will be discussed in the following section.

To conclude our analysis of the equations of motion, we sketch briefly an alternate perturbative approach for finding the dependence of z^2 and ω on ϵ. The technique we consider has been used in Coddington and Levinson (1955) for studying bifurcation of periodic solutions. We note that any real solution of equation (12) can be written in the form

$$x = \sum_{j=2}^{4} (\alpha_{\lambda_j} u_j + \alpha_{\lambda_j}^* u_j^*)$$

Substituting this expression into equation (12) and using the linear independence of the vectors $\{\alpha_{\lambda_l}, \alpha_{\lambda_l}^*\}$ $j = 2, 4$ we find that

$$z^2\ddot{u}_j + \dot{u}_j + \lambda_j u_j = \epsilon G_j\{u_2, \dot{u}_2, u_4, \dot{u}_4\}; j = 2, 4 \tag{28}$$

where

$$G_j = 2|Re(\dot{u}_2 + \lambda_2 u_2)|\left\{\sum_{k=2,4}(u_k\beta_{kj} + u_k^*\gamma_{kj}^*) + \dot{u}_j\right\} \tag{29}$$

and we have expressed $M\alpha_{\lambda_j}$ as the sum

$$M\alpha_{\lambda_j} = \sum_{i-2,4}(\beta_{jl}\alpha_{\lambda_l} + \gamma_{jl}\alpha_{\lambda_l}^*)$$

For future reference, we note that $\beta_2 = \beta_{22} + i\omega_0$ and $\gamma_{22} = \beta_1^*$.

In equation (28) we introduce the change of independent variable $\tau = t/z^2$, and let $v(\tau) = u_2(t)$, $\omega(\tau) = u_4(t)$, and $' = d/d\tau$. Then we obtain the pair of equations

$$L_2(v) = v'' + v' + z_0^2\lambda_2 v = \epsilon G_2\{v, v'/z^2; w, w'/z^2\} + \lambda_2(z^2 - z_0^2)v \tag{30}$$

$$L_4(w) = w'' + w' + z_0^2\lambda_4 w = \epsilon G_4\{v, v'/z^2; w, w'/z^2\} + \lambda_4(z^2 - z_0^2)w \tag{31}$$

Assuming that $z^2 = z_0^2 + \epsilon z_1^2 + o(\epsilon)$, we see that for $\epsilon = 0$ equations (30) and (31) have the periodic solution:

$$v(\tau) = C_0 \exp(i\Omega_0\tau), w(\tau) = 0, \quad \Omega_0 = z_0^2 w_0$$

Let $K_2(\tau)$ be the impulse response function for the right-hand side of equation (30), i.e., $L_2(K_2) = 0$, $k_2(0) = 0$, $K_2'(0) = 1$. Then $k_2(\tau) = (1 + 2i\Omega_0)^{-1}\{\exp(i\Omega_0\tau - \exp(-(1 + i\Omega_0)\tau\}$

and any periodic solution (v, w) of equations (30) and (31) must satisfy

$$v(\tau) = C_0 \exp(i\Omega_0\tau) + \int_0^\tau K_2(\tau - \hat{\tau}) H_{2\epsilon}(\hat{\tau}) d\hat{\tau}$$

where

$$H_{2\epsilon}(\tau) = \epsilon x^2 G_2\{v, v'/z^2; w, w'/z^2\} + \lambda_2(z^2 - z_0^2)v \tag{32}$$

The corresponding equation for w is of the form $w(\tau) = \int_0^\tau K_4(\tau - \hat{\tau}) H_{4\epsilon} d\hat{\tau}$ with K_4 and $H_{4\epsilon}$ defined in a similar way as K_2, $H_{2\epsilon}$. In order for (v, w) to be periodic with period $T = T_0 + \epsilon T_1 + 0(\epsilon)$, $T_0 = 2\pi/\Omega_0$, it is necessary and sufficient that $v(T) = v(0)$, $v'(T) = v'(0)$ and that similar conditions hold for $\omega(\tau)$. We impose these conditions on equation (32) and the corresponding expression for w. Thus, we obtain

$$C_0\{\exp(i\Omega_0 T) - 1\} + \int_0^T K_2(T - \hat{\tau}) H_{2\epsilon}(\hat{\tau}) d\hat{\tau} = 0$$

$$i\Omega_0 C_0\{\exp(i\Omega_0 T) - 1\} + \int_0^T K_2'(T - \hat{\tau}) H_{2\epsilon}(\hat{\tau}) d\hat{\tau} \tag{33}$$

and two similar equations involving integrals of H_{45}. We divide both of the above equations by ϵ and let $\epsilon \to 0$. Assuming $(v, \omega) \to (C_0 \exp(i\Omega_0\tau), 0)$ as $\epsilon \to 0$, we obtain

$$iC_0\Omega_0 T_1 + \int_0^{T_0} K_2(T_0 - \hat{\tau}) J(\hat{\tau}) d\hat{\tau} = 0 \tag{34}$$

$$-C_0\Omega_0^2 T_1 + \int_0^{T_0} K_2'(T_0 - \hat{\tau}) J(\hat{\tau}) d\hat{\tau} = 0 \tag{35}$$

where

$$J(\tau) = \lim_{\epsilon \to 0}\{(1/\epsilon) H_{2\epsilon}(\tau)\} = z_0^2 G_2\{v_0, v_0'/Z_0^2; 0, 0\} + \lambda_2 z_1^2 v_0$$

and $v_0 = C_0 \exp(i\Omega_0\tau)$. Noting that $K_2'(\tau) + (1 + i\Omega_0)K_2(\tau) = \exp(i\Omega_0\tau)$ and forming the corresponding linear combination of equations (34) and (35), we obtain

$$C_0\{-\lambda_2 z_1^2 T_0 - 2\Omega_2^2 T_1 + i\Omega_0 T_1\} + z_0^2 \int_0^{T_0} \exp(-i\Omega_0\hat{\tau}) G_2(v_0(\hat{\tau}), v_0'(\hat{\tau})/z_0^2; 0, 0) d\hat{\tau} = 0 \tag{36}$$

where

$$G_2 = 2z_0^2\omega_0^2 |Re(v_0(\hat{\tau}))| \{(\beta_2 v_0(\hat{\tau}) + \beta_1^* v_0^*(\hat{\tau})\}$$

This equation and a similar relation derived from the combination of equation (35) and $-i\Omega_0$ times equation (34) give the results found in equation (25) for z_1^2 and ω_1. As with our previous approach, this method requires that the parameters $z^2(\epsilon)$ and $\omega(\epsilon)$ have derivatives with respect to ϵ at $\epsilon = 0$. On the other hand, it appears to demand only that the solution (v, w) depends continuously on ϵ as $\epsilon \to 0$.

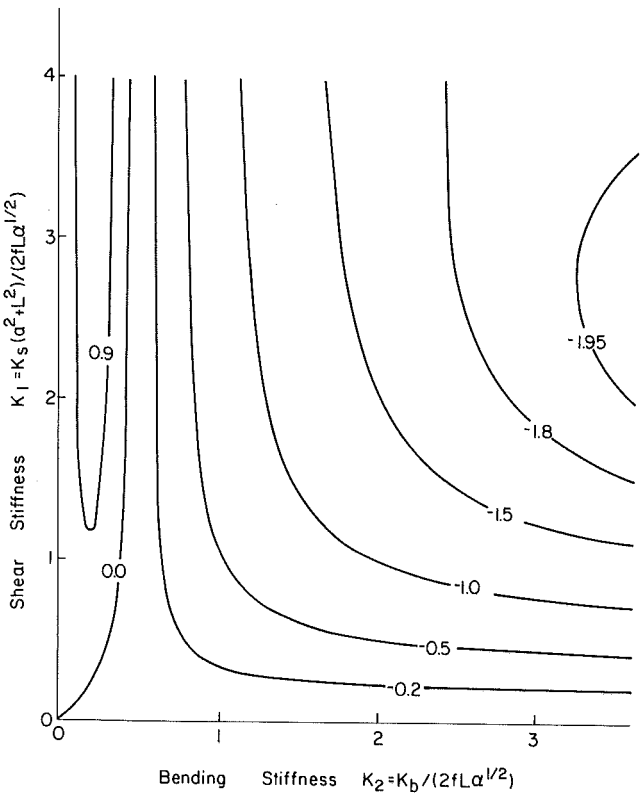


Fig. 2 Contours of the coefficient of $\epsilon\Delta/3\pi$ in equation (26)

Discussion of Results

The main results obtained here are the initial deviation of the hunting frequency, equation (27), and critical speed, equation (26), from their linear values as a function of the amplitude of the limit cycle.

From equation (27) we can see that the frequency is reduced by the creep saturation effect, and that the reduction is symmetric in k_1 and k_2. In fact, it depends only on their series combination, k_1, to leading order.

The effect on critical speed is more complex. In Fig. 2 we plot a universal measure of the deviation, the coefficient of $\epsilon\Delta/3\pi$ in equation (26), of critical speed from the linear result, as a function of the two stiffnesses. It can be seen from this figure that over a large part of the k_1, k_2 plane, the critical speed is decreased as a result of creep force saturation. This reduction is, however, not symmetrical in k_1 and k_2. Indeed there is even a small region (of positive valued contour curves) in which the critical speed increases with increasing motion amplitude. Of course one cannot take this figure literally. The bifurcation problem that occurs here is an example of a Hopf bifurcation, and it is known that the positive valued contour curves (supercritical case) correspond to stable limit cycles while the negative valued curves (subcritical case) correspond to unstable ones. Nonetheless one can infer that contours with larger negative values correspond to trucks with lower hunting speeds.

It has long been known in the rail vehicle design community that radial trucks (those with large k_1 and small k_2) have better hunting characteristics than conventional trucks (those with small k_1 and large k_2), but the reasons for this were not clear. On the basis of a purely linear theory (Whitman, 1983), there should be no difference between the two types of truck,

because in that case the critical speed is symmetric in the two stiffness parameters. The result noted in the previous paragraph, however, does serve as an explanation of this physical fact. Interestingly, other attempts at explanation based on using geometric nonlinearities have failed (Sweet, 1980). We regard this as support for the argument that creep force saturation is the dominant nonlinear effect.

A further point of interest is that, because of the nonanalytic nature of the creep force nonlinearity, the forcing function, even to first order, contains harmonics of all orders, equation (22). Consequently any numerical solution algorithm must account for this in order to be reasonably accurate. For this reason, any single frequency approximation method such as equivalent linearization is inadequate for this problem.

References

Burton, T. D., and Whitman, A. M., 1980, "Nonlinear Contact Geometry Effects on Wheelset Dynamics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 155–160.

Burton, T. D., Whitman, A. M., 1978, "Nonlinear Kinematics of Wheel-Rail Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 664–668.

Coddington, E. A., and Levinson, N., 1955, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York.

Huilgol, R. R., 1979, "Hopf-Friedricks Bifurcation and the Hunting of a Railway Axle," *Quarterly of Applied Mathematics*, Vol. 36, pp. 85–95.

Nayfeh, A. H., 1973, *Perturbation Methods*, Wiley, New York.

Scheffel, H., 1979, "The Influence of the Suspension on the Hunting Stability of Railway Vehicles," *Rail International*, Vol. 10, pp. 662–677.

Shen, Z. Y., et. al., 1984, "A Comparison of Alternative Creep Force Models for Rail Vehicle Dynamics," *The Dynamics of Vehicles on Roads and Tracks, 8th IAVSD Conference*, Hedrick, J. K., ed., Swets and Zeitlinger, Lisse, pp. 591–605.

Sweet, L., 1980, "Experimentally Derived Reformulation of the Wheelset Nonlinear Hunting Problem," *New Approaches in Nonlinear Problems in Dynamics*, Holmes, P., ed., SIAM, Philadelphia, PA, pp. 463–469.

Vermeulin, P. J., Johnson, K. L., 1964, "Contact of Nonspherical Elastic Bodies Transmitting Tangential Forces," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 338–340.

Whitman, A. M., 1983, "On the Lateral Stability of a Flexible Truck," ASME *Journal of Dynamic Systems, Measurement, and Control*, Vol. G105, pp. 120–125.

Whitman, A. M., Khaskia, A. R., 1984, "Freight Car Lateral Dynamics – An Asymptotic Sketch," ASME *Journal of Dynamic Systems, Measurement, and Control*, Vol. G106, pp. 107–113.

Wickens, A. H., 1975, "Steering and Dynamic Stability of Railway Vehicles," *Vehicle System Dynamics*, Vol. 5, pp. 15–46.

Coupled Gas-Liquid-Structure Systems: Part 1—Theory[1]

F. M. Joos
Creare, Inc.,
Hanover, NH 03755
Assoc. Mem. ASME

P. W. Huber
Science Concepts, Inc.,
Washington, DC 20003

A complex hydrodynamic transient due, for example, to the injection of gas into a liquid, creates pressure forces on adjacent structures. These structures, together with gas cavities in the liquid itself, represent flexible boundaries to the distributed, time-varying liquid mass. The response of the gas-liquid-structure system depends on the intrinsic flexibility of the gas cavities and on the flexibility of structural boundaries. In this paper we analyze the dynamics of such systems where the liquid is incompressible. We present systematic procedures for driving the response of one system from the known response of a geometrically identical system with different flexibility. Finally, we outline the analysis for the compressible case.

I Introduction

Consider a system consisting of a bounded or unbounded liquid in motion, interfacing with deformable structures and pockets of gas. Pressures and displacements throughout the liquid depend on the characteristics of the flow and the properties of the boundary structures and gases. Of the many problems in this class, we consider in this paper those in which the boundaries oscillate so as to produce oscillations in the liquid that are rapid and of small amplitude compared to the times and lengths that characterize the liquid's bulk motion.

Flows of this type may be encountered in flexbile tanks partially filled with liquid and subjected to excitations during an earthquake or the acceleration of a rocket, in underwater explosions, in implosions due to vapor condensation in tanks, in waterhammer effects in piping systems, in the buoyant rise and vibration of large gas bubbles through a liquid, in the injection of gas or steam into a liquid-filled flexible tank, or in the impact of a flexible structure such as a ship's hull on the surface of a liquid. Illustrative references for these types of flows can be found in Joos (1982).

These flows are generally analyzed in one of two ways. When there is no bulk deformation, that is, when the liquid appears "frozen" except for small amplitude, oscillatory motion, the cross-convective terms in the Navier-Stokes equations are ignored and viscosity is neglected because the frequency of the boundary motion is high. The momentum and continuity equations in the liquid then reduce to the wave equation, or to Laplace's equation if the liquid can be considered incompressible. Because displacements are small, the boundaries of the liquid are assumed fixed, as if the gas cavities or structures were rigid bodies. With these assumptions it is fairly easy to analyze the free or forced vibrations of such systems. For example, Kana et al. (1968) studied vibrations of liquid-filled cylindrical containers using this approach. Nevertheless, a paper by Antony-Spies (1979) is the only previous work we have seen in which the coupling of gas, liquid, and structures is studied in such a system.

When the liquid cannot be treated as frozen, that is when there *is* transient bulk flow to consider, the usual approach has been to set about solving the liquid momentum and continuity equations (including all or most of the terms in those equations) simultaneously with equations that specify the constitutive relations for the structures and gas cavities in the liquid. The solution is complicated by various factors: structural equations are typically formulated in Lagrangian form while the liquid equations are Eulerian, and surfaces (free or otherwise) change shape during the transient altering the computational domain. The result is, invariably, a complex computer code that is lengthy, costly to run, and difficult to verify. Belytschko and Schumann (1980) review some analyses of this type applied to nuclear energy containment systems.

Sonin (1980), and Kalumuck and Huber (1981), have proposed a simpler method for analyzing flows of this second type. Their approach effectively decouples the solution of the deforming flow in a hypothetical "rigid" system from the small amplitude high frequency oscillations in the actual "flexible" system. But the method is restricted to cases where the motions of all gas cavities are entirely independent of the structural oscillations. The possibility of coupling between the oscillations at the structural boundaries and those at the gaseous ones is excluded.

In this paper we bring together the two evolving bodies of learning in this area and develop theoretical tools for dealing with the more general problem in which both structures and gas cavities are allowed to oscillate, while the liquid domain also undergoes bulk deformations. Many of the problems that have been analyzed previously reduce to special cases of our theory. In the final section of the paper we show how the methods we have developed may be extended to account for the effects of liquid compressibility. In the companion paper

[1]The bulk of the work presented in this paper was performed while the authors were at the Massachusetts Institute of Technology, Cambridge, MA.

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 17, 1986; final revision July 9, 1987.

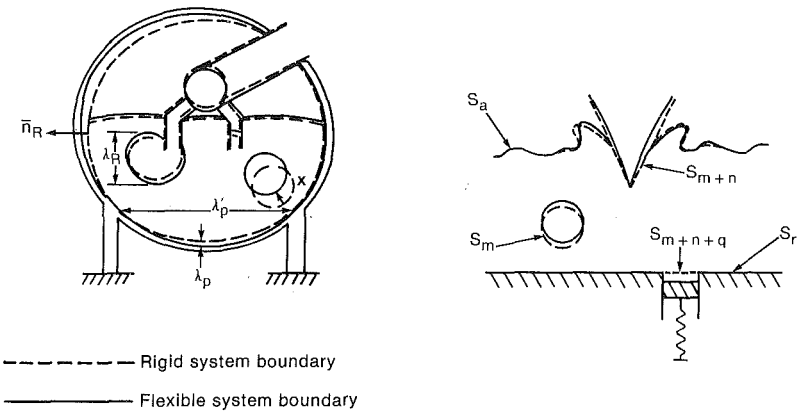


Fig. 1 Typical systems with slow bulk motion and small amplitude high frequency oscillations: (*a*) bounded liquid domain and typical length scales; (*b*) unbounded liquid domain and different types of boundary surface

(Part 2), Joos and Huber (1987) apply the analytical tools we have developed to two sets of experiments conducted in rigid and flexible liquid-filled tanks.

II Theory for Incompressible Liquid

A Liquid Domain. In Fig. (1), one system consists of a liquid totally surrounded by a finite number of surfaces, the other a liquid in an infinite domain surrounding a finite number of bodies. In each case the liquid is totally connected. The boundary surfaces may represent gas cavities or structural elements. Boundaries at which conditions are uniform are termed pressure or displacement *nodes*; boundaries at which conditions are nonuniform are termed *surfaces*. In general, a boundary may undergo two types of motion, "translation" and "oscillation," distinguished by their characteristic times—long for translation, short for vibration.

Consider also a second pair of systems, identical in all respects to the first, except that the surfaces can only translate, not oscillate. The liquid still obeys the usual equations of motion but boundary conditions at the liquid surfaces are modified so as to exclude any oscillatory motion. These systems will be termed *idealized rigid systems*, to be distinguished from the *flexible systems* in which both types of displacement are allowed. The difference between the pressure and velocity fields in the flexible and rigid systems will be termed the *perturbation fields*.

Conservation of mass and momentum of the liquid in the flexible system require

$$\bar{\nabla} \cdot \bar{v} = 0 \tag{1}$$

$$\rho\left[\frac{\partial \bar{v}}{\partial t} + \bar{v} \cdot \bar{\nabla}\bar{v}\right] = -\bar{\nabla}P - \rho g \bar{\nabla} z + \mu \nabla^2 \bar{v} \tag{2}$$

We shall assume here that the liquid motion is incompressible and the liquid viscosity is constant; compressible liquid motion is addressed later.

The dependent variables may be expressed in terms of their rigid and perturbation components:

$$P = P_R + P_P \tag{3}$$

$$\bar{v} = \bar{v}_R + \bar{v}_P \tag{4}$$

where the subscripts R and P denote the rigid and perturbation fields, respectively. The rigid system fields must, by definition, independently obey the equations of mass and momentum conservation:

$$\bar{\nabla} \cdot \bar{v}_R = 0 \tag{5}$$

$$\rho\left[\frac{\partial \bar{v}_R}{\partial t} + \bar{v}_R \cdot \bar{\nabla}\bar{v}_R\right] = -\bar{\nabla}P_R - \rho g \bar{\nabla} z + \mu \nabla^2 \bar{v}_R \tag{6}$$

Table 1 Order of magnitude estimates of velocity terms – incompressible case

$\bar{v}_R \sim V_R$	$\cdot \bar{\nabla}\bar{v}_P \sim \dfrac{\lambda_P}{\lambda'_P \tau_P}$
$\bar{v}_P \sim \dfrac{\lambda_P}{\tau_P}$	$\nabla^2 \bar{v}_P \sim \dfrac{\lambda_P}{\lambda'^2_P \tau_P}$
$\cdot \bar{\nabla}\bar{v}_R \sim \dfrac{V_R}{\lambda_R}$	$\dfrac{\partial \bar{v}_P}{\partial t} \sim \dfrac{\lambda_P}{\tau_P^2}$

The requirement that the rigid system motion be characterized exclusively by low frequency components plainly need not extend to the rigid system pressure, P_R. Equation (6) reveals that the only high frequency rigid system pressure components are spatially uniform throughout the liquid domain.

Substituting equations (3) and (4) into equations (1) and (2), and subtracting from them equations (5) and (6) yields:

$$\bar{\nabla} \cdot \bar{v}_P = 0 \tag{7}$$

$$\rho\left[\frac{\partial \bar{v}_P}{\partial t} + \bar{v}_R \cdot \bar{\nabla}\bar{v}_P + \bar{v}_P \cdot \bar{\nabla}\bar{v}_R + \bar{v}_P \cdot \bar{\nabla}\bar{v}_P\right] = -\bar{\nabla}P_P + \mu \nabla^2 \bar{v}_P \tag{8}$$

Order of magnitude estimates for the velocity terms of equation (8) are given in Table 1, where the symbols represent:

V_R: a speed characterizing the magnitude of the velocity of the rigid system (of order λ_R/τ_R);

λ_R: a typical length over which velocities change in the rigid system (see Fig. 1(*a*)). λ_R is assumed to be comparable to the length by which the liquid in the rigid system is displaced in the time τ_R. (λ_R might be a gas cavity's or structural element's dimension or the size of a container.)

λ_P: a characteristic displacement amplitude of a surface in the flexible system relative to its corresponding position in the rigid system (see Fig. 1(*a*)). λ_P might represent, for example, the amplitude of oscillation of a bubble surface or a structural boundary element.

λ'_P: a typical length over which perturbation velocities change (for example, a dimension of a node or a surface wavelength). This length also defines the characteristic length over which the gradient of the perturbation velocity changes.

τ_R: a time chartacterizing the fluid motion in the rigid system (for example, the time taken by a gas cavity or a structural element to translate through the system).

τ_P: a time characterizing oscillations in the flexible system (for example, a vibration period).

Table 2 Simplifying assumptions on the fluid domain – incompressible case

$\left[\frac{\lambda_R}{c\tau_P}\right]^2 \ll 1$	A
$\frac{\tau_P}{\tau_R} \ll 1$	B
$\frac{\lambda_P}{\lambda_R} \ll 1$	C
$\frac{\lambda_P}{\lambda_P'} \ll 1$	D
$\frac{\lambda_R \tau_P}{\lambda_P' \tau_R} \ll 1$	E
$\frac{\mu \tau_P}{\rho \lambda_P'^2} \ll 1$	F

In Table 2 we set out six order-of-magnitude criteria that are presumed to be satisfied in the following analysis. Criterion (A) permits the liquid flow to be treated as incompressible. Criteria (B) and (C) restate our assumptions that the vibration of the system can be distinguished from its translation by the relatively high frequency and small amplitude of the vibratory motion. Condition (D) requires that the amplitude of oscillation at the boundaries be much smaller than its surface wavelength. Applying condition (B), criterion (E) would require that the typical geometric lengths of the system be comparable to, or smaller than, the boundary wavelengths. Criterion (F) is a requirement of inviscidity of the perturbations.

If criteria (A, B, D, E, and F) are satisfied, equation (8) reduces to:

$$\rho \frac{\partial \bar{v}_P}{\partial t} = -\bar{\nabla} P_P \tag{9}$$

Taking the divergence and applying equation (7), the perturbation pressure field is found to satisfy Laplace's equation:

$$\nabla^2 P_P = 0 \tag{10}$$

At the boundaries equation (9) can be used to relate normal pressure gradients to the local surface displacements x (see Fig. 1(a))

$$\frac{\partial P_P}{\partial n_R} = -\rho \frac{\partial^2 x}{\partial t^2} \tag{11}$$

where the left-hand side represents the components of the pressure gradient normal to the boundary.

The boundary values of equation (10) depend on conditions at the material surfaces of the flexible system. It is, however, computationally convenient to apply these boundary conditions at nonoscillatory control surfaces defined by the corresponding rigid system surfaces. This approximation is reasonable when criteria (C) and (D) of Table 2 are satisfied.

We can now solve equation (10) in the liquid domain D subject to the following boundary conditions:

$$P_P(\bar{r}) = 0 \quad \text{for} \quad \bar{r} \epsilon S_a \tag{12}$$

$$P_P(\bar{r}) = P_{Pj} \quad \text{for} \quad \bar{r} \epsilon \{S_j \,|\, 1 \le j \le m\} \tag{13}$$

$$\frac{\partial P_P}{\partial n_R}(\bar{r}) = 0 \quad \text{for} \quad \bar{r} \epsilon S_r \tag{14}$$

$$\frac{\partial P_P}{\partial n_R}(\bar{r}) = Dp_P(\bar{r}) \quad \text{for} \quad \bar{r} \epsilon \{S_j \,|\, m+1 \le j \le m+n\} \tag{15}$$

$$\frac{\partial P_P}{\partial n_R}(\bar{r}) = Dp_{Pj} \quad \text{for} \quad \bar{r} \epsilon \{S_j \,|\, m+n+1 \le j \le m+n+q\} \tag{16}$$

Here S_a defines free boundaries on which perturbation pressures are zero throughout the transient and S_r defines stationary boundaries which do not oscillate during the transient (see Fig. (1(b)). The pressure node boundaries (areas of uniform—but time-dependent—pressure) are defined by $\{S_j \,|\, 1 \le j \le m\}$; surfaces with spatially varying displacements are $\{S_j \,|\, m+1 \le j \le m+n\}$; and displacement node boundaries (surfaces of uniform—but time-dependent—displacement) are $\{S_j \,|\, m+n+1 \le j \le m+n+q\}$. The left-hand sides of equations (15) and (16) denote outward normal gradient components of the perturbation pressure. We will assume that at least one free surface S_a exists in any system. In open systems (Fig. 1(b)), surfaces at infinity will be assumed to have homogeneous boundary conditions.

We now develop the solution to Laplace's equation (10) with Dirichlet and Neumann type boundary conditions (equations (12–13) and (14–16), respectively) as a superposition of a set of functionals. The functionals $G_j(\bar{r})$ satisfy Laplace's equation in the domain D and also the same types of boundary conditions, but these are homogeneous on all surfaces except S_j.

$$\nabla^2 G_j(B_j \,|\, \bar{r}) = 0 \tag{17}$$

$$G_j(B_j \,|\, \bar{r}_i) = \delta_{ij} B_j(\bar{r}_i) \quad 1 \le i \le m \text{ and } \bar{r}_i \in S_i \tag{18}$$

$$\frac{\partial G_j}{\partial n_R}(B_j \,|\, \bar{r}_i) = \delta_{ij} B_j(\bar{r}_i) \quad \text{for } m+1 \le i \le m+n+q \text{ and } \bar{r}_i \in S_i \tag{19}$$

Here δ_{ij} represents the Kronecker delta. The inhomogeneous boundary condition B_j is arbitrary but continuous over S_j. The functionals G_j will exist provided there exists a solution to Laplace's equation in the domain D with the same boundary condition types. Each G_j is linearly dependent on the distribution of its inhomogeneous boundary condition. The general solution is

$$P_p(\bar{r}) = \sum_{j=1}^{m+n+q} G_j(B_j \,|\, \bar{r}) \tag{20}$$

Replacing the boundary conditions defined in equations (12–16) for B_j in equation (20) and invoking linearity of G_j with respect to B_j yields

$$P_p(\bar{r}) = \sum_{j=1}^{m} g_j^o(\bar{r}) P_{pj} + \sum_{j=m+1}^{m+n} G_j(Dp_p(\bar{r}_j) \,|\, \bar{r}) + \sum_{j=m+n+1}^{m+n+q} g_j^o(\bar{r}) Dp_{pj} \tag{21}$$

where

$$g_j^o(\bar{r}) = G_j(1.0 \,|\, \bar{r}) \tag{22}$$

Taking the normal pressure gradient at a point on the control surface yields

$$\frac{\partial P_P}{\partial n_R}(\bar{r}) = \sum_{j=1}^{m} h_j^o(\bar{r}) P_{Pj} + \sum_{j=m+1}^{m+n} h^o(Dp_p(\bar{r}_j) \,|\, \bar{r}) + \sum_{j=m+n+1}^{m+n+q} h_j^o(\bar{r}) Dp_{Pj} \tag{23}$$

where

$$h_j^o(\bar{r}) = \frac{\partial g_j^o}{\partial n_R}(\bar{r}) \tag{24a}$$

and

$$h^o(Dp_p(\bar{r}_j)\,|\bar{r}) = \frac{\partial G_j}{\partial n_R}\,(Dp_p(\bar{r}_j)\,|\bar{r}) \tag{24b}$$

Equation (23) is integrated along the liquid-gas interfaces to yield an average normal pressure gradient $\{Dp_{Pi}, 1 \le i \le m\}$:

$$Dp_{pi} = \sum_{j=1}^{m} h_{ij}P_{Pj} + \sum_{j=m+1}^{m+n} h_i(Dp_p(\bar{r}_j)) + \sum_{j=m+n+1}^{m+n+q} h_{ij}Dp_{Pj} \tag{25}$$

where

$$h_{ij} = \frac{1}{A_i}\int_{\bar{r}\epsilon S_i} h_j^o(\bar{r})\,dS_i \tag{26a}$$

and

$$h_i(Dp_p(\bar{r}_j)) = \frac{1}{A_i}\int_{\bar{r}\epsilon S_i} h^o(Dp_p(\bar{r}_j)\,|\bar{r})\,dS_i \tag{26b}$$

Here A_i is the area of the interface of the control surface S_i.

Average perturbation pressures P_{Pi} on surfaces of uniform displacement ("lids") are determined by integrating equation (21) over these surfaces. For $m+n+1 \le \ell \le m+n+q$.

$$P_{P\ell} = \sum_{j=1}^{m} g_{\ell j}P_{Pj} + \sum_{j=m+1}^{m+n} g_\ell(Dp_p(\bar{r}_j)) + \sum_{j=m+n+1}^{m+n+q} g_{\ell j}Dp_{Pj} \tag{27}$$

where

$$g_{\ell j} = \frac{1}{A_\ell}\int_{\bar{r}\epsilon S_\ell} g_j^o(\bar{r})\,dS_\ell \tag{28a}$$

and

$$g_\ell(Dp_p(\bar{r}_j)) = \frac{1}{A_\ell}\int_{\bar{r}\epsilon S_\ell} G_\ell(Dp_p(\bar{r}_j)\,|\bar{r})\,dS_\ell \tag{28b}$$

Next we combine the previous equations to construct an operator relating average and local boundary perturbation pressures to average and local boundary perturbation pressure normal gradients. Equation (25) expresses the average normal pressure gradients on each of the pressure nodes in terms of the boundary conditions; equation (27) relates the average perssures on the displacement nodes to the boundary conditions; equation (21) provides a functional relation of the perturbation pressures on the displacement surfaces to the boundary conditions. From equation (25) we can determine the pressure of the pressure nodes in terms of the appropriate local or average boundary normal pressure gradients by inverting the matrix formed by the coefficients h_{ij}, $\{1 \le (i, j,) \le m\}$ and rearranging. Replacing and rearranging the pressure-node pressures in terms of the normal pressure gradients in equations (21), (25), and (27), we obtain a linear expression of pressure in terms of normal pressure gradient.

$$p_P = H\,dp_P \tag{29}$$

Here p_P represents the local boundary surface pressure (except at the displacement nodes where p_P is the average surface pressure), dp_P represents the boundary surface normal pressure gradient (except at the pressure nodes where it is the average normal pressure gradient), and H is a linear operator determined from equations (21), (25), and (27). Note that when there are no displacement surfaces equation (29) represents uniform and average pressures as a set of linear combinations of uniform and average normal pressure gradients. Equation (29) then represents a matrix operation.

The flexible system may have more than one free surface, and each free surface may have a different forcing pressure applied to it. If the forcing pressures contain different high frequency components, then a new perturbation field must exist that satisfies the high frequency pressure boundary conditions at these surfaces. This perturbation field is not included in the solution leading to equation (29) because, according to equation (12), the perturbation pressure boundary conditions at free surfaces are homogeneous. In addition, these boundary conditions are independent of flexible system conditions. Because this perturbation component is independent of the flexible system response, it will also be present in any practical rigid system that includes high frequency forcing pressure components at the free surfaces. It is appropriate, therefore, to include this perturbation field as part of the rigid system fields. The practical rigid system pressure's high frequency component then becomes a perturbation pressure field satisfying Laplace's equation (10) with the following boundary conditions: high frequency forcing pressure components (the forcing pressures referred to here and later on are assumed to be imposed on the non-liquid side of the boundary) at the free surfaces, zero average normal pressure gradient and uniform (but unknown) pressure at each pressure node, and zero normal pressure gradient at all liquid-structure interfaces.

B Constitutive Equations for the Boundaries. Boundary conditions at the pressure and displacement nodes and displacement surfaces, equations (12–16), will depend on the constitutive relations governing the surfaces from the non-liquid side. A crucial step in the analysis is decoupling the constitutive equations into high and low frequency terms. Consistent with our assumption that only low frequency liquid motions are permitted in the idealized rigid system, the low frequency terms in the decoupled constitutive equations provide the boundary conditions for the rigid system while the high frequency terms provide the boundary conditions for the flexible system.

We consider here three illustrative constitutive relations for the boundaries. Decoupling the equations into their high and low frequency components requires linearity; apart from this requirement the method is general and can readily be extended to other types of boundaries.

1. *Pressure nodes* are formed by gas cavities in the liquid. We will assume here that the pressure in such a cavity in the flexible system depends on the cavity volume and the gas properties, as well as on an arbitrary forcing pressure:

$$P_j = f_j(V_{Rj}) + \frac{df_j}{dV_j}\,(V_{Rj})A_j y_{Pj} + P_{fj},\quad 1 \le j \le m \tag{30}$$

where:

P_j is the total pressure at the jth pressure node.

P_{fj} is any specified pressure forcing function at the node. It might represent boundary pressures due to gas absorption or condensation at the liquid-gas interface or independently imposed acoustic waves in the cavity.

$f_j(V_j)$ is a function of the cavity volume V_j. It might model the isentropic expansion of a bubble. The linearized approximation of $f_j(V_j)$ is justified by assumption (*C*) in Table 2.

y_j is the average perturbation displacement over the jth cavity's surface area, A_j.

V_{Rj} is the volume of the cavity in the rigid system.

The first term on the right-hand side of equation (30) represents the pressure in the rigid system as a function of rigid system cavity volume. Since gas cavity volumes change only in response to bulk liquid deformations, this term can involve only low frequency pressure components. On the other hand, the second term depends on the average perturbation displacement, and can therefore contain only high frequency components. Thus:

$$P_j^{\ell} = P_{Rj}^{\ell} = f_j(V_{Rj}) + P_{fj}^{\ell} \tag{31}$$

$$P_j^{h} = P_{Rj}^{h} + P_{Pj} = k_{bj} y_{Pj} + P_{fj}^{h} \tag{32}$$

where $\ell \le j \le m$ and the superscripts ℓ and h denote low and high frequency components. The coefficient k_{bj} replaces the first derivative of $f_j(V_j)$ multiplied by the cavity area A_j in equation (25). The total pressure P_j has been separated into its low and high frequency components, P_j^{l} *and* P_j^{h}. The high frequency component, equation (32), may be written as

$$k_{bj} y_j = P_j^{h} - P_{fj}^{h} \tag{33}$$

2. *Displacement surfaces* are governed by a second order differential equation. By assumptions (*C*) and (*D*) of Table 2 the load-displacement relation for the perturbation displacements should be linearly elastic. We assume, in addition, that the structure in the rigid system is also linearly elastic, that shear stresses are absent, and that displacements are exclusively perpendicular to the structural surfaces. The differential equations for the displacement surfaces can then be written in the form:

$$m_s \ddot{y} + K_s y = P - P_f, \quad m+1 \le j \le m+n \tag{34}$$

Here m_s represents the local mass per unit surface area, K_s is a stiffness linear integro-differential operator of the local surface displacements y, and P and P_f are local total and forcing pressures. Equation (34) is the general constitutive equation for a shell. The analysis of other linearly elastic structures should not differ significantly.

Equation (34) may be decoupled into high and low frequency terms as follows:

$$K_s y_R = P_R^{\ell} - P_f^{\ell} \tag{35}$$

$$m_s \ddot{y}_P + K_s y_P = P^{h} - P_f^{h} \tag{36}$$

where y_R and y_P are the local rigid and perturbation surface displacements of the structures. Equation (35) will be satisfied by the idealized rigid system, but because the low frequency response will be quasi-static, the acceleration term is neglected. This convenient simplification does not affect later steps of the analysis.

3. *Displacement nodes* ("lids") are special cases of displacement surfaces for which displacement is uniform across the surface. Here the stiffness operator represents a spring constant:

$$m_{Lj} \ddot{y}_j + k_{Lj} y_j = P_j - P_{fj}, \quad m+n+1 \le j \le m+n+q \tag{37}$$

where m_{Lj} and k_{Lj} are the jth lid mass and spring constant divided by its surface area, y_j is the node's displacement and P_j and P_{fj} the average total and forcing pressure across the lid. Equation (37) can be decoupled in the same way:

$$k_{Lj} y_{Rj} = P_{Rj}^{\ell} - P_{fj}^{\ell} \tag{38}$$

$$m_{Lj} \ddot{y}_{Pj} + k_{Lj} y_{Pj} = P_j^{h} - P_{fj}^{h} \tag{39}$$

where the assumption of negligible inertia at low frequency has again been made.

4. *Combined Boundary Conditions.* Combining equations (3), (33), (36), and (39) yields a complete set of boundary conditions to the perturbation equations:

$$m \ddot{y}_P + K y_P = p_R^{h} + p_P - p_f^{h} \tag{40}$$

In equation (40), y_P represents average surface perturbation displacement for the pressure nodes, or local displacements for the displacement surfaces and nodes; m is zero for the pressure nodes, the local representative mass per unit surface area for the displacement surfaces and the average mass per unit area for the displacement nodes; K is a linear integro-differential operator of displacement which is a constant for pressure and displacement nodes; p_R^{h}, and p_f^{h} are local rigid and forcing high frequency pressure components (except for displacement nodes, where average values are applied) and p_P is the perturbation pressure of equation (29).

C System Perturbation Equations. The system perturbation equations are obtained by combining the fluid domain perturbation equation, equation (29), with the boundary condition relations, equation (40). Equation (11) must first be rewritten in terms of the displacements y_P:

$$dp_P = -\rho \ddot{y}_P \tag{41}$$

Here equation (11) has been averaged over the surface of the pressure nodes. Combining equations (29) and (41) yields a relationship between accelerations and pressures at the boundaries:

$$p_P = -\rho H \ddot{y}_P \tag{42}$$

Combining equations (40) and (42), we obtain system equations in terms of perturbation displacement or perturbation pressure

$$(m + \rho H)\ddot{y}_P + K y_P = p_R^{h} - p_f^{h} \tag{43}$$

$$(m H^{-1} + \rho)\ddot{p}_P + K H^{-1} p_P = -\rho(\ddot{p}_R^{h} - \ddot{p}_f^{h}) \tag{44}$$

In deriving equations (43) and (44) we have assumed, for simplicity, that all structural nodes and surfaces are wetted. To include the nonwet case is simple but the notation becomes more cumbersome. In addition, time derivatives of the Laplace operator H have been neglected. This is valid when the time scales characterizing the idealized rigid system, τ_R, are large compared to those characterizing the perturbation oscillations, τ_P, as required by condition (*B*), Table 2. We have also assumed that H^{-1}, the inverse operator to H, exists and is unique. This assumption is reasonable because H is derived from a Laplace operator, having a unique inverse.

In the absence of displacement surfaces, equations (43) and (44) reduce to linear combinations of the perturbation displacements or pressures at the pressure and displacement nodes. There is one such linear combination for each node, their synthesis being a matrix differential equation of dimension equal to the number of nodes. Thus, each node contributes one and only one natural mode and frequency to the system.

Clearly, equations (43) or (44) can be used to predict flexible system behavior once the rigid system's pressure history, the liquid's gross distribution history, and the flexible system boundary constitutive relations are known. It is not so obvious that the same equations can also be used to predict the rigid system's pressure if, besides knowing the gross liquid distribution history and the flexible system's boundary constitutive relations, we also have total pressure or displacement history data at all boundaries. One method of predicting the rigid pressure histories at the boundaries would be to transform all total boundary pressure data into total boundary displacement data via the constitutive equations (30), (34), and (37); then these displacements would be filtered to obtain their high frequency or perturbation components; the rigid system pressure is then predicted from equation (43). Ultimately, by combining these two exercises, it is possible to predict a flexible system's behavior from data of another system of different flexibility.

III Extension to a Compressible Liquid

When the liquid is compressible the liquid domain equations, together with additional simplifying assumptions, can be obtained in much the same way as before.

The liquid domain in the flexible and rigid systems must obey the compressible mass and momentum conservation equations and the isentropic relation:

$$\frac{\partial \rho}{\partial t} + \bar{\nabla} \cdot (\rho \bar{v}) = 0 \tag{45}$$

Table 3 Order of magnitude estimates of velocity and density terms—compressible case

$\bar{v}_R \sim V_R$	$\bullet \bar{\nabla} \bar{v}_P \sim \frac{\lambda_P}{\lambda_P' \tau_P}$
$\bar{v}_P \sim \frac{\lambda_p}{\tau_P}$	$\nabla^2 \bar{v}_P \sim \frac{\lambda_P}{\lambda_P'^2 \tau_P}$
$\bullet \bar{\nabla} \bar{v}_R \sim \frac{V_R}{\lambda_R}$	$\frac{\partial \bar{v}_P}{\partial t} \sim \frac{\lambda_P}{\tau_P^2}$
$\frac{\partial \rho_P}{\partial t} \sim \frac{\rho_P}{\tau_P}$	$\bar{\nabla} \bullet \bar{v}_P \sim \frac{\lambda_p}{\lambda_P' \tau_P}$
$\bar{\nabla} \rho_R \sim \frac{\rho_R}{\lambda_R}$	$\bar{\nabla} \rho_P \sim \frac{\rho_P}{\lambda_P'}$
$\bar{\nabla} \rho_0 \sim \rho_0 \frac{g}{c^2}$	

$$\rho\left[\frac{\partial \bar{v}}{\partial t} + \bar{v} \bullet \bar{\nabla} \bar{v}\right] = -\bar{\nabla} P - \rho g \bar{\nabla} z + \mu \nabla^2 \bar{v} \tag{46}$$

$$\left.\frac{\partial P}{\partial \rho}\right|_s = c^2 \tag{47}$$

Excursions in pressure and liquid densities will be assumed sufficiently small for c^2 to remain constant. Pressure, density, and velocity fields are again decomposed into rigid and perturbation components. In addition, the rigid system total pressure and density can be separated into a hydrostatic and a rigid system excursion component:

$$P = P_o + P_R + P_P \tag{48}$$

$$\rho = \rho_o + \rho_R + \rho_P \tag{49}$$

$$\bar{v} = \bar{v}_R + \bar{v}_P \tag{50}$$

Here, terms subscripted o represent static components; those subscripted R represent rigid system excursions from the corresponding static terms. The isentropic relation requires the rigid and perturbation pressure components to be proportional to the rigid and perturbation density components.

We require both the rigid and flexible systems to obey the conservation and isentropic relations independently. The perturbation equations are reduced to the wave equation through assumptions similar to those made before in analyzing incompressible perturbations:

$$\frac{1}{c^2} \frac{\partial^2 P_P}{\partial t^2} = \nabla^2 P_P \tag{51}$$

Table 3 provides order of magnitude estimates characterizing different variables and operations, which are similar but more complex than those of Table 1. As before, the perturbation equation will accurately predict flexible system perturbations only when certain order of magnitude criteria are satisfied. The new criteria are (*A*)-(*J*) in Table 4; note that the previously adopted conditions (*B*)-(*F*), Table 2, are included here. At the boundaries, the normal perturbation pressure gradient is related to boundary accelerations by

$$\frac{\partial P_P}{\partial n_R} = -\rho_o \frac{\partial^2 x}{\partial t^2} \tag{52}$$

The first noteworthy feature of the perturbation equation for a compressible liquid, equation (51), is that it has an inherent characteristic time, λ_R/c. Since the shortest time period characterizing the rigid system response must be large compared with the longest period characterizing the perturbation fields, the rigid system flow fields must have the

Table 4 Simplifying assumptions on the fluid domain—compressible case

$\frac{\rho_0}{\rho_P} \frac{\lambda_P g}{c^2} \ll 1$	*A*
$\frac{\lambda_P}{\lambda_R} \frac{\rho_R}{\rho_P} \ll 1$	*B*
$\frac{\lambda_R}{\lambda_P'} \frac{\tau_P}{\tau_R} \ll 1$	*C*
$\frac{\lambda_P}{\lambda_P'} \ll 1$	*D*
$\frac{\tau_P}{\tau_R} \ll 1$	*E*
$\frac{\rho_P}{\rho_o} \ll 1$	*F*
$\frac{\rho_R}{\rho_o} \ll 1$	*G*
$\frac{\rho_P \lambda_R \tau_P^2}{\rho_o \lambda_P \tau_R^2} \ll 1$	*H*
$\frac{\rho_P g \tau_P^2}{\rho_o \lambda_P} \ll 1$	*I*
$\frac{\mu \tau_P}{\rho_o \lambda_P'^2} \ll 1$	*J*
$\frac{\lambda_P}{\lambda_R} \ll 1$	*K*

characteristics of an incompressible flow. All compressibility effects in the liquid must therefore be part of the perturbation fields.

As before, the perturbation equation can be reduced to a boundary value problem. The formulation is simplified considerably by imposing conditions (*E*) and (*K*), Table 4, which permit us to solve the wave equation for stationary boundaries. Using Morgans's (1930) formulation, the perturbation pressures at the boundaries at any time t can be expressed as a linear operation of the boundary condition history (perturbation pressures and normal pressure gradients at all times prior to t).

$$P_P = H\, dP_P \tag{53}$$

we have assumed, for simplicity, that initial perturbation pressures and their time derivatives throughout the liquid domain are homogeneous. In equation (53), boundary conditions at the liquid-structure surfaces are of the Neumann type, and those at the gas-liquid surfaces are Dirichlet boundary conditions. At rigid walls and free surfaces the boundary conditions are homogeneous.

Constitutive relations for the boundary surfaces can be defined as before, modified if necessary to account for compressibility effects in gaseous cavities. The high frequency components of the constitutive relations can be summarized once again in the form:

$$M\ddot{x} + Kx = P^h - P_f^h \tag{54}$$

For structural boundaries the constitutive equations here will be no different from those used previously in the incompressible analysis. For gaseous boundaries, however, the constitutive equation will normally depend on acoustic effects within the gas itself. The linear operator M for such bound-

aries will therefore be similar in form to equation (53), with the instantaneous boundary pressures in the gas depending on the prior history of boundary pressures in the cavity.

The high frequency components of the constitutive relations are combined with the boundary value expression of the perturbation equation to form a system equation:

$$(M+\rho\ H)\ddot{x}+Kx=P_R^h-P_f^h \tag{55}$$

where we have used equation (52). If initial perturbation pressures throughout the domain are inhomogeneous, additional terms will appear on the right-hand side of this expression.

The system equation, independent of the rigid system pressures at the boundaries, can be used to predict the perturbation pressures and displacements in a system of given flexibility. Required as inputs are the gross liquid distribution, forcing pressures at the boundaries, and initial conditions.

We identify a practical rigid systems pressure high frequency component as the perturbation field component which is independent of system flexibility. Thus, the practical rigid system high frequency component satisfies the wave equation. It has zero normal pressure gradient boundary conditions at the walls, and at the free surfaces the pressure is the high frequency component of the forcing pressure function. At other gas-liquid surfaces, the boundary condition depends on the constitutive relation of the gas cavity: if acoustic effects are not important in the cavity, then a uniform pressure with zero average normal pressure gradient must be applied; alternatively, the normal pressure gradient is homogeneous.

Rigid system data can be extracted from a flexible system as before by combining equations (48) and (52)–(53) yielding:

$$P_o+P_R=P+\rho_o H\ddot{x} \tag{56}$$

Thus, the rigid system pressure can be obtained once the liquid operator H and the total pressure P and perturbation accelerations at the boundaries have been identified.

IV Conclusion

The analysis of coupled gas-liquid-structure flow fields by computation or experiment is difficult. When simplifying criteria that we have formulated are satisfied, the analysis can be decomposed into two simpler ones. The first prong of the analysis focuses on the flow fields created by the gross deformations of the liquid, fields that are governed by the full, nonlinear equation of motion for the bulk flow. The second prong addresses the superposed perturbation fields, governed by the linear perturbation equations we have formulated.

Based on this analysis, a rigid system flow field can be computed, measured directly, or extracted from data obtained in a flexible system. The rigid system characteristics can then be used as inputs to compute flow fields in a system of different flexibility. Applications of this theory are given in the companion paper.

Acknowledgments

This work was sponsored by the United States Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, under Grant No. NRC-G-04-81-028. Part of the clerical and art work was performed by personnel of the employers of one of the authors (F.M. Joos).

References

Antony-Spies, P., 1979, "Fluid-Structure Interaction in the Pool of Pressure Suppression Systems Containing Oscillating Bubbles," *Trans. of the 5th Int. Conf. on Struct. Mech. in Reactor Technology*, Vol. B, paper B7/3.

Belytschko, T., and Schumann, U., 1980, "Fluid-Structure Interactions in Light Water Reactor Systems," *Nuclear Eng. and Design*, Vol. 60, pp. 173–195.

Joos, F. M., 1982, Ph.D. Thesis, Dept. of Mech. Eng., Massachusetts Institute of Technology, Cambridge, MA.

Joos, F. M., and Huber, 1987, "Coupled Gas-Liquid-Structure Systems: Part 2 – Applications," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 942–950.

Kana, D. D., Glaser, R. F., Eulitz, W. K., and Abramson, H. N., 1968, "Longitudinal Vibration of Spring Supported Membrane Shells Containing Liquid," *J. Spacecraft*, Vol. 5, pp. 189–196.

Kalumuck, K. M., and Huber, P. W., 1981, "A Perturbation Analysis of Fluid-Structure Interactions in a Model Test System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 7–14.

Morgans, W. R., 1930, "The Kirchhoff Formula Extended to Moving Surfaces," *Phil. Mag.*, Vol. 9, pp. 141–161.

Sonin, A. A., 1980, "Rationale for a Linear Perturbation Method for the Flow Field Induced by Fluid-Structure Interactions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 725–728.

Coupled Gas-Liquid-Structure Systems: Part 2—Applications[1]

F. M. Joos
Assoc. Mem. ASME
Creare, Inc.
Hanover, NH 03755

P. W. Huber
Science Concepts, Inc.,
Washington, DC 20003

In the companion paper, Joos and Huber (1987) have developed an analysis that applies to certain types of systems in which a hydrodynamic transient creates pressure forces dependent on the flexibilities of the liquid boundary surfaces. To test the analysis we performed two pairs of experiments. Each pair involved identical experiments conducted in two geometrically identical cylindrical tanks. One tank was rigid, the other flexible. In the first set of experiments a gas bubble oscillated freely in a container partially filled with liquid. We compared experimentally measured natural frequencies with those predicted theoretically. In the second set of experiments a complex hydrodynamic transient was produced by injecting air into a tank partially filled with liquid. Using data from the rigid system experiments we predict the pressures and wall displacements in the flexible system and compare these predictions with measurements made in the flexible test tank. The inverse problem—predicting the pressures in the rigid tank from data obtained in the flexible tank—is also analyzed using the theoretical tools we have developed, and again predictions are compared with experiment.

I Introduction

In the companion paper, Joos and Huber (1987) develop theoretical tools to predict pressures and displacements in a system consisting of a liquid undergoing a hydrodynamic transient in a flexible system, provided data is available from an identical transient in a geometrically identical system of different flexibility. At the heart of the analysis is the rigid system: a system that undergoes the same gross liquid redistribution and in which all nonfree surface boundaries displace but do not oscillate. In this paper we apply these analytical tools to two sets of experiments we conducted in rigid and flexible liquid-filled tanks. The first set of experiments examines the natural frequencies of vibrating bubbles. In the other set, involving a "poolswell" transient, first rigid system experimental data is used to predict the flexible system's behavior, then we address in the inverse problem of predicting the rigid system's behavior from the flexible system's experimental data. Throughout this article we will freely apply the concepts, definitions, tables, and equations of the companion paper which we will refer to as "Part 1."

II Free Vibrations

When the rigid and forcing pressures are zero, equations (43) and (44) of Part 1 are homogeneous. This greatly simplified case, representing free vibrations, provides a useful test of the model and illustrates its application. In this section we summarize experiments and analyses of the coupled, free vibrations of a rigid and a flexible tank containing liquid and a submerged bubble of gas.

A Experiments. Two geometrically identical but structurally different cylindrical tanks were tested. The first (Fig. 1(*a*)) was made of rigid acrylic. The second, similar to that used by Kana et al. (1968), had thin polyester walls (Mylar, type D) and a rigid flat bottom. This tank is shown in Fig. 1(*b*) (set up for the second series of experiments). The flexible shell was stiffened by acrylic ribs, which were rigid, stationary, and chemically bonded to the polyester shell. Each test system was partially filled with water, as shown in Fig. 1(*a*). In successive tests, the height of the column of water was varied between 102 mm and 254 mm in 25.4 mm increments.

At the beginning of each experiment a thin rubber balloon, filled with 8.2 ml of air, was suspended in the liquid on the axis of the cylinder approximately 50 mm above the base. The initial overpressure in the balloon was about 5 kPa. The balloon was burst using a metal spike. By means of high speed films, we obtained sequences of photographs of bubbles so formed in the rigid test tank during the interval of interest here. The vertical displacement and gross deformations of the bubble were seen to be negligible. Further experimental details are provided by Joos (1982).

The pressure transient immediately following the rupture of the balloon was recorded at the center of the rigid bottom plate. Figure 2 shows typical pressure histories obtained in the rigid and flexible tanks tested under otherwise identical conditions. Power spectral densities (PSDs) of the pressure records were calculated from fast Fourier transforms of the individual pressure signals. Figure 3 shows average and envelopes of ten

[1]The work presented in this article was performed while the Authors were at the Massachusetts Institute of Technology, Cambridge, MA.

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 17, 1986; final revision July 9, 1987.

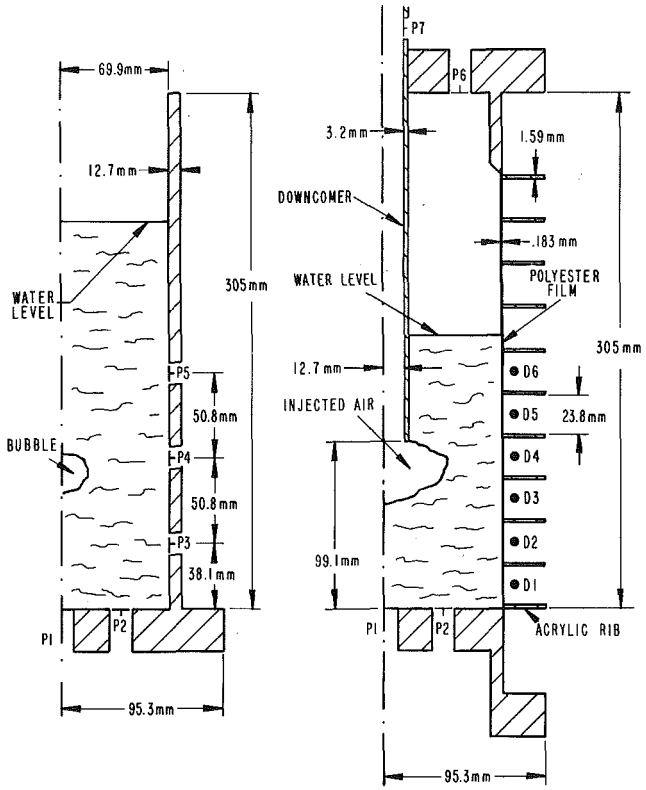


Fig. 1 Schematics of experimental systems: (*a*) rigid tank, set up for free vibrations experiment; (*b*) flexible tank, set up for forced transient experiment. Key: *P*1–*P*7, pressure probes; *D*1–*D*6, displacement probes.

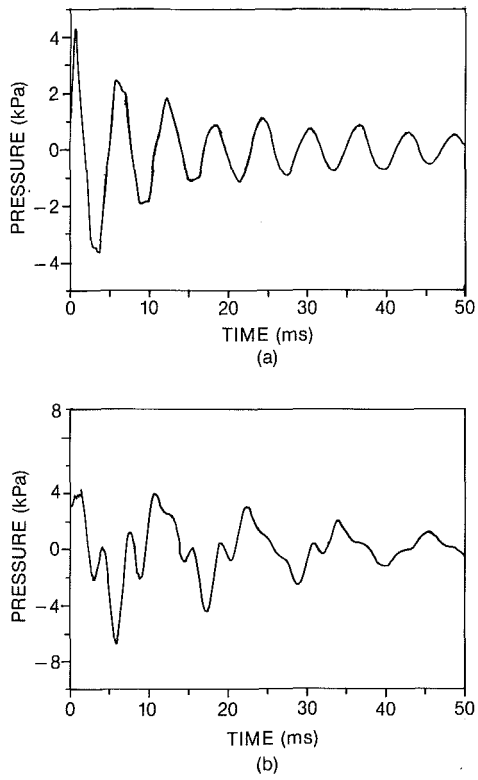


Fig. 2 Typical pressure histories (water depth = 254 mm): (*a*) rigid tank; (*b*) flexible tank

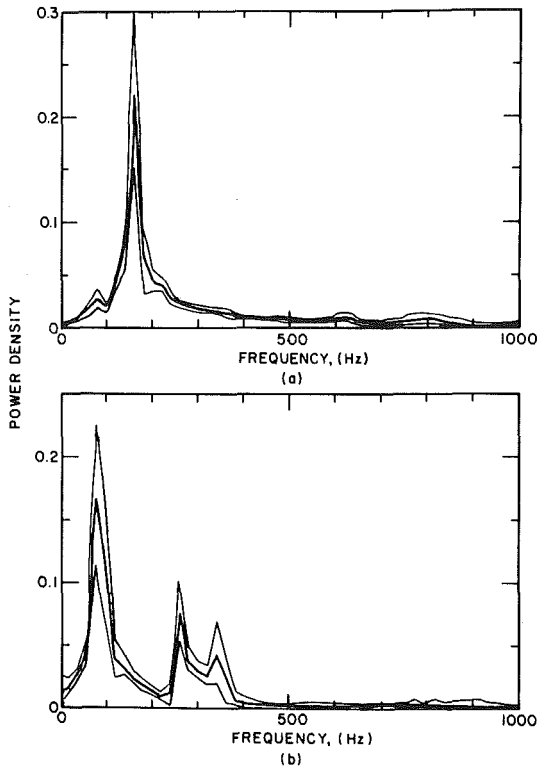


Fig. 3 Pressure trace power spectral densities (water depth = 254 mm): (*a*) rigid tank; (*b*) flexible tank

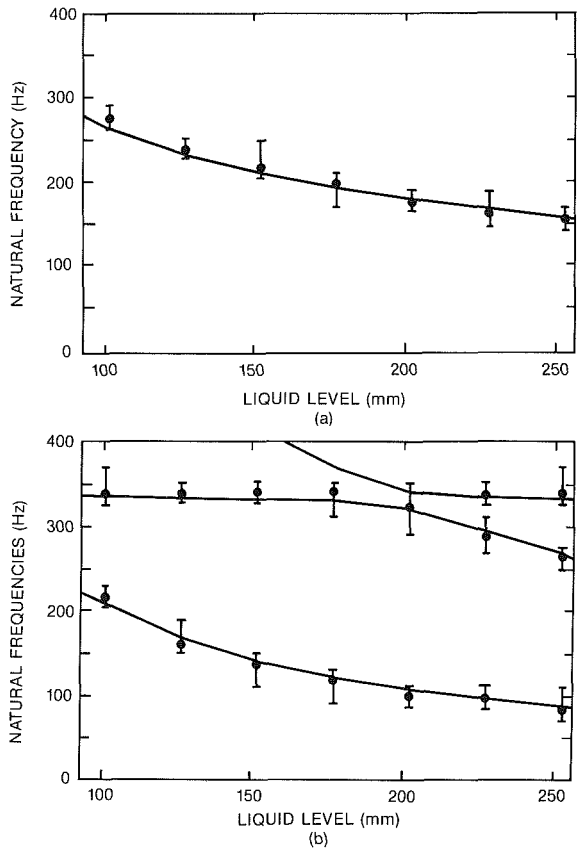


Fig. 4 Comparison of experiment with theory: (*a*) rigid system; (*b*) flexible system. Circles represent experimental average values.

such PSDs for repetitions of one of the experiments in each of the two systems. The results for the rigid structure clearly show a single system natural frequency, while the data for the geometrically identical flexible structure reveal several natural frequencies. none of which coincide with the natural frequency of the rigid-walled system.

Joos (1982) includes pressure traces and PSDs obtained in

Table 1 Verification of simplifying assumptions—free vibrations case

Estimated Characteristic Values from Experiment

$\lambda_R \simeq 25.4$ mm
$\lambda'_P \simeq 25.4$ mm
$\lambda_P \simeq 0.15$ mm
$\tau_P \simeq 10$ ms
$\tau_R \simeq 200$ ms

$c \simeq 1500$ m/s
$c_g \simeq 330$ m/s
$\mu \simeq 10^{-3}$ kg/ms
$\rho \simeq 1000$ kg/m^3
$m_s \simeq 0.26$ kg/m^2
$k_s \simeq 2\times10^8$ Pa/m

Conditions of Table 1

$\left(\frac{\lambda_R}{c\tau_P}\right)^2 \simeq 3\times10^{-6}$ (A)

$\frac{\tau_P}{\tau_R} \simeq 5\times10^{-2}$ (B)

$\frac{\lambda_P}{\lambda_R} \simeq 6\times10^{-3}$ (C)

$\frac{\lambda_P}{\lambda'_P} \simeq 6\times10^{-3}$ (D)

$\frac{\lambda_R\tau_P}{\lambda'_P\tau_R} \simeq 5\times10^{-2}$ (E)

$\frac{\mu\tau_R}{\rho\lambda'^2_P} \simeq 2\times10^{-5}$ (F)

Quasi-Static Conditions

(a) gas cavity

$\left(\frac{\lambda_R}{\tau_P C_g}\right)^2 \simeq 6\times10^{-5}$

(b) structure

$\frac{m_s}{\tau_P^2 k_s} \simeq 1.3\times10^{-5}$

Table 2 Verification of simplifying assumptions—forced transient case

Estimated Characteristic Values from Experiment

$\lambda_R \simeq 70$ mm
$\lambda'_P \simeq 25.4$ mm
$\lambda_P \simeq 0.1$ mm
$\tau_P \simeq 10$ ms
$\tau_R \simeq 400$ ms
$k_g \simeq 8\times10^6$ Pa/m

$c \simeq 1500$ m/s
$c_g \simeq 330$ m/s
$\mu \simeq 10^{-3}$ kg/m^3
$\rho \simeq 1000$ kg/m^3
$m_s \simeq 0.26$ kg/m^2
$k_s \simeq 2\times10^8$ Pa/m

Conditions of Table 2

$\left[\frac{\lambda_R}{c\tau_P}\right]^2 \simeq 9\times10^{-5}$ (A)

$\frac{\tau_P}{\tau_R} \simeq 10^{-2}$ (B)

$\frac{\lambda_P}{\lambda_R} \simeq 10^{-3}$ (C)

$\frac{\lambda_P}{\lambda'_P} \simeq 4\times10^{-3}$ (D)

$\frac{\lambda_R\tau_P}{\lambda'_P\tau_R} \simeq 3\times10^{-2}$ (E)

$\frac{\mu\tau_R}{\rho\lambda'^2_P} \simeq 8\times10^{-6}$ (F)

Quasi-Static Conditions

(a) gas cavity

$\left[\frac{\lambda_R}{c_g\tau_P}\right]^2 \simeq 2\times10^{-3}$

(b) structure

$\frac{m_s}{k_s\tau_P^2} \simeq 5\times10^{-5}$

Condition for Homogeneous b.c. at Gas Cavity

$\frac{k_g}{k_s} \simeq 4\times10^{-2}$

experiments performed at the other water depths in the two test tanks. Natural frequencies obtained from these experiments are plotted as a function of pool depth in Fig. 4. The error bars are due largely to the course (20 Hz) numerical resolution of the fast Fourier transforms.

B Analysis. The solid lines in Fig. 4 are the natural frequencies predicted as the eigenvalues of equations (43) and (44) of Part 1, applied to the two test systems, in question. A brief description of the theoretical analysis of these two test systems follows.

Table 1 demonstrates that the criteria of Table 2 of Part 1 are met in the experiments. In addition, it is easy to show that the bubble vibrations are quasistatic and that for our flexible test tank the inertia of the flexible polyester wall can be neglected. Finally, over the time periods of interest in our experiments, the liquid configuration was essentially stationary. The linear operators of equations (43) and (44) of Part 1 are therefore constant in time. Equation (43) of Part 1 thus reduces to:

$$\rho H\ddot{y}_P + Ky_P = 0 \tag{1}$$

For the purpose of numerical calculations we approximate equation (1) by a set of linear equations:

$$\rho[K]^{-1}[H]\ddot{\mathbf{y}}_P + y_P = \mathbf{0} \tag{2}$$

where $[K]$ and $[H]$ represent the linear operators K and H and $\mathbf{y}_P$ is an array representing discretized effective displacements. The term representing the bubble in the $\mathbf{y}_P$ array measures the average displacement of the bubble surface. The implementation of the theoretical model for our axisymmetric experimental systems is thus reduced to calculating the matrices $[H]$ and $[K]^{-1}$ and the eigenvalues of their product.

To calculate the liquid matrix $[H]$ we solved Laplace's equation, equation (10) of Part 1, using a boundary integral method. The bubble surface was idealized as a sphere of radius $r_b = 12.5$ mm, located at the axis of the cylinder 60 mm above the bottom of the plate. All computational nodes on the bubble surface are reduced to a single row and column in $[H]$, operating on the variable bubble pressure and average normal pressure gradient over the bubble surface.

The top boundary of the liquid pool is a free surface and the bottom a rigid surface. In the rigid-walled test system, the $[H]$ matrix, therefore, has only a single term representing the gas cavity; all the other boundary conditions are homogeneous. In the flexible system the walls of the cylinder were displacement surfaces, rigid at the location of the ribs. Additional details of the computation appear in Joos (1982).

To calculate the influence matrix $[K]^{-1}$ for each test system, we require a pressure-displacement relations for the bubble and (for the flexible-wall system only) a pressure-displacement relation for the wall. The influence coefficient for a spherical gaseous bubble is easily shown to be $r_b/3\gamma p_o$, where $\gamma = 1.4$ and p_o is atmospheric pressure. This is the only source of flexibility in the rigid-walled test tank; here, therefore, the influence matrix collapses to a single number.

For the flexible structure, the bubble influence coefficient provides only one row and column of the influence matrix; other terms in the matrix account for the flexibility of the wall. We define shell rings consisting of the shell segment between each pair of ribs. Because the rings are fixed at the ribs, each ring vibrates independently. The linear operator K_s describing the static displacement of a cylindrical shell under an axisymmetric load is, following Timonshenko and Woinowsky-Krieger (1959):

$$K_s = D_s\left(\frac{d^4}{dz^4} + 4\beta^4\right) \tag{3}$$

where

$$\beta^4 = \frac{3(1-\nu_s^2)}{h^2 r_c^2}\,\frac{E_t}{E_l} \tag{4a}$$

and

$$D_s = \frac{E_l h^3}{12(1-\nu_s^2)} \tag{4b}$$

Boundary conditions at the edges of the ribs are $y = dy/dz = 0$. Here z is the axial coordinate, ν_s is Poisson's ratio (assumed to be 0.3, which is Poisson's ratio for Mylar A), h is the thickness of the shell, r_c is its radius, and E_l and E_t are the Young's moduli in the axial and transverse direction (4.72 and 5.20 G Pa, respectively, for our system). We extract the influence matrix for a ring from the discretized solution of equations (3)–(4). Combining the influence matrices for each ring yields the influence matrix for the entire flexible shell; adding one line and one row in the influence matrix for the bubble completes the assembly of the $[K]^{-1}$ matrix for the flexible test tank.

The eigenvalues of the matrix $[[K]^{-1}[H]]$represent the system natural frequencies. In either the flexible or the rigid system, if the eigenvalues are λ_i, the system natural frequencies are

$$f_i = \frac{1}{2\pi\sqrt{\rho\lambda_i}} \tag{5}$$

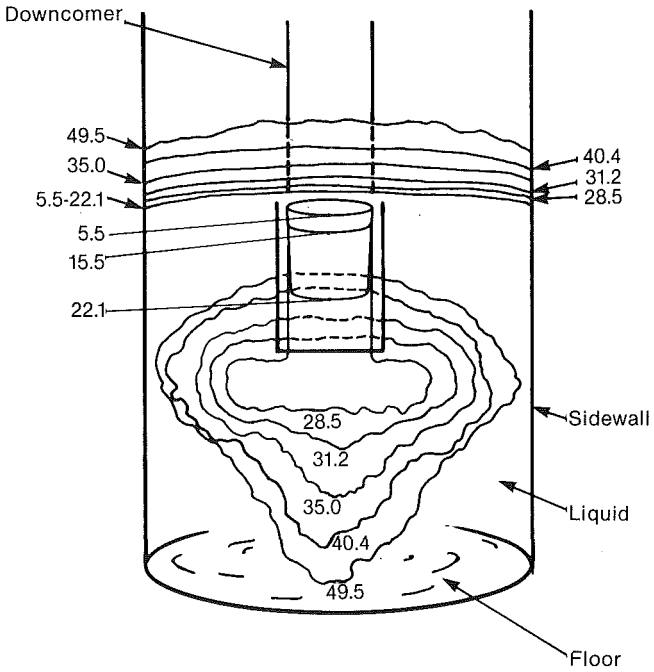


Fig. 5 Profiles obtained from high speed film showing liquid redistribution during the poolswell transient. The test system's side wall curvature radially magnifies the interior surfaces by twenty percent.

where $\rho = 1000$ kg/m^3. The rigid system has only a single natural frequency.

It is plain from Fig. 4 that for both the rigid structure and for the flexible structure this analysis yields a very accurate prediction of the experimental results. It should be emphasized that the only inputs to the theory are a description of the system geometry and the liquid, gas, and structure properties (ρ, γ, E_1, E_t, ν_s).

III Forced Transients

We next consider application of the theory to a transient flow that involves large liquid deformations. We demonstrate how rigid system data may be used to predict the flexible system response, and how flexible system data may be used to predict the rigid system transient. This pair of calculations, applied to experimental tests, illustrates the strengths as well as some of the practical difficulties that attend implementation of the analysis.

A Experiments. The same two test tanks were used. The tanks were filled with water to a depth of 152 mm. A drywell (1.57×10^{-3}m^3 capacity) was connected through a solenoid valve to a downcomer, which extended below the liquid surface as shown in Fig. 1(*b*). The initial pressure in the wetwell and downcomer was atmospheric. The drywell was pressurized with nitrogen gas to a gauge pressure of 25.0 kPa. A transient was produced by opening the solenoid valve. Nitrogen flowed rapidly through the downcomer into the liquid, forming a large bubble, redistributing the liquid mass, and causing the pool to "swell" abruptly. During poolswell the pressure above the pool remained close to atmospheric. The 400 ms transient terminated when the pressure in the drywell had dropped to atmospheric and the pool had subsided following bubble breakthrough. Our experiments and analysis focus on the first 50 ms of the transient.

We obtained high speed films of the rigid system transient to determine the liquid geometry as a function of time (Fig. 5). We also used piezoelectric transducers to record pressures at various locations in both systems (Fig. 1): at the center and

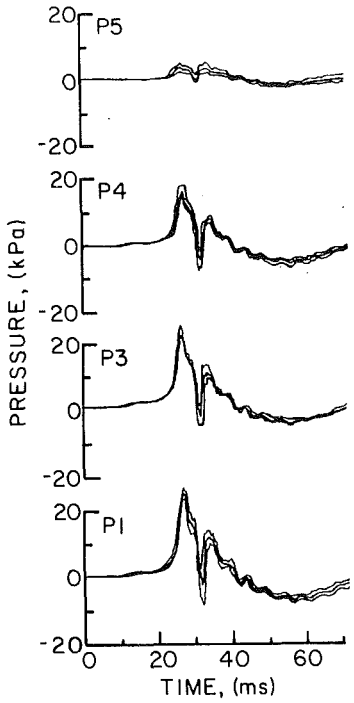


Fig. 6 Experimentally measured rigid system response

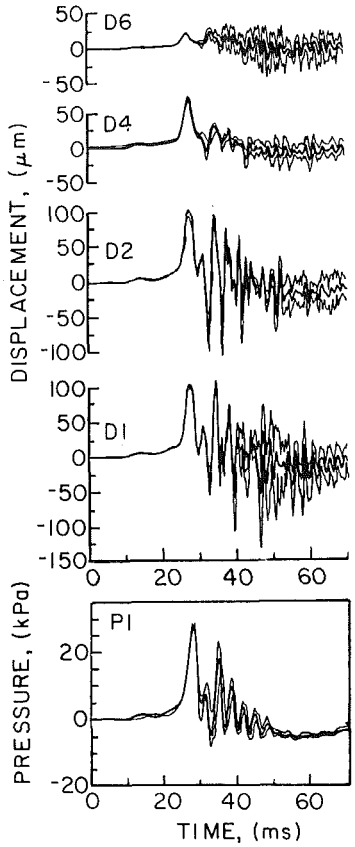


Fig. 7 Experimentally measured flexible system response

half-a-pool radius off-center on the pool floor, at three heights along the sidewall (in the rigid tank only), and in the unwetted region of the tank and downcomer in both systems. Using a photonic sensor we recorded transient boundary oscillations in the flexible tank at eight locations on the cylindrical walls (six locations are shown in Fig. 1(*b*)). Additional details of the experimental method and results are presented in Joos (1982).

Typical pressure and displacement data at selected sites in the rigid and flexible tanks are shown in Figs. 6 and 7. Both average results and the data-bounding envelopes of ten or more identical tests are shown. The first pool pressure peak (trace *P*1, Fig. 6, at $t=26.2$ ms) occurs just after the downcomer clears the gas-liquid interface is about 15 mm below the submerged tip of the downcomer. Thereafter the bubble grows and pressures in the pool first fall and then rise again. Average wall displacements in the flexible tank roughly follow the pressure histories of the rigid tank, but oscillations of significant amplitude are superimposed (Fig. 7). These are best seen in the oscillatory pressures on the pool floor in the flexible system tests (trace *P*1, Fig. 7).

B Application of the System Equations. The data from the pair of transient tests just described present two paradigm problems for application of our analytical framework. In Table 2 we demonstrate that, in general, the conditions of our experimental tests satisfied the criteria for applying our analysis. Once the pressure and liquid distribution histories in a rigid system are known, direct application of the system equations, equations (43) or (44) of Part 1, should therefore yield perturbation displacements and pressures in a geometrically identical system of known boundary flexibility. Alternatively, the system equations may be used to extract the rigid system characteristics from data obtained in a flexible system.

Examination of the high speed films from which the profiles of Fig. 5 were taken reveals there is a region in the liquid domain that does not obey the simplifying assumptions (*B*) and (*E*) of Table 2 of Part 1 during the early stages of the transient. These assumptions require that the mixed cross-convective terms be small compared to the perturbation's acceleration in the flexible system. During the downcomer clearing stage, ending at $t=26.2$ ms, these terms are at best comparable in the liquid being injected from downcomer. Similar observations can be made of the next stage, consisting approximately of the following 2.5 ms, when the downcomer gas-liquid interface rapidly expands mostly in the horizontal direction. At the end of this stage the experimental pressures and displacements (Figs. 6 and 7) are halfway down the negative slope between the first two peaks. The total amounts of liquid displaced by the end of these stages are very small: about 1.5 percent when the downcomer clears and 2 percent by the end of the second stage, at about $t=28.5$ ms. Thus, with the exception of the small region near the downcomer, most of the remaining liquid is virtually at rest and satisfies all the simplifying assumptions. Comparison of the bottom-center pressure histories of the rigid and flexible systems, trace *P*1 of Figs. 6 and 7, suggest that in these early stages the perturbation fields are still small compared to the rigid system fields, and the most evident perturbative effects are yet to come. Therefore, bearing in mind the limitations imposed on the comparison by experimental inadequacies in a region of the liquid and stage of the transient of small interest, we proceed with the comparison as if the perturbation equations were fully valid.

1 *Determining the Transient Liquid and Influence Matrices.* Records of the transient liquid redistribution were traced from high speed films of tests in the rigid system. A total of 16 profiles of the transient (some of the them shown in Fig. 5) were used to define the transient; intermediate distributions were determined by interpolation. After correcting for optical magnification, we assumed that the profiles accurately represented the outer boundary of the bubble. We then assumed a crude inner boundary using conservation of the liquid volume as a criterion. The precise shape of those inner surfaces should have little effect on the analysis.

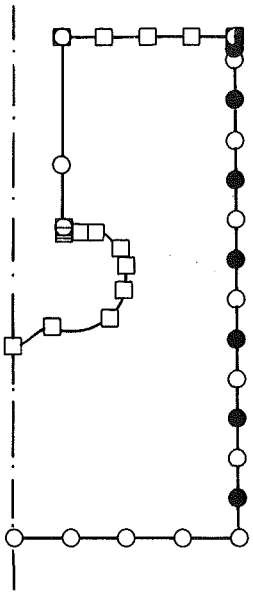


Fig. 8 Nodal distribution on the liquid interfaces at time $t = 31.25$ ms. Squares represent pressure boundary conditions; circles represent normal pressure gradient boundary conditions. Only solid nodal points have inhomogeneous boundary conditions. The two boundary conditions for the top right corner are an inhomogeneous pressure gradient on one side and a homogeneous pressure on the other.

As in Section II, we approximated the liquid and influence operators by their discretized counterparts. We used a boundary integral method to solve Laplace's equation, and so to determine $[H]$ as a function of time. Figure 8 shows a typical distribution of the computational nodes along the boundary. Inhomogeneous boundary conditions occur only at the midpoint of each flexible ring, and on the topmost, partially wetted ring. All other nodes represent homogeneous boundary conditions—either zero perturbation pressure (at free surfaces) or zero nominal perturbation pressure gradient (at rigid boundaries, including the nodes located on the sidewall ribs). Conservation of mass requires that the vibration amplitudes at the gas-liquid interface be comparable to those at the flexible walls. As Table 2 indicates, the stiffness constant k_g of the gas cavity is negligible compared to that of the structure, k_s; therefore, the gas-liquid interface created by air injected through the downcomer can reasonably be modelled as a free surface.

To calculate the influence matrix $[K]^{-1}$ for the flexible system, we defined shell rings consisting of shell segments between each pair of ribs. The operator governing the shell's displacement is again given by equations (3) and (4). The rings are assumed fixed at the ribs and each ring, therefore, vibrates independently.

2 *Calculating Flexible System Behavior from Rigid System Data.* Most real systems containing gas cavities are "flexible," because boundary displacements at such interfaces usually contain both low and high frequency components, the latter kind modifying the cavity volume and thus producing perturbation pressures. Although liquid-structure interfaces can be made rigid through appropriate stiffening, it is usually not possible to force gas-liquid interfaces to behave "rigidly," i.e., to displace only at low frequency. For most systems, therefore, no "rigid" system *experiment* is possible—all experimental tests necessarily involve "flexible" systems, and the idealized system response must either be calculated theoretically or extracted from flexible system data as described in the following subsection.

An exception arises, however, when all gas-liquid interfaces can be treated as free surfaces, that is, when the pressures at all such surfaces are independent of any perturbation displacements of the surfaces. In such cases the rigid system pressure history can effectively be measured directly by experiment, in a test system with rigid structural boundaries.

Our poolswell tests in the structurally rigid system represent an experiment of the latter type: the gas cavity generated at the downcomer exit was large enough to be treated as a free surface. For the conditions of our transient tests, all inhomogeneous boundaries therefore have surface displacement constitutive relations. Adding equations (35) and (43), both of Part 1, and considering that low frequency wall accelerations are negligible compared to the perturbation accelerations, we obtain

$$\rho H\ddot{y} + K_s y = p_R \tag{6}$$

Equation (6) can be solved numerically, with the operators H and K_s replaced by the corresponding matrices $[H]$ and $[K]$ and the displacements y by the array $\mathbf{y}$. Equation (6) is first rewritten:

$$\rho[K]^{-1}[H]\ddot{\mathbf{y}} + \mathbf{y} = [K]^{-1}\mathbf{p_R} \tag{7}$$

We determine total pressures by combining equations (3) and (42), both of Part 1.

$$p = p_R - \rho H\ddot{y}_P \tag{8}$$

which, in matrix form is

$$\mathbf{p} = \mathbf{p_R} - \rho[H]\ddot{\mathbf{y}} \tag{9}$$

where, again, we have applied the assumptions that rigid system accelerations at the structural boundaries are negligible compared to the perturbation accelerations.

Because the liquid distribution varies slowly compared with the lowest characteristic frequencies of $\mathbf{p}$ and $\mathbf{y}$, the H and K operators also vary slowly. The matrices $[H]$ and $[K]^{-1}$ were updated fifteen times during the portion of the transient we analyzed, $0 \le t \le 50$ ms. The interval between consecutive updatings ranged from 1 to 10 ms. During each interval an eigenvalue solution was used to solve equation (7). The decoupled differential equations which resulted from the decomposition of the system into its natural modes were solved by Duhamel integrations. Total pressures were then computed using equation (9).

Figure 9 shows typical predictions of the flexible system's pressure and displacement responses. For comparison with the experimental results we show displacements at the midpoints of several of the flexible wall rings and pressures at the center of the pool floor. Corresponding experimental results are shown in Fig. 7.

Clearly the theoretical predictions are in good general agreement with the experiments. The predicted first peak in floor pressure (Fig. 9, $t = 27$ ms, trace $P1$) is almost 12 percent higher than the average experimental trace; subsequent predicted peaks and troughs fall within the envelopes of the experimental results. The calculations somewhat underpredict the decay of the flexible system oscillations. There is some phase shift between the predicted oscillations and the experiments. The predicted frequencies of the oscillating component of the flexible system response agree well with the experimental observations.

Agreement between predicted and experimentally measured displacements is also good, though less so than for the pressures. For the lowest two rings (traces $D1$, $D2$, Fig. 9) where the displacements are greatest, the amplitude of the first peak is overpredicted by about 23 percent; many of the troughs are under-predicted and certain high frequency displacement components observed experimentally are not predicted at all. Nevertheless, the basic features of the predicted displacement responses agree well with the experiments. Both prediction and experiment reveal that the

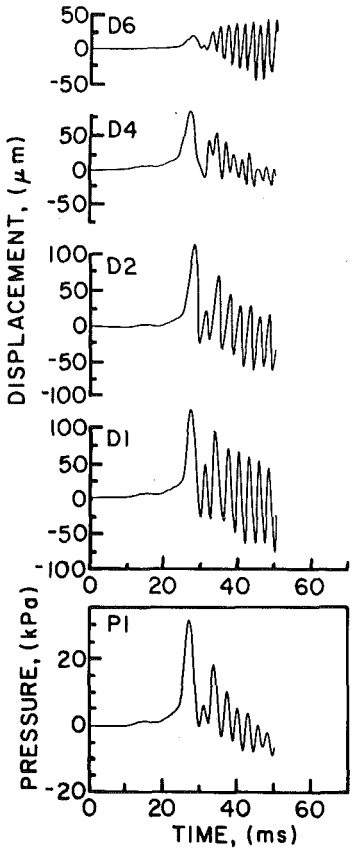


Fig. 9 Flexible system response predicted from rigid system data

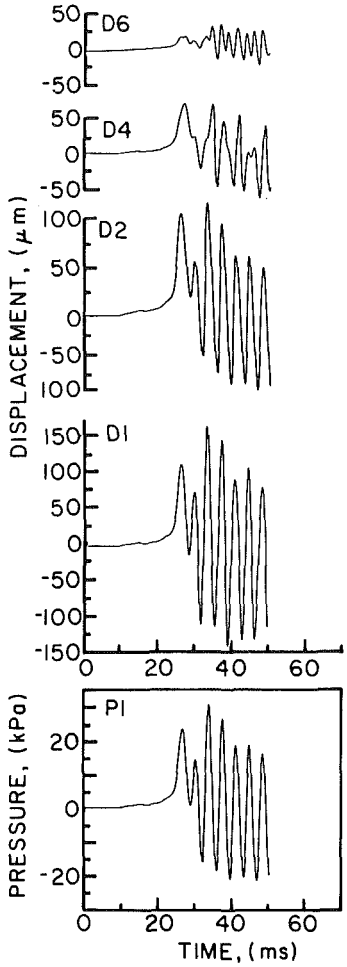


Fig. 10 Flexible system response predicted from rigid system data using a frozen liquid distribution corresponding to the actual distribution at time t = 31.25 ms

oscillation frequencies change with depth in the pool, increasing with decreasing distance from the free surfaces. Both predicted and observed displacement amplitudes vary significantly from ring to ring along the side wall and the variation is at first surprising—displacement amplitudes at the sixth ring are larger than those at the fourth ring, even though the former is much closer to the pool's surface. The rate of decrease of displacement amplitude in both prediction and experiment is greater at the second ring (traces $D2$, Figs. 7 and 9) than at the first (traces $D1$, Figs. 7 and 9). The analysis suggests that this is due to the rapid rate at which the free surface attached to the downcomer approaches the wall at that point. The liquid-gas interface is a surface of zero perturbation pressure and acts as a sink, reducing the amplitude of the perturbation pressures near it.

The greatest discrepancy between displacement predictions and the experiment occurred at ring 3, not shown here, where the predicted amplitude of the first peak was 100 μm, while experiments recorded 44.6 μm. The error may derive from the coarseness of the sidewall pressure history assumed for the calculation, which consisted of linearly interpolating the rigid system pressures between the sites of the transducer ports.

It is informative, finally, to compare the predictions we have presented with those of an identical computation that fails to account for the dynamic redistribution of the liquid mass during the transient. In Fig. 10 we show the results of such a computation. Equations (7) and (9) have been solved here as before, but the liquid distribution throughout the transient has been frozen in the distribution that occurs at time $t = 31.2$ ms (see Fig. 8). (This particular time is somewhat arbitrarily chosen "average" configuration.) The frequency of the predicted oscillations is, of course, constant and the oscillation amplitudes, low in the pool and late in the transient $t > 32$ ms, are grossly overestimated.

3 *Calculating Rigid System Pressures From Flexible System Data.* A more difficult application of the system equation involves the converse problem—extracting idealized rigid system characteristics from pressure or displacement data obtained in a flexible system. Total pressure and total displacement at each boundary are related by the constitutive equations, equations (30), (34), and (37) of Part 1; once one is known the other can thus be found directly.

The perturbation displacements are the high frequency components of the known total displacement. A zero phase shift, high-pass filter with a cutoff frequency between $1/\tau_R$ and $1/\tau_P$ can, therefore, be used to extract the displacement perturbation y_P or the component of the total pressure which produces y_P. The rigid pressures are then obtained by rearranging equation (8).

We can now predict pressure and displacement histories in a flexible system from information obtained in a geometrically identical system of different, known flexibility. The rigid system information is extracted from the flexible system data, and is then used in one of the system equations, (43) or (44) of Part 1, to predict flexible system pressures and displacements for different flexibilities.

For the conditions of our experiments, equations (34) of Part 1 and (8) can be rewritten in matrix terms:

$$\mathbf{p} = [K]\mathbf{y} \tag{10}$$

and

$$\mathbf{p}_R = \mathbf{p} + \rho[H]\ddot{y} \tag{11}$$

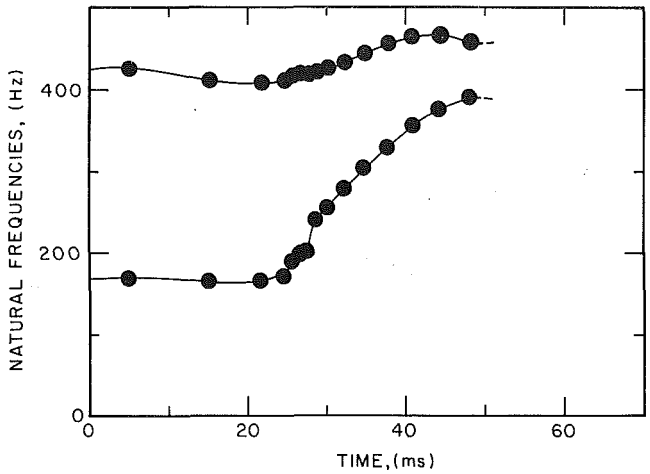


Fig. 11 Predicted variation of the lowest two system natural frequencies during the transient

where the structural mass and forcing pressure in equation (34) of Part 1 have been neglected, and we leave low frequency component displacement terms in equation (11) because their contribution to the acceleration term is negligible. As inputs we used the average flexible system experimental displacement results (Fig. 7) and the liquid distribution history.

Equation (11) requires second derivatives in time to be calculated from the experimental signal. Performing these calculations mechanically yields predictions that deviate significantly from the rigid system experimental data (Fig. 6). Although the basic features of the rigid system transient emerge from the computations, the predictions contain significant high frequency components that are not observed experimentally. But their presence is not altogether surprising; the input flexible system data were essentially unfiltered (only electrical noise had been removed with a 1 kHz digital low-pass filter) and differentiating such experimental signals twice in time, as required by equation (11), greatly amplifies any high frequency content. This situation could have been avoided by also measuring normal surface accelerations; we decided, however, to solve this problem as indicated below.

Suitable filtering of the rigid system predictions might successfully remove these parasitic high frequency signals while preserving the rigid system features that the calculation is intended to extract. But because the system natural frequencies vary with time it is not obvious which pass and stop-band filtering frequencies to choose.

Figure 11 shows our predictions of the two lowest natural frequencies of the flexible system (found from the lowest eigenvalues of equation (7) as a function of time during the transient). It is, from inspection of the flexible system experiments, not clear *a priori* when in the transient the flexible oscillations are excited. In trace $P1$, Fig. 7, for example, it appears that flexible system oscillations have been initiated at about $t = 33$ ms in the transient, although one cannot assert with any confidence that further significant excitation of the system does not continue beyond that time. Inspection of Fig. 11 shows that at $t = 33$ ms the lowest system natural frequency is about 290 Hz. On this somewhat tenuous basis we filter our predictions using a low-pass, zero phase shift digital filter with cutoff and rejection frequencies 290 and 330 Hz. These two frequencies correspond to the lowest system natural frequencies at $t = 33$ and 36 ms, respectively. Figure 12 shows the results of this modification in our rigid system predictions. The choice of pass and stop-band frequencies could be systematized and improved as follows. Following a choice of these frequencies the newly calculated (filtered) rigid pressure histories are used to recalculate the flexible system pressures

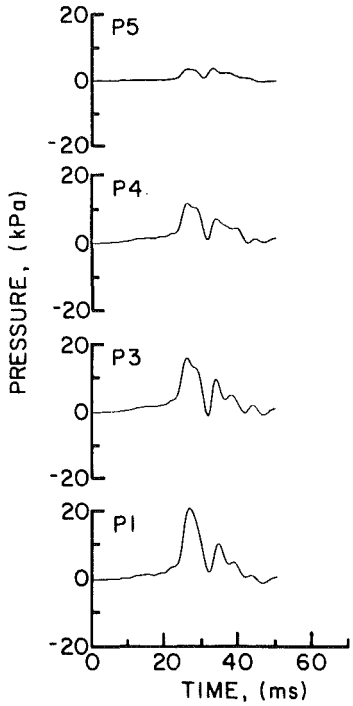


Fig. 12 Filtered rigid system response predicted from flexible system data

and displacements. These predictions are compared with the original data. If the amplitudes of the predictions in the fundamental modes of the the flexible system do not compare well with the experimental data, the predicted rigid system pressures must lack (or contain an excess of) some frequency component that excites the flexible system. It should thus be possible, by iteration, to improve the choice of pass and stop-band frequencies until the predicted rigid system histories prove acceptable at repredicting the flexible system responses from which they are derived.

A comparison of Figs. 6 and 12 reveals that the filtered predictions match the rigid system experiments quite successfully. If the traces in Fig. 6 are passed through the same digital filter, the amplitudes of the first peak in Fig. 8 are underpredicted by less than 15 percent of the maximum peak, an error that is consistent with our *over*-prediction of the *flexible* system responses in our calculations using the rigid data as input (Fig. 9). It is possible that one of the calculated matrix coefficients in equation (7), either $[K]^{-1}$ or $[H]$, is somewhat in error. That has led to an overprediction of the perturbation field components. When we add the predicted perturbations to rigid system data we overpredict the flexible transient somewhat; when we subtract the predicted perturbations from flexible system data we under-predict the rigid transient by about the same amount. Another possible source of error is the failure of the perturbation equations in the neighborhood of the downcomer when this peak occurs.

The pressure traces in Fig. 12 have been obtained by passing the calculated rigid system pressures through the low-pass digital filter. Predictions essentially identical to those in Fig. 12 can be obtained by filtering the experimental flexible pressure and displacement data first, then performing the equation (11) calculations. Because the liquid matrix $[H]$ varies slowly in time, it has no high frequency components; thus filtering may be performed before or after the equation (11) computations.

It is important to emphasize, however, that filtering alone will not successfully translate flexible data into rigid system predictions. Figure 13 shows the flexible system data filtered

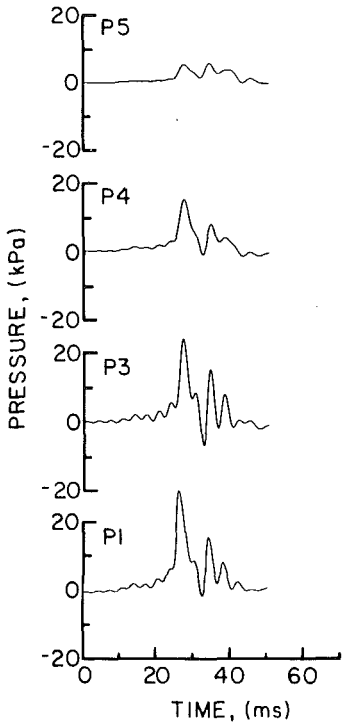


Fig. 13 Flexible system data filtered through the low-pass filter used in Fig. 12

at 290-330 Hz. A comparison with Figs. 6 and 12 demonstrates that the liquid-structure coupling is not eliminated simply by filtering the total pressure traces in the flexible system. The acceleration terms in equation (11) are important and must be calculated and subtracted in going from flexible data to rigid system predictions.

4 *Verification of the Numerical Methods.* To check the internal consistency of the two computational schemes used in the preceding subsections, we performed a circular calculation as follows. Starting with the rigid system experimental data and the matrix coefficients from equation (7), we calculated flexible system responses according to the Subsection III.B.2 recipe. The predicted flexible system displacements were then used as the input "data" in the scheme described in Subsection III.B.3 to recompute the rigid system pressures. Apart from some sharp numerical peaks introduced by the abrupt changes in liquid distribution, the original rigid pressures were reproduced very accurately.

IV Conclusion

The prediction of large-scale coupled gas-liquid-structure flow fields by computation or experiment is difficult. Here we have demonstrated that, when simplifying criteria formulated in the companion paper are satisfied, the problem can be decomposed into two simpler ones. The first focuses on the gross deformations of the liquid; the second addresses oscillatory motion.

Based on the analysis, a rigid system flow field can be computed, measured directly, or extracted from data obtained in a flexible system. The rigid system characteristics can then be used as inputs to compute flow fields in a system of different flexibility. In addition, natural frequencies of a coupled gas-liquid-structural system can be determined through a relatively simple calculation.

Acknowledgments

This work was sponsored by the United States Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, under Grant No. NRC-G-04-81-028. Part of the clerical and art work was performed by personnel of the employers of one of the authors (F. M. Joos).

References

Joos, F. M., 1982, Ph.D. Thesis, Dept. of Mech Engg., Massachusetts Institute of Technology, Cambridge, MA.

Joos, F. M., and Huber, P. W., 1987, "Coupled Gas-Liquid-Structure Systems: Part 1—Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 935-941.

Kana, D. D., Glaser, R. F., Eulitz, W. K., and Abramson, H. N., 1968, "Longitudinal Vibration of Spring Supported Membrane Shells Containing Liquid," *J. Spacecraft*, Vol. 5, pp. 189-196.

Timoshenko, S., and Woinowsky-Krieger, S., 1959, Theory of Plates and Shells, 2d ed., McGraw-Hill, New York, p. 468.

Mechanics of a Free-Surface Liquid Film Flow

Cyrus K. Aidun[1]
School of Chemical Engineering and
Mathematical Sciences Institute,
Cornell University,
Ithaca, NY 14853-5201
Assoc. Mem. ASME

The mechanics of a free surface viscous liquid curtain flowing steadily between two vertical guide wires under the influence of gravity is investigated. The Navier-Stokes equation is integrated over the film thickness and an integro-differential equation is derived for the average film velocity. An approximate nonlinear differential equation, attributed to G. I. Taylor, is obtained by neglecting the higher order terms. An analytical solution is obtained for a similar equation which neglects the surface tension effects and the results are compared with the experimental measurements of Brown (1961).

1 Introduction

The process of film flow is used in single and multi-layer coating of photographic bases, spinning of liquid films, and protective organic coatings at high speed to continuous steel or tinplate strip. One method of film coating used in industry is to pump fluid through a narrow slot; as the liquid passes through this slot, it forms a continuous sheet or curtain which falls on the workpiece as it travels underneath. The thickness of the coating material applied depends on the mass flux of the fluid, the velocity and distance of the workpiece from the slot, and the velocity of the fluid.

The dynamics of liquid sheets have been examined by Taylor (1959a, b, c), who studied the shapes of axially symmetric "water bells," capillary waves, and the disintegration of thin sheets of fluid. The stability of a liquid film with large viscosity was studied by Yeow (1974), who considered only the viscous and pressure terms in the momentum equation. By means of linear stability analysis, he formulated a set of eigenvalue problems and obtained the neutral-stability curves.

The stability of a thin liquid curtain with respect to spatially as well as temporally changing disturbances is studied by Lin (1981). He showed that the liquid curtain is unstable to spatially growing sinuous disturbances if the Weber number exceeds 1/2.

The velocity distribution in a thin Newtonian liquid curtain falling between two guide wires (see Fig. 1) is measured by Brown (1961), and an equation based on the free-fall velocity of a particle is deduced by him.

The following nonlinear differential equation, for the axial film velocity, is also cited in the appendix of Brown (1961), and is attributed to G. I. Taylor.

$$U\frac{dU}{dx} = 4\nu U\frac{d}{dx}\left(\frac{1}{U}\frac{dU}{dx}\right) + g \tag{1}$$

here, U is the velocity in the direction of gravity g, and ν is the kinematic viscosity. In citing this equation, Brown makes the assumption that the lateral component of the stress is constant, and the surface tension effects are negligible. Although the result predicted by equation (1) agrees well with the experimental data of Brown, the origin and limitations of this equation, however, are not well understood. A formal derivation would clarify some ambiguities associated with this equation. Also, in the standpoint of engineering applications, a closed-form solution for the liquid curtain velocity is highly desirable.

The objectives of the present analysis are to apply the Navier-Stokes equation to a falling liquid curtain and present a formal derivation of equation (1), and to obtain an analytical solution of a similar equation valid for thin liquid curtains. In addition to engineering applications, an analytical solution will serve as a first order approximation of the veloci-

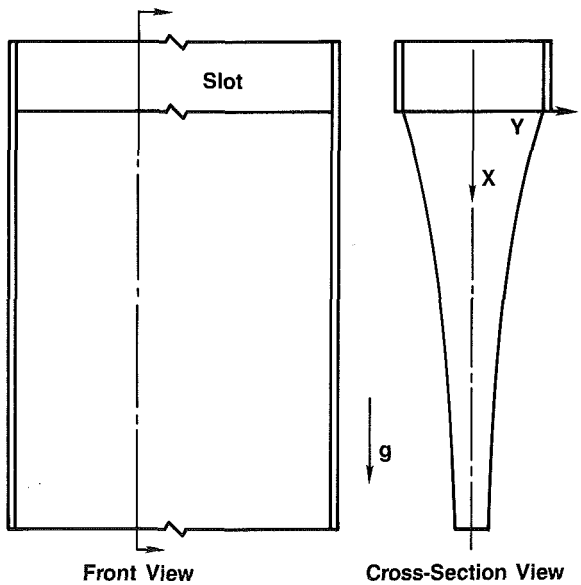


Fig. 1 Schematic diagram of a free-surface liquid film flow

[1]Present address: Battelle Memorial Institute, 505 King Avenue, Columbus, Ohio 43201-2693.

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 25, 1986; final revision April 23, 1987.

ty profile in a spinning or a falling curtain. The domain of validity for each equation is established by comparing the analytical results with the numerical solution of equation (1), the experimental results of Brown (1961), and the results from the free-fall equation.

2 Formulation

Normally, in the spinning or coating processes, the film thickness is much smaller than the width and end effects are negligible; therefore, we assume two-dimensional incompressible flow governed by the following equations of motion:

$$\text{continuity} \quad \frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \tag{2}$$

$$x\text{-momentum} \quad u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=\frac{1}{\rho}\left(\frac{\partial}{\partial x}\pi_{xx}+\frac{\partial}{\partial y}\pi_{xy}\right)+g \tag{3}$$

$$y\text{-momentum} \quad u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}=\frac{1}{\rho}\left(\frac{\partial}{\partial x}\pi_{xy}+\frac{\partial}{\partial y}\pi_{yy}\right) \tag{4}$$

Here, u and v are the x and y components of velocity and π_{xx}, π_{xy}, etc. are components of the total stress tensor π, given in the standard notation. The density, ρ, is assumed to be constant throughout the process.

We define the free surface of the film by the function $F(x,y)$ where

$$F(x,y)\equiv y-\eta(x)=0 \tag{5}$$

and $\eta(x)$ is the thickness of the film at x. The boundary of the film is a streamline; therefore, the material derivative of F, i.e., DF/Dt must vanish at $F=0$. This results in the following boundary condition

$$v=u\eta' \text{ at } y=\eta \tag{6}$$

where $'$ indicates differentiation with respect to the single argument. The surface tension σ, creates a stress on the free surface boundary. Balancing the surface forces on the boundary yields

$$\pi_{xy}n_y+\pi_{xx}n_x=\frac{\sigma}{Z}n_x \tag{7a}$$

$$\pi_{yx}n_x+\pi_{yy}n_y=\frac{\sigma}{Z}n_y \tag{7b}$$

where the components of the unit normal vector $\mathbf{n}$ are given by

$$n_x=-\eta'(\eta'^2+1)^{-1/2},\quad n_y=(\eta'^2+1)^{-1/2} \tag{8a,b}$$

and the radius of curvature Z is given by

$$Z=(\eta'^2+1)^{3/2}/\eta'' \tag{9}$$

We decompose velocity u and normal stress π_{xx} as

$$u(x,y)=\bar{u}(x)+\epsilon u_1(x,y) \tag{10}$$

$$\pi_{xx}(x,y)=\overline{\pi_{xx}}(x)+\epsilon\pi_{xx1}(x,y) \tag{11}$$

here ϵ is in someway related to η'. The functions $\bar{u}(x)$ and $\overline{\pi_{xx}}(x)$ are unknowns at this point and will be derived later in the analysis. Further below we show that if $\eta'<<1$, then $\epsilon<<1$, and the functions u and π_{xx} are only weakly dependent on y.

With this decomposition, the integral of the continuity equation over the film thickness is given by

$$\int_0^\eta\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)dy=\eta\bar{u}'-\epsilon u_1(x,\eta)\eta'+v(x,\eta)=0 \tag{12}$$

where we have employed Leibnitz's rule to evaluate this integral. From boundary condition (6), we have

$$v(x,\eta)=\bar{u}\eta'+\epsilon u_1(x,\eta)\eta' \tag{13}$$

substituting this relation into equation (12), and integrating once with respect to x, results in the global form of the mass conservation,

$$Q=\bar{u}\eta \tag{14}$$

where Q is a constant representing the volumetric mass flow and $\bar{u}$ is, therefore, the average velocity defined by

$$\bar{u}\equiv\frac{1}{\eta}\int_0^\eta u\,dy \tag{15}$$

We integrate the momentum equation over the film thickness in the same manner. The integrals of the nonlinear inertia terms in the x component of the momentum equation are

$$\int_0^\eta u\frac{\partial u}{\partial x}dy=\eta\bar{u}\bar{u}'-\epsilon\eta' u_1(x,\eta)\bar{u}+\frac{\epsilon^2}{2}A(x) \tag{16}$$

$$\int_0^\eta v\frac{\partial u}{\partial y}dy=\epsilon v(x,\eta)u_1(x,\eta)+\frac{\epsilon^2}{2}A(x) \tag{17}$$

where

$$A(x)\equiv\frac{d}{dx}\int_0^\eta u_1^2dy-u_1^2(x,\eta)\eta'$$

Using the decomposition for π_{xx} given by equation (11), one can show that

$$\int_0^\eta\frac{\partial}{\partial x}\pi_{xx}dy=\eta\overline{\pi'_{xx}}+\epsilon\left[\frac{d}{dx}\int_0^\eta\pi_{xx1}dy-\pi_{xx1}(x,\eta)\eta'\right] \tag{18}$$

and

$$\int_0^\eta\frac{\partial}{\partial y}\pi_{xy}dy=\pi_{xy}(x,\eta) \tag{19}$$

here, $\pi_{xy}(x,0)$ vanishes due to symmetry of the flow with respect to centerline. From equation (7), $\pi_{xy}(x,\eta)$ is given by

$$\pi_{xy}(x,\eta)=-\left[\frac{\sigma}{Z}-\pi_{xx}(x,\eta)\right]\eta' \tag{20}$$

The integral of the x component of the momentum equation over the film thickness, after some manipulations, can now be written as

$$\bar{u}\bar{u}'-\frac{\epsilon}{\eta}\left[\eta'\bar{u}u_1(x,\eta)-v(x,\eta)u_1(x,\eta)-\epsilon A(x)\right]$$

$$=\frac{1}{\rho}\overline{\pi'_{xx}}+\frac{\epsilon}{\rho\eta}\left[\frac{d}{dx}\int_0^\eta\pi_{xx1}dy\right]-\frac{1}{\rho\eta}\left(\frac{\sigma}{Z}-\bar{\pi}_{xx}\right)\eta'+g \tag{21}$$

If the magnitude of x, $\bar{u}$, and u_1 are assumed to be 0(1), then for a thin film, y and η' are $0(\eta)$, where $\eta<<1$. Also, one can show from equation (6) that $v=0(\eta)$. Employing the Newtonian fluid model, i.e.,

$$\pi_{ij}=-p\delta_{ij}+\mu\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right) \tag{22}$$

where $(x_1,x_2)\equiv(x,y)$ and $(u_1,u_2)\equiv(u,v)$, we can express the y component of equation (7) by

$$\mu\epsilon\left(\frac{\partial u_1}{\partial y}\right)_{y=\eta}=-\frac{dv(x,\eta)}{dx}-\frac{\sigma}{Z}\eta'+2\mu\left(\frac{\partial v}{\partial y}\right)_{y=\eta}\eta' \tag{23}$$

Since the right-hand side of equation (23) is of order (η), we conclude that $\epsilon=0(\eta^2)$. It can be shown from equation (3) that

$$\frac{\partial}{\partial y}\pi_{yy}=0(\eta) \tag{24}$$

therefore, to the first order, we can assume that

$$\pi_{yy}=-P_0 \tag{25}$$

where P_0 is a constant representing the ambient pressure. The expression for pressure is then given by

$$p(x,y)=P_0-2\mu\bar{u}'-2\epsilon\mu\left[\frac{\partial u_1}{\partial x}+\ \ldots\ \right] \tag{26}$$

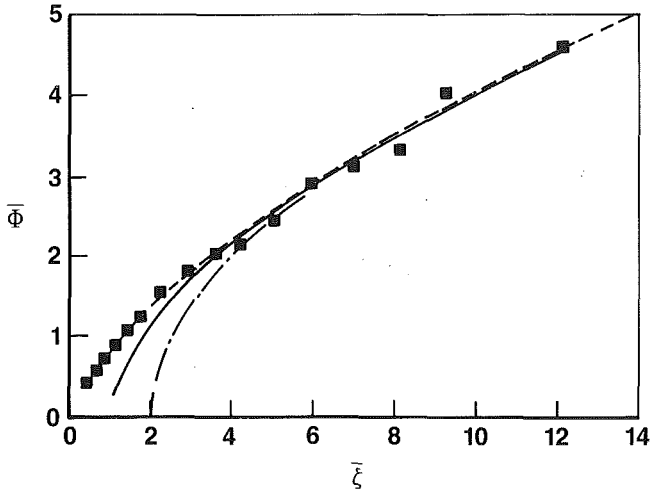


Fig. 2 Average curtain velocity versus distance from the slot; ■ experimental result (Brown 1961); _ _ _ _ , numerical solution of equation (1); ____ , analytical solution of equation (29); __._ , free-fall equation (Brown, 1961)

The x component of the normal stress tensor can now be written as

$$\pi_{xx} = -P_0 + 4\mu\bar{u}' + 4\epsilon\mu\left[\frac{\partial u_1}{\partial x} + \ \ldots\ \right] \tag{27}$$

Note that the functional form of π_{xx} is consistent with the decomposition given by equation (11), and therefore,

$$\bar{\pi}_{xx} \equiv -P_0 + 4\mu\bar{u}'$$

Substituting equation (27) for the x component of the normal stress tensor in equation (21), and neglecting the higher order terms results in

$$\bar{u}\bar{u}' = \frac{4\mu}{\rho}\bar{u}'' - \frac{\eta'}{\rho\eta}\left(\frac{\sigma}{Z} + P_0 - 4\mu\bar{u}'\right) + g \tag{28}$$

If we neglect the surface tension σ and normalize pressure such that $P_0 = 0$, then this equation reduces to equation (1).

3 Thin Films

The effect of the surface curvature can be neglected for thin films and equation (28) reduces to

$$\bar{u}\bar{u}' = 4\nu\bar{u}'' + g \tag{29}$$

We define the dimensionless variables

$$\bar{\Phi} \equiv \left(\frac{1}{4\nu g}\right)^{1/3}\bar{u} \tag{30}$$

$$\bar{\zeta} \equiv \left(\frac{1}{4\nu}\right)^{2/3} g^{1/3}x \tag{31}$$

and substitute (30) and (31) into equation (29) to get

$$\bar{\Phi}'' = \bar{\Phi}\bar{\Phi}' - 1 \tag{32}$$

Integrating this equation results in a first order differential equation given by

$$\bar{\Phi}' = \frac{1}{2}\bar{\Phi}^2 - \bar{\zeta} - C \tag{33}$$

where C is the integration constant. Equation (33) can be written in a more convenient form as

$$\Phi' - \Phi^2 + \zeta = 0 \tag{34}$$

where Φ and ζ are defined by

$$\Phi \equiv 2^{-2/3}\bar{\Phi}, \zeta \equiv 2^{-1/3}(\bar{\zeta} + C)$$

To solve equation (34) we use a nonlinear transformation of the dependent variable given by

$$\Phi(\zeta) = -\frac{\Psi'(\zeta)}{\Psi(\zeta)} \tag{35}$$

Substituting equation (35) into equation (34) yields the well-known Airy's equation

$$\Psi'' - \zeta\Psi = 0 \tag{36}$$

The general solution to this equation is given by

$$\Psi(\zeta) = \zeta^{1/2}\left[C_1 f_1\left(\frac{2}{3}\zeta^{3/2}\right) + C_2 f_2\left(\frac{2}{3}\zeta^{3/2}\right)\right] \tag{37}$$

where f_1 and f_2 are the linearly independent solutions of the Bessel equation of order one-third and C_1 and C_2 are the integration constants. The solution can also be expressed in terms of the linearly independent Airy functions Ψ_1 and Ψ_2 (see Abramowitz and Stegun, 1964).

$$\Psi = a_1\Psi_1(\zeta) + a_2\Psi_2(\zeta) \tag{38}$$

here a_1 and a_2 are the integration constants. The properties of these solutions are: (*i*) For $\zeta > 0$, one solution decays exponentially as the other independent solution grows exponentially; (*ii*) Infinity is an irregular-singular point; and (*iii*) For $\zeta < 0$ the solutions are oscillatory, similar to trigonometric sine and cosine functions. An oscillatory solution is not physically acceptable, since it contradicts the primary assumption of thin film with a smooth surface, therefore ζ and C are both positive.

Brown (1961) showed that for $\bar{\zeta} > 10$, the curtain velocity can be closely approximated by the equation for the free-fall velocity of a particle originating at $\bar{\zeta} = 2$. This equation will be referred to as the free-fall equation.

Equation (1) is solved numerically for the nondimensional average curtain velocity and the results are compared in Fig. 2 with the analytical solution of equation (29) and the experimental results of Brown (1961). The velocity profile obtained from the free-fall equation is also included in this figure for comparison purposes. The downstream velocity at $\bar{\zeta} = 12$ is extracted from the experimental results and the velocity gradient at this position is computed based on the free-fall equation, i.e.,

$$\bar{\Phi}(12) = 4.6 \tag{39a}$$

and

$$\frac{d\bar{\Phi}}{d\bar{\zeta}} = 1/\bar{\Phi} \ \text{ at } \ \bar{\zeta} = 12 \tag{39b}$$

The constant C is obtained directly from equation (33) and the boundary condition given by (39); the second arbitrary constant in the system is a_1/a_2 which is given by

$$\frac{a_1}{a_2} = -\frac{\Psi'_2 + \Phi_b\Psi_2}{\Psi'_1 + \Phi_b\Psi_1}, \ \bar{\zeta} = 12 \tag{40}$$

where

$$\Phi_b = 2^{-2/3}\bar{\Phi}(12)$$

The symbolic manipulation software, MACSYMA (1983), is used to evaluate the Airy functions and their derivative.

The results from equation (29) agree with the experimental results in the range where $\bar{\zeta} > 3$. For $\bar{\zeta} < 3$, however, the surface effects due to the surface curvature become important and equation (1) is more appropriate.

4 Conclusion

Information regarding the mechanics of a free-surface liquid film is important in industrial coating and spinning processes. Film spinning, also referred to as film casting, involves drawing a molten liquid from a slot into a thin sheet. The flow is extensional type since the draw speed downstream is larger than the extrusion velocity. The analytical solution for the average velocity, derived above, is applied here to a coating process which involves a liquid film falling freely under the force of gravity. The analytical solution can also be applied to

spinning processes by using the appropriate boundary conditions to compute the constants of integration. The downstream draw velocity is a controlled parameter in spinning processes and can serve as one of the boundary conditions. The choice of a second condition, however, is not trivial and depends on the particular process or experiment under investigation. Although the analytical solution was derived for thin films, it can also by used as a model for a first order approximation of the thickness and average velocity of a broader class of free-surface liquid films.

Acknowledgments

The author is grateful to Professor S. P. Lin for suggesting this problem to him.

References

Abramowitz, M., and Stegun, A., 1964, *Handbook of Mathematical Functions*, US Government Print, Washington, D.C.

Brown, D. R., 1961, "A Study of the Behavior of a Thin Sheet of Moving Liquid," *J. Fluid Mech.*, Vol. 10, p. 297.

Lin, S. P., 1981, "Stability of a Viscous Liquid Curtain," *J. Fluid Mech.*, Vol. 104, p. 111.

MACSYMA Reference Manual, MACSYMA Group Symbolic Inc., Cambridge, MA, December 1983.

Taylor, G. I., 1959a, "The Dynamics of Thin Sheets of Fluid, I. Water Bells," *Proc. Roy. Soc. A*, Vol. 253, p. 289.

Taylor, G. I., 1959b, "The Dynamics of Thin Sheets of Fluid, II. Waves on Fluid Sheets," *Proc. Roy. Soc. A*, Vol. 253, p. 296.

Taylor, G. I., 1959c, "The Dynamics of Thin Sheets of Fluid, III. Disintegration of Fluid Sheets," *Proc. Roy. Soc. A*, Vol. 253, p. 313.

Yeow, Y. L., 1973, "On the Stability of Extending Films: A Model for the Film Casting Process," *J. Fluid Mech.*, Vol. 66, Part 3, p. 613.

Z. M. Weng
Visiting Scholar and Associate Professor of Power Engineering,
Xian Jiaotung University,
China

Alice A. L. Ting
Resarch Assistant.

W. L. Chow
Professor of Mechanical Engineering.
Fellow ASME

Department of Mechanical and Industrial Engineering,
University of Illinois at Urbana-Champaign,
Urbana, IL 61801

Discharge of a Compressible Fluid Through a Control Valve

The original analysis on incompressible flow discharge from a vessel through an axisymmetric control valve has been extended to the discharge of a compressible fluid. The inviscid anlaysis is based on the method of hodograph transformation. While it is simple to account for the effect of compressibility for conditions of subcritical pressure ratios, special treatment must be applied to establish the sonic line and the free jet boundaries under conditions of supercritical pressure ratios. Discharge characteristics have been established for different pressure ratios and positions of the control valve. This series of investigations provides ample evidence that the hodograph transformation coupled with numerical computations is effective in dealing with problems of this nature.

Introduction

The study of discharge from a vessel has been of considerable interest in its simulation of practical flow conditions. Early studies were exclusively restricted to two-dimensional inviscid incompressible flow discharge, and the method of conformal mapping can be readily employed to solve these problems (e.g., Milne-Thompson, 1950; Mises, 1971). Inviscid flow analysis is adequate since the viscous effect often has negligible influence on the flow. For practical applications, the flow always assumes an axisymmetric geometry. Some theoretical studies were directed to inviscid compressible nozzle discharge to simulate flows from jet-propulsive systems. Brown and Chow (1974) studied the discharge from a conical nozzle on the basis of the tangent gas approximation. Their interest was directed toward its application to ejector-propulsive systems (Anderson, 1974). Later, Fenain et al. (1974) examined this type of problem on the basis of the hodograph transformation. Due to the complicated geometry, flow through a control valve has not been studied.

Recently, Chow et al. (1987) examined the incompressible flow discharge from a vessel through an axisymmetric control valve. As a usual practice, the effect of gravitation has been ignored. This boundary value problem has an unusual feature in that part of the boundary in the hodograph plane is overlapped. A specific manipulation which will be explained in more detail later was applied to produce the realistic jet boundaries under this condition. The discharge characteristics of the control valve have also been established.

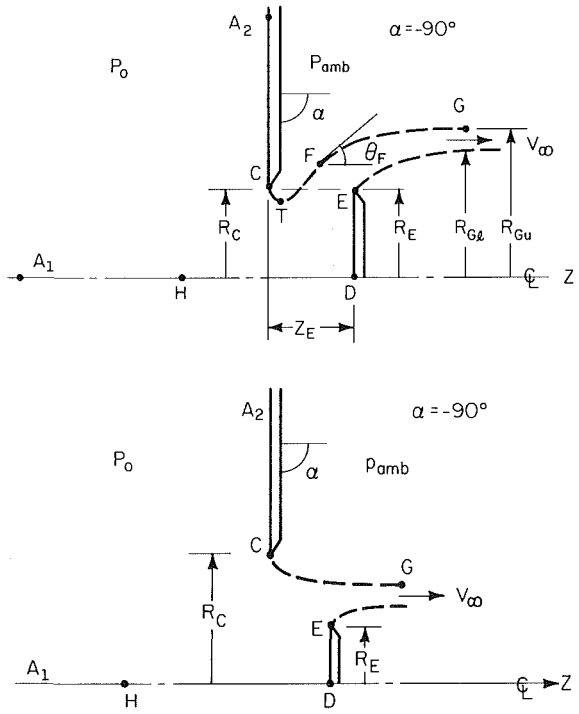


Fig. 1 Discharge through an axisymmetric control valve: (a) $R_E = R_c$, $\theta_F \neq 0$; (b) $R_E < R_c$, $\theta_F = 0$

The present effort is an extension of this study to the compressible flow regime. While similar considerations and treatment can be employed for conditions of subcritical pressure ratios, the sonic line and free jet boundaries under conditions of supercritical pressure ratios lead to additional complications. Details of this study follow.

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 10, 1986; final revision April 16, 1987.

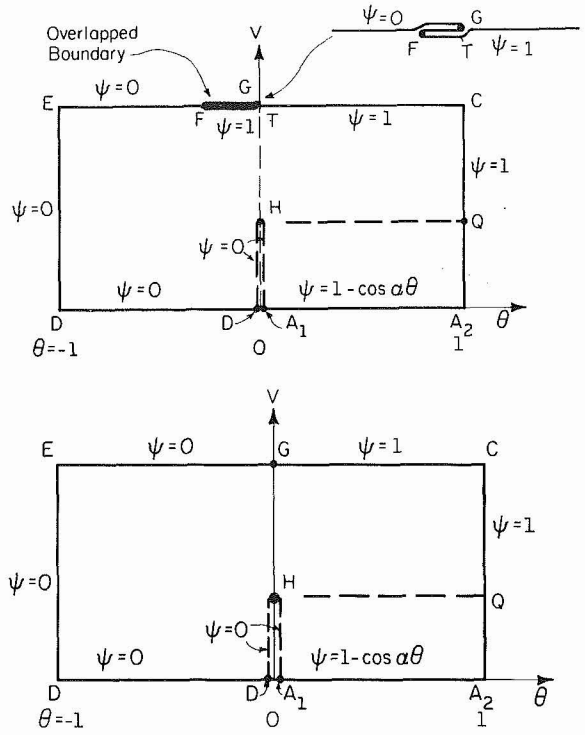


Fig. 2 The hodograph for subcritical discharge: (a) with overlapping boundary ($\theta_F < 0$); (b) without overlapping boundary ($\theta_F = 0$)

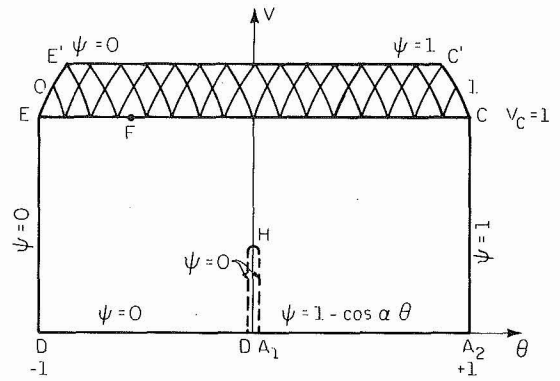


Fig. 3 The hodograph for supercritical discharge

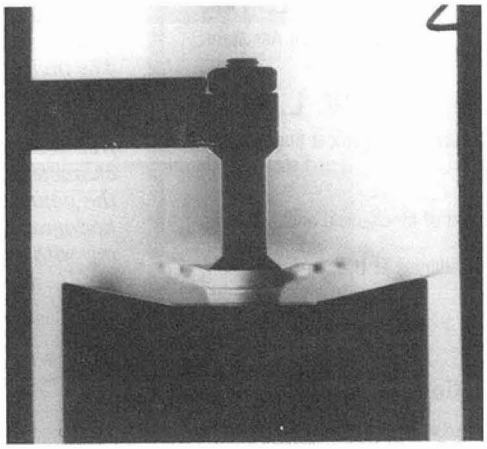


Fig. 4 A Schlieren photograph of a certain supercritical flow condition

Analytical Consideration

The general problem under consideration is illustrated in Fig. 1. The valve radius R_E may be smaller or equal to the opening radius R_C so that a variety of flow patterns may result. The reservoir is large enough that the approaching flow velocity is vanishingly small. The flow pattern may also be further classified to be subcritical or supercritical depending upon the pressure ratio P_{amb}/P_o. Since the effect of gravity is always ignored, the free streamlines are also lines of constant velocity according to Bernoulli's principle.

For subcritical flows, the asymptotic state must occur at a far downstream position with a uniform velocity in the horizontal direction. For a valve with a smaller radius, this asymptotic state can be reached with the top free streamline angle increasing monotonically after it leaves the valve opening (Fig. 1(b)) and its hodograph is shown in Fig. 2(b) where the streamline angle θ has been normalized by $\alpha(\alpha = -90$ deg). For an ordinary valve with $R_E = R_C$, the top free streamline must change itself into a state of positive maximum streamline angle (point F in Fig. 1(a)) before it settles down toward the asymptotic state. This behavior implicates a special feature that part of the boundary of its hodograph is overlapped (Fig. 2(a)). Since the stream function has already been normalized by the unknown discharge from the reservoir, the boundary value of the stream function changes discontinuously from zero to unity at the asymptotic state G, as it should.

For supercritical flows, the free streamlines assume supersonic Mach numbers. With a given pressure ratio, the sonic line and the initial part of the free jet boundaries must be established together with the internal elliptic flowfield. The corresponding hodograph is shown in Fig. 3 where curves on top of the figure represent the well-known characteristics of the simple wave of two-dimensional flow. Point F separates the sonic line EC into two branches. Details of this part of the hodograph will be explained later. It is to be noted that the present study shall not consider the case where the valve is located far enough downstream that the problem becomes a nonuniform supersonic jet past a blunt body. A Schlieren photograph of this situation is shown in Fig. 4, where a detached shock wave can be clearly seen.

In any event, the center streamline initially accelerates until an unknown velocity V_H is reached. Afterwards it decelerates toward the stagnation state D. A barrier of $\psi = 0$ is thus set up in the middle of the hodograph which corresponds to two different parts of the center streamline. Clearly V_H and θ_F for both subcritical and supercritical flows are important parameters for the problem.

Governing Equations. Hodograph transformation employs the magnitude V and the streamline angle θ of the velocity as independent variables while the potential function ϕ, stream function ψ, and also the physical coordinates Z and r are dependent variables. In parallel with the derivation of the hodograph transformation for a two-dimensional flow, Fenain et al. (1974) obtained the corresponding equations for the axisymmetric flow. This derivation is briefly outlined below.

From the conventional definitions for potential flow, one has $\phi_z = u = V\cos\theta$, $\phi_r = v = V\sin\theta$,[1] and $\psi_r = r\rho u$, $\psi_z = -r\rho v$ where ρ, the density, has already been normalized by ρ_{ref}, which is ρ_∞ for subcritical flow and ρ_o, the stagnation density, for supercritical flow. It follows immediately that

$$d\phi = V\cos\theta dz + V\sin\theta dr,\quad d\psi = -r\rho V\sin\theta dz + r\rho V\cos\theta dr. \tag{1a,b}$$

Upon treating ϕ and ψ as functions of V and θ, so that $d\phi =$

[1] Here subscripts denote differentiations.

$\phi_V dV + \phi_\theta d\theta$ and $d\psi = \psi_V dV + \psi_\theta d\theta$ one may solve dz and dr from equation (1) to obtain

$$dz = \left(\frac{\cos\theta}{V}\phi_V - \frac{\sin\theta}{r\rho V}\psi_V\right)dV + \left(\frac{\cos\theta}{V}\phi_\theta - \frac{\sin\theta}{r\rho V}\psi_\theta\right)d\theta \tag{2a}$$

$$dr = \left(\frac{\sin\theta}{V}\phi_V + \frac{\cos\theta}{r\rho V}\psi_V\right)dV + \left(\frac{\sin\theta}{V}\phi_\theta + \frac{\cos\theta}{r\rho V}\psi_\theta\right)d\theta \tag{2b}$$

Equation (2) implies that

$$Z_V = \frac{\cos\theta}{V}\phi_V - \frac{\sin\theta}{r\rho V}\psi_V,\ Z_\theta = \frac{\cos\theta}{V}\phi_\theta - \frac{\sin\theta}{r\rho V}\psi_\theta \tag{3a}$$

$$r_V = \frac{\sin\theta}{V}\phi_V + \frac{\cos\theta}{r\rho V}\psi_V,\ r_\theta = \frac{\sin\theta}{V}\phi_\theta + \frac{\cos\theta}{r\rho V}\psi_\theta \tag{3b}$$

since $Z_{\theta V} = Z_{V\theta}$ and $r_{\theta V} = r_{V\theta}$, equations (3$a$) and (3$b$) may be cross differentiated to yield

$$\phi_\theta = \frac{V\psi_V}{r\rho},\ \phi_V = \frac{1}{r\rho V}\left[\psi_\theta(M^2-1) + \frac{V}{2y}(y_\theta\psi_V - y_V\psi_\theta)\right] \tag{4a,b}$$

where $y = r^2/2$, $M = V/C$ and the relationship of $d\rho/dV = -\rho V/C^2$ has been introduced; C being the velocity of sound. Equation (3b) can be written now as

$$y_\theta = \frac{1}{\rho}\left[\sin\theta\,\psi_V + \frac{\cos\theta}{V}\psi_\theta\right] \tag{5a}$$

$$y_V = \frac{1}{\rho V}\left[\psi_V\cos\theta + \frac{(M^2-1)\sin\theta}{V}\psi_\theta + \frac{S}{V}\right] \tag{5b}$$

with

$$S = \frac{V(y_\theta\psi_V - y_V\psi_\theta)\sin\theta}{2y} \tag{5c}$$

Equation (3a) can also be written as

$$rz_\theta = \frac{1}{\rho}\left[\cos\theta\,\psi_V - \frac{\sin\theta}{V}\psi_\theta\right] \tag{6a}$$

$$rz_V = \frac{1}{\rho V}\left[\frac{\cos\theta(M^2-1)}{V}\psi_\theta - \sin\theta\,\psi_V + \frac{S}{V}\cot\theta\right] \tag{6b}$$

Since $y_{V\theta} = y_{\theta V}$, cross differentiations of equations (5a) and (5b) yield

$$V^2\psi_{VV} + V(M^2+1)\psi_V + (1-M^2)\psi_{\theta\theta} = \frac{1}{\sin\theta}\frac{\partial S}{\partial\theta} \tag{7}$$

where S can also be written as

$$S = \frac{(V^2\psi_V^2 + (1-M^2)\psi_\theta^2)\sin^2\theta}{2y\,\rho V + \sin\theta\,\psi_\theta} \tag{8}$$

on account of equations (5a) and (5b). It is recognized that the right-hand side of equation (7) is the only difference in the governing equations between a two-dimensional and an axisymmetric flow.

Perhaps it is worthwhile to note that the original motivation of adopting the method of hodograph transformation lies in its reduction of the nonlinear partial differential equation governing the compressible flow into a linear one in the hodograph plane. For the axisymmetric problem, this advantage is lost as a result of the existence of the nonlinear term on the right-hand side of equation (7). Nevertheless, the simplicity in its specification of the boundary condition of the free boundary problem in the hodograph plane and the ease of finding the solution to the equation through iterations as described later, make this approach very attractive to the present problem.

So far, all variables, other than density, are the original flow quantities. It will be convenient to normalize V, θ, z, r, ψ by $V_{\text{ref}}, \alpha, R_c, \psi_o$, respectively; ψ_o being the unknown rate of discharge per radian from the reservoir while $V_{\text{ref}} = V_\infty$ for the subcritical flow and $V_{\text{ref}} = V^*$ for the supercritical flow.[2] After normalization, equations (5) through (8) are changed, respectively, into

$$V^2\psi_{VV} + V(M^2+1)\psi_V + \frac{(1-M^2)}{\alpha^2}\psi_{\theta\theta} = \frac{1}{\alpha\sin\alpha\theta}\frac{\partial S}{\partial\theta} \tag{9}$$

$$S = \frac{(V^2\psi_V^2 + (1-M^2)\psi_\theta^2/\alpha^2)\sin^2(\alpha\theta)}{2yV\rho\left(\frac{R_c^2 V_{\text{ref}}}{\psi_o}\right) + \frac{\sin(\alpha\theta)}{\alpha}\psi_\theta} \tag{10}$$

$$y_\theta = \left(\frac{\psi_o}{R_c^2 V_{\text{ref}}}\right)\frac{1}{\rho}\left(\alpha\sin(\alpha\theta)\psi_V + \frac{\cos(\alpha\theta)}{V}\psi_\theta\right) \tag{11a}$$

$$y_V = \left(\frac{\psi_o}{R_c^2 V_{\text{ref}}}\right)\frac{1}{\rho}\left(\frac{\cos(\alpha\theta)\psi_V}{V} + \frac{(M^2-1)\sin(\alpha\theta)}{\alpha V^2}\psi_\theta + \frac{S}{V^2}\right) \tag{11b}$$

$$rz_\theta = \left(\frac{\psi_o}{R_c^2 V_{\text{ref}}}\right)\frac{1}{\rho}\left(\alpha\cos(\alpha\theta)\psi_V - \frac{\sin(\alpha\theta)}{V}\psi_\theta\right) \tag{12a}$$

$$rz_V = \left(\frac{\psi_o}{R_c^2 V_{\text{ref}}}\right)\frac{1}{\rho}\left(\frac{(M^2-1)\cos(\alpha\theta)}{\alpha V^2}\psi_\theta - \frac{\sin(\alpha\theta)}{V}\psi_V + \frac{S\cot(\alpha\theta)}{V^2}\right) \tag{12b}$$

where all variables are dimensionless. $\psi_o/R_c^2 V_{\text{ref}}$ becomes an important parameter. Since the actual rate of discharge from the vessel is $2\pi\rho_{\text{ref}}\psi_o$, this parameter is directly related to the discharge coefficient defined by

$$C_d = \psi_o/[(R_c^2 V_{\text{ref}})/2]. \tag{13}$$

Since the case of supercritical discharge needs additional treatment, the case of subcritical discharge is first discussed.

Subcritical Discharge. For a high back-pressure ratio, i.e., $P_{\text{amb}}/P_o \geq P^*/P_o$; P^* being the pressure corresponding to unity Mach number, the treatment is not much different from the case of incompressible flow (Chow et al., 1987). The S function on the right-hand side of equation (9) is initially assumed to be vanishing everywhere. With a pair of values of V_H and θ_F, the ψ function in the hodograph domain may be established through computations of equation (9) in its finite difference form. Partial derivatives ψ_V and ψ_θ may be formed accordingly. The y values may be obtained by integrating equation (11a) from H to Q in the hodograph (see Fig. 2(a)) and equation (11 (b)) from Q to C. The parameter $\psi_o/R_c^2 V_\infty$ is determined by imposing the end condition of $y_c = 1/2$ ($R_c = 1$). Once this is satisfied, the y values may be established throughout the whole domain. This allows the evaluation of S function and the right-hand side of equation (9). For all grid points along the line GH where $\theta = 0$ the right-hand side of equation (9) is evaluated by

[2] After normalization V is M^* for supercritical flow.

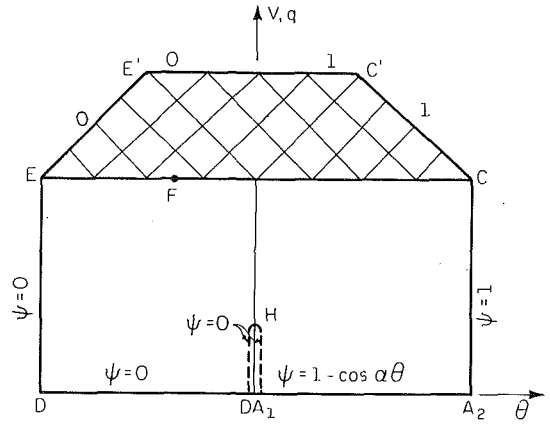


Fig. 5 Hodograph after transformation in the supersonic domain

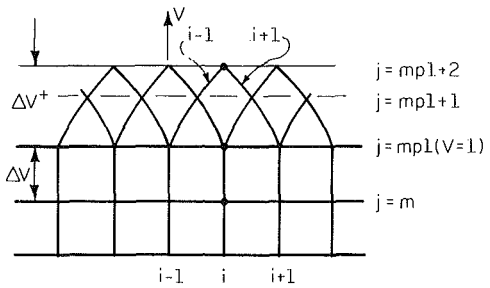


Fig. 6 Computational grid for sonic point

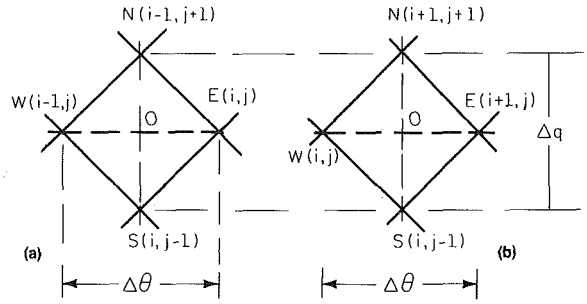


Fig. 7 Computational grid for supersonic point

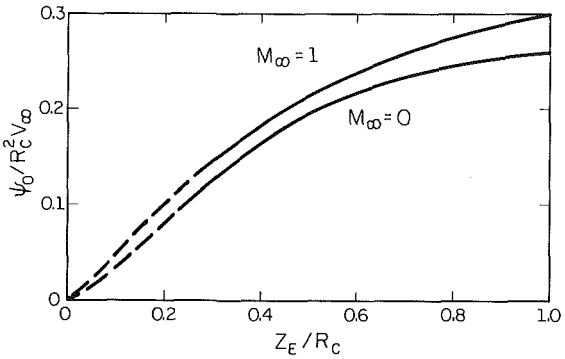


Fig. 8 Discharge characteristics for subcritical discharge

$$\frac{1}{\alpha \sin \alpha\theta}\left.\frac{\partial S}{\partial \theta}\right|_{\theta=0} = \frac{V^2\psi_V^2 + \dfrac{(1-M^2)}{\alpha^2}\psi_\theta^2}{\rho y V\left(\dfrac{R_c^2 V_\infty}{\psi_o}\right)}$$

Computations may be repeated until the change in the S value is less than a small margin for every grid point in the plane. It should be emphasized that under-relaxation of the S function must be used to reach a convergent solution. The solution is now established for the selected parameters V_H and θ_F. The physical location and the size (R_E) of the valve can be determined through additional integrations of equation (11). The geometry of the free streamlines may also be traced.

It should be stressed that for these computations with $\theta_F < 0$, the value of ψ on the overlapping boundary FG in Fig. 2(a) is unity since it is the value of the streamline immediately adjacent to the domain. The value of ψ on EF is zero, leaving the natural discontinuity of ψ at the point F. This is inherent to the method of hodograph transformation. In tracing the free streamline, there is no difficulty in establishing configurations corresponding to segments CF and EF. However, if unity value for ψ were used for the segment of the top streamline FG, so that the values of ψ_V there were the same as those of the segment TF, the configuration of FT in the physical plane would be retraced since the only difference between them is the direction of tracing. Similarly, if zero were used for ψ for the segment FG of the lower streamline, a corresponding curve with sharp curvature bending downward would be produced. They all do not realistically present these free jet configurations. After some experimentations, it was decided to adopt their average value ($\psi = 0.5$) for both portions of the free jet boundary. The supporting rationalities are (1) this manipulation yielded realistic free jet configurations, (2) it indicates that the streamline of $\psi = 0.5$ is tracing such a geometry, and (3) only under the condition of overlap, use of this average value of ψ is justified.

Usually, the valve radius R_E so established is smaller than R_c. By varying both values of θ_F and V_H, the flow pattern corresponding to the valve of $R_E = R_c$ situated at a specific location may be established. θ_F should always be negative for $R_E = R_c$. Computational experience has indicated that the size (R_E) of the valve depends strongly upon θ_F while the axial locations of the valve relies more on V_H.

Supercritical Discharge. For $P_{amb}/P_o < P^*/P_o$, the free streamlines are supersonic and a sonic line spans across the opening CE. Naturally the sonic line and the initial part of the free jet boundary must be solved together with the internal elliptic region of the flow. Even through the flow beyond the sonic line is supersonic, the problem may be classified as quasi-elliptic (i.e., elliptic although locally hyperbolic) since this portion of the flowfield including the shape of the sonic line is entirely influenced and determined by the pressure ratio P_{amb}/P_o. The present study benefited tremendously from a recent investigation (Wu and Chow, 1985) on a two-dimensional asymmetric nozzle as a vector-thrust device. Point F is identified on the sonic line which is the only point where waves of two opposite families intersect. These waves are either directly originated from the two corners of the nozzle or are multipally reflected from the upstream sonic line and jet boundaries. This point naturally divides the sonic line EC into two branches for the present problem. For convenience of numerical computations, the curved characteristics in the region $V > 1$ on the hodograph are transformed into straight lines making 45 deg angles with the horizontal and vertical axes (see Fig. 5). It is well known that

$$\theta = \pm f(V) + \text{constant} \tag{14}$$

represents the simple wave characteristics in a two-dimensional flow

where

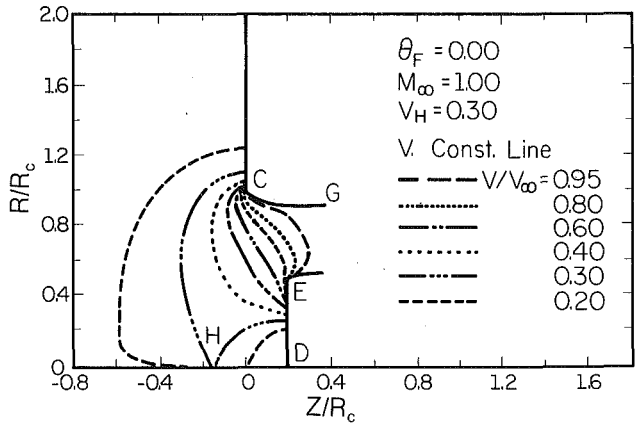


Fig. 9 Flow pattern for $\theta_F = 0$ (subcritical flow)

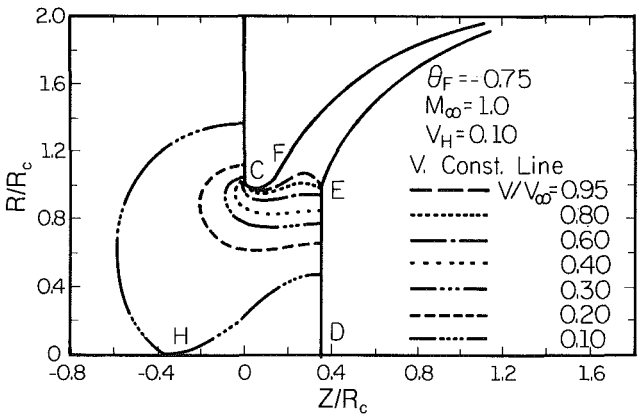


Fig. 10 Flow pattern for $\theta_F < 0$ (subcritical flow)

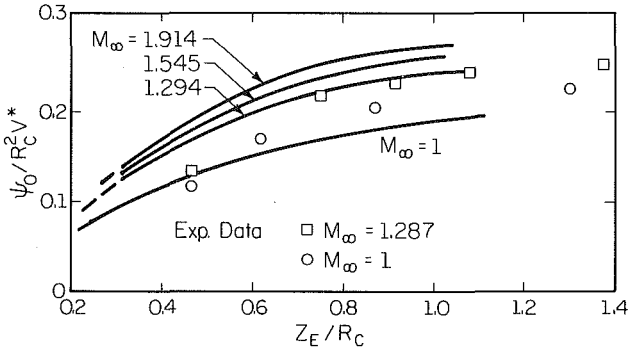


Fig. 11 Discharge characteristics for supercritical discharge

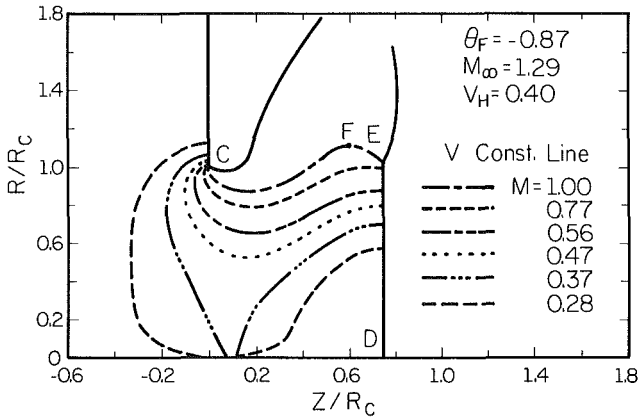


Fig. 12 Flow pattern for $M_\infty = 1.294$

$$f(V) = -\frac{1}{\alpha}\left(\sqrt{G}\tan^{-1}\sqrt{\frac{V^2-1}{G-V^2}} - \tan^{-1}\sqrt{\frac{G(V^2-1)}{G-V^2}}\right)$$

with $G = (\gamma+1)/(\gamma-1)$

Under this condition one introduces a variable q defined by

$$q = f(V) \text{ for } V > 1 \tag{15}$$

and the characteristic equation (equation (14)) becomes

$$\theta \pm q = \text{constant}. \tag{16}$$

In the subsonic region, the governing equation (9) remains in the same form. In the supersonic region, equation (9) is transformed into

$$\frac{1}{\alpha}(V^2-1)\psi_{qq} - \frac{2V^4}{(\gamma-1)\sqrt{G(V^2-1)(G-V^2)}}\psi_q + \frac{(1-V)^2}{\alpha}\psi_{\theta\theta} = \frac{1}{\sin(\alpha\theta)}\frac{\partial \bar{S}}{\partial \theta} \tag{17}$$

where $\bar{S} = [S(G - V^2)/G]$. This locally hyperbolic character can now be easily identified from this equation.

While the numerical treatment in the subsonic region follows the scheme similar to the case of subcritical discharge, the sonic line and the supersonic region under this condition must be dealt with in different special manners. For a grid point on the sonic line, equation (9) is simplified into

$$\psi_{VV} + 2\psi_V = \frac{1}{\alpha\sin(\alpha\theta)}\frac{\partial S}{\partial \theta}. \tag{18}$$

Referring to the transonic portion of the hodograph diagram (Fig. 6), the finite difference form of equation (18) is

$$C\psi_{i1,mp1+2} - (C+D)\psi_{i,mp1} + D\psi_{i,m} = \frac{\Delta V^+ \Delta V(\Delta V + \Delta V^+)}{2\alpha\sin(\alpha\theta)}\frac{\partial S}{\partial \theta} \tag{19}$$

where $C = \Delta V(1+\Delta V)$ and $D = \Delta V^+(1-\Delta V^+)$ (20)

with $i1 = i - 1$ for the left branch (i.e., $\theta_i < \theta_F$), and $i1 = i + 1$ for the right branch (i.e., $\theta_i > \theta_F$). ΔV^+ is given by

$$\Delta V^+ = V - 1 \tag{21}$$

where V satisfies

$$\Delta\theta = -\frac{1}{\alpha}\left[\sqrt{G}\tan^{-1}\sqrt{\frac{V^2-1}{G-V^2}} - \tan^{-1}\sqrt{\frac{G(V^2-1)}{G-V^2}}\right]. \tag{22}$$

For an unknown grid point (i, j) in the supersonic region of the hodograph (Fig. 7), we may find $\psi_{i,j}$ by writing equation (17) in its finite difference form around the point 0. It would yield

$$\psi_{i,j} = \psi_{i+1,j+1} + \psi_{i,j-1} - \psi_{i+1,j} - g(V_j)(\psi_{i+1,j+1} - \psi_{i,j-1}) - \frac{\alpha(\Delta q)^2}{4(V_j^2-1)}\left(\frac{1}{\sin(\alpha\theta)}\frac{\partial \bar{S}}{\partial \theta}\right)_o \tag{23}$$

for the right branch, where point W is the unknown grid point in Fig. (7(a)) and

$$\psi_{i,j} = \psi_{i-1,j+1} + \psi_{i,j-1} - \psi_{i-1,j} - g(V_j)(\psi_{i-1,j+1} - \psi_{i,j-1}) - \frac{\alpha(\Delta q)^2}{4(V_j^2-1)}\left(\frac{1}{\sin\alpha\theta}\frac{\partial \bar{S}}{\partial \theta}\right)_o \tag{24}$$

for the left branch where point E is the unknown grid point in Fig. (7(b)). $g(V_j)$ is given by

$$g(V_j) = \frac{\alpha\Delta\theta V_j^4}{2(\gamma-1)(V_j^2-1)\sqrt{G(V_j^2-1)(G-V_j^2)}} \tag{25}$$

It is now obvious that from the given boundary conditions on the hodograph, computations are possible by sweeping the hodograph from the two sides of the two branches toward the middle. Although the subsonic part of the flowfield is linked in a way similar to the subcritical discharge, there is no linkage between the two branches of the supersonic flow region even though there is a small overlapping triangular domain on the hodograph. Again the S (or $\bar{S}$) function is iterated and built up slowly until its change at every grid point is smaller than a small margin. The solution is thus established for the supercritical discharge condition.

With the established sonic line, the adjacent supersonic flow region, and the free jet boundaries, additional computations are continued with the well known method of characteristics for axisymmetric supersonic flow. Other than a small modification introduced near point F on the sonic line, the computational scheme is well known and is not presented here. Wave coalescence will set in soon and computations cannot proceed further unless one is prepared to account for the entropy gradient prevailing within the flow (Chang and Chow, 1974).

Results and Discussion

For the range of the subcritical pressure ratios, the discharge characteristics have been established for a valve of $R_E = R_c$ with $M_\infty = 1$ by using a 41 × 21 uniform grid. It is presented in Fig. 8 together with previous results for incompressible flow. Typical flow patterns and free streamline configurations for both $\theta_F = 0$ ($R_E < R_c$) and $\theta_F = -0.75$ (when $R_E = R_c$) are presented in Figs. 9 and 10. It is obvious that the velocity changes tremendously near the lip of the opening.

For supercritical pressure ratios, the discharge characteristics have also been obtained for $M_\infty = 1.294$, 1.544, and 1.914. They are presented in Fig. 11 including the previous results for $M_\infty = 1$ modified by a factor of ρ^*/ρ_o due to the change in the reference density from the two formulations. In both Figs. 8 and 11, solid lines indicate computed results with $V_H \geqq 0.05$; 0.05 being the smallest grid value of V in a 41 × 21 grid, while broken lines indicate expected values for $V_H < 0.05$ from extrapolations.

A series of simple experiments was also carried out in the laboratory. Other than data points at small valve opening conditions, the agreement between the results of computation and the corresponding experimental data is reasonably good. It was also learned that at small opening positions, the outside wall of the model interfered with the discharge. A typical flow pattern obtained for $M_\infty = 1.294$ is also presented in Fig. 12 where the sonic line and the free jet boundaries are also shown. The extremely large curvatures of the sonic line at edges of the nozzle opening made it impossible to show the details within these regions.

In the case of supercritical discharge, it has been learned through computations that the ψ function varies drastically in the vicinity of point F along the sonic line indicating a fairly uniform flow there. It is for this reason that a coarser 31 × 21 uniform grid has been employed for computations of supercritical discharge. Furthermore, it is also convenient to keep point F in the middle of the two grid points. Since point F must be a point of large curvature on the sonic line and the sonic line has been obtained from finite grid computations, the Mach number of the two adjacent grid points on the sonic line has been modified to 1.035 for the subsequent computations of the jet by the method of characteristics. It is believed that this modification is not serious enough to affect the accuracy of the results of subsequent computations.

It should be mentioned that for the case of subcritical discharge such as that shown in Fig. 1, the pressure within the jet is always higher than the ambient pressure and the later portion of the free jet boundary must assume a downward-curved geometry. For the top free jet boundary in Fig. 1, gravitation is responsible for balancing effects of both curvature and pressure gradient (Chow et al., 1987). It is, however, the usual practice that the gravity effect be ignored for the present problem, especially so for a compressible fluid. Otherwise, the problem becomes three-dimensional and it would be impossible to treat the problem under the present scheme. Judging from the agreement between the computational results and experimental data, the present scheme is useful, at least in predicting the discharge characteristics under this condition. There is no doubt that the present study can be extended to consider cases of valves with different geometries including valves with stems, finite approaching flow velocities, and reservoir walls of different angles.

The longest computational time for each case of the present series of investigation required only 60 seconds on the Cyber 175 computing system.

Acknowledgment

The authors are grateful to Mr. L. Westendorf for making the valve models for the testing and to Messrs. M. Morris and B. Carroll for their help in conducting the tests in the laboratory. The authors are also grateful to Forsyth Chair Professor Robert H. Page at the Texas A & M University for his constant interest and encouragement of this work.

References

Anderson, B. H., 1974, "Assessment of an Analytical Procedure for Predicting Supersonic Ejector Nozzle Performance," NASA TND-7601.

Brown, E. F., and Chow, W. L., 1974, "Critical Flow Through Convergent Conical Nozzles," *Proc., 1st Symp. on Flow, The Measurement and Control in Science and Industry*, Inst. Soc. of America, Vol. 1, pp. 231–240.

Chang, I. S., and Chow, W., L., 1974, "Mach Disk from the Underexpanded Axisymmetric Nozzle Flows," *AIAA J.*, Vol. 12, No. 8, pp. 1079–1082.

Chow, W. L., Ting, Alice A. L., and Tsai, P. H., 1987, "Discharge from a Vessel through An Axisymmetric Control Valve," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, No. 2, pp. 447–452.

Fenain, M., Dutouquet, L., and Solignac, J. L., 1974, "Calcul des Performances d'une Tuyere Propulsive Convergente," *Proc. of the 9th Congress of the International Council of the Aeronautical Sciences*, Haifa, Vol. 2, pp. 691–705.

Milne-Thompson, L. M., 1950, *Theoretical Hydrodynamics*, 2nd Ed., McMillan, New York, p. 285.

Mises, R. V., 1917, "Berechnung von Ausfluss und Veherfallzahlen," *Zeitschrift des Vereines Deutschen Ingenieure*, Vol. 61, No. 22, pp. 469–474.

Wu, C., and Chow, W. L., 1985, "Study of an Asymmetric Flap Nozzle as a Vector-Thrust Device," *AIAA J. of Propulsive and Power*, Vol. 1, No. 4, pp. 286–291.

Journal of Applied Mechanics

Design Data and Methods

Utilizing Dynamic Stability to Orient Parts

N. C. Singer[1,3] and W. P. Seering[2,3]

The intent of this research is to study the dynamic behavior of a solid body resting on a moving surface. Results of the study are then used to propose methods for controlling the orientation of parts in preparation for automatic assembly. Two dynamic models are discussed in detail. The first examines the impacts required to cause reorientation of a part. The second investigates the use of oscillatory motion to selectively reorient parts. This study demonstrates that the dynamic behaviors of solid bodies, under the conditions mentioned above, vary considerably with small changes in geometry or orientation.

1 Introduction and Problem Definition

Conventionally, small parts are oriented by bowl feeders. These machines vibrate and thereby convey parts through a series of filters which reject all but a particular orientation. Rejected parts are returned to storage. Many researchers have examined the implementation of programmable or adjustable filter stages (Boothroyd, 1975; Boothroyd and Ho, 1977; Boothroyd and Murch, 1970; Boothroyd et al., 1982; Boothroyd et al., 1977; Lozano-Pérez, 1986; Murch, 1977; Murch and Boothroyd, 1975; Murch and Poli, 1977; Redford et al., 1983a; Redford et al., 1983b; Singer, 1985; Suzuki and Kohno, 1981). These techniques are often successful for limited classes of parts. The scope of this paper is to present theoretical results which may be useful in feeder designs. It is hoped that a more detailed understanding of the dynamic behavior of bodies resting on a moving surface will facilitate the design of new types of programmable parts feeders.

2 Impact Reorientation

In this section, we determine the conditions necessary to cause part reorientation. A part starts in a stable orientation (a natural resting aspect, (Boothroyd and Ho, 1977)) on a flat surface. It is then given an initial horizontal velocity after which it impacts a wall near the surface upon which it is resting (Fig. 1). The equations describing the impact are based on conservation of angular momentum about the impact point. After impact, the principle of conservation of energy is applied in order to determine whether reorientation occurs.

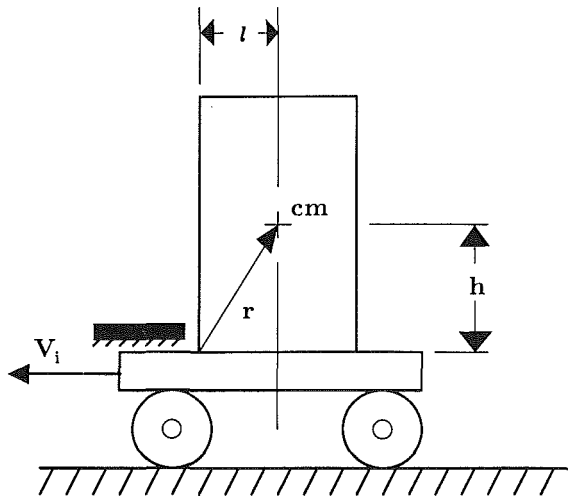


Fig. 1 Impact reorientation model

The part shown in Fig. 1 is a cylinder, but the analysis easily generalizes to any object. The important parameters are the mass, m of the part, the moment of inertia, I, and the constants, l and h, which locate the center of mass, cm, with respect to the impact point. The expressions for angular momentum, L_i, before and, L_f, after impact are (Klepner and Kolenkow, 1973):

$$L_i = mV_ih$$
$$L_f = \frac{I_aV_f}{\sqrt{h^2+l^2}}. \tag{2.1}$$

where V_i is the speed of the center of mass prior to impact, V_f is its speed just after impact, and I_a, the inertia about the impact point, is defined as

$$I_a = I_c + (h^2+l^2)m$$

with I_c the inertia about the center of mass. Equating kinetic energy just after impact to the change in potential energy which will cause the part to reorient yields

$$\frac{1}{2}I_a\left(\frac{V_f^2}{h^2+l^2}\right) = mg[(h^2+l^2)^{1/2}-h] \tag{2.2}$$

Solving equations (2.1) for V_f, substituting into equation (2.2), and then solving the result for the critical value of V_i yields

$$V_{\text{crit}} = \frac{\sqrt{2g}[(h^2+l^2)^{1/2}-h]^{1/2}[(h^2+l^2)+I_c/m]^{1/2}}{h}, \tag{2.3}$$

[1]Research Assistant, Mem. ASME.
[2]Associate Professor, Mem. ASME.
[3]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.
Manuscript received by ASME Applied Mechanics Division, April, 2, 1987.

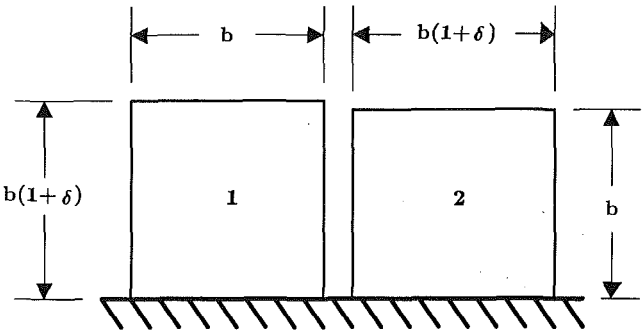


Fig. 2(*a*) Two similar parts

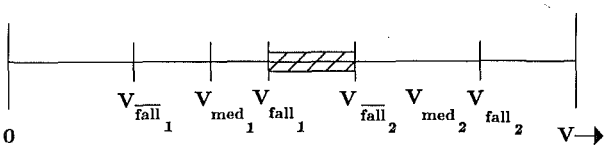


Fig. 2(*b*) Distinguishing the parts in Fig. 2(*a*)

which is an equation for the minimum initial horizontal velocity that will cause reorientation. For cylinders, I_c/m is $\frac{1}{3}\ h^2 + \frac{1}{4}\ l^2$; for rectangular parts, I_c/m is $\frac{1}{3}\ (h^2 + l^2)$.

In order to test the ability of impact reorientation to discriminate among different orientations, it is useful to consider a part which has two very similar orientations. Figure 2(*a*) shows a rectangular solid part in two possible orientations. One edge is $b\delta$ longer than the other where δ is expressed as a fraction less than one. As the part exchanges orientations, the center of mass location changes in both height and horizontal distance from the impact point by $b\delta/2$. The ratio of V_{crit} of orientations one and two (denoted by V_1 and V_2) expanded in a power series in δ is expressed as:

$$\frac{V_2}{V_1} = 1 + 2.2\delta + 1.33\delta^2 \ldots . \qquad (2.4)$$

This equation illustrates that small differences in location of the center of mass result in significant variations in impact reorientation velocity.

To verify the equations and to get a measure of experimental error, tests were performed. Cylinders of different size were placed on a moving conveyor belt. The cylinders impacted a steel wall, 0.889 mm (0.035 in.) high. Fifteen trials were made for each velocity set point and the number of pegs which reoriented was counted. The velocity of the belt was monitored by a tachometer which was friction driven directly by the belt.

Several tests were performed. First, the lower threshold velocity, $V_{\overline{fall}}$, was determined by finding the maximum velocity at which no pegs fell over in fifteen trials. The upper threshold velocity, V_{fall}, was determined similarly by finding the lowest velocity at which all pegs fell over. The median velocity V_{med} is simply the center of the velocity band defined by $V_{\overline{fall}}$ and V_{fall}. Finally, two percentages have been determined from this data. The percentage of error is calculated between the theoretical velocity, V_{Theory}, and the median experimental velocity, V_{med}. The uncertainty band of the experiment is calculated from $V_{\overline{fall}}$ and V_{fall}; this band is represented as a percentage of V_{med}.

Results confirm that the system model adequately represents the experimental situation. All tests performed had narrow uncertainty bands. This indicates that there is a sharp cutoff between not having enough energy to cause reorientation and having enough energy to guarantee reorientation. In addition, the error percentages between V_{Theory} and V_{med} are small. Furthermore, the scaling of the velocity agrees with the scaling predicted in equation (2.3). Two pegs which vary greatly in size and length to diameter ratio can have the same reorientation velocity. Table 1 gives typical results for two different cylinders.

Table 1 Selected impact reorientation results for two parts which are physically different yet possess the same reorientation velocity

Test Object		Velocities (Meters/Sec)				Errors	
No	Size	V_{Theory}	$V_{\overline{fall}}$	V_{fall}	V_{med}	% Error	Unc. Band
1	50.8 mm dia 113 mm long	0.409	0.399	0.429	0.414	1.2 %	± 3.7 %
2	25.4 mm dia 34 mm long	0.409	0.384	0.429	0.406	0.6 %	± 5.6 %

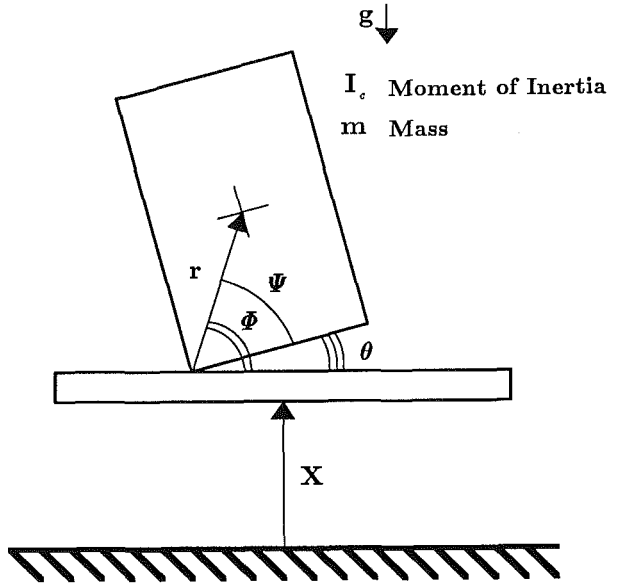


Fig. 3 Vibration reorientation model

The ability to discriminate between orientations, or to reorient one object while not reorienting another, depends heavily on the quality of the velocity source. Fortunately, excellent velocity sources are inexpensive and easy to build. Figure 2(*b*) presents a velocity line with the reorientation velocities for the two parts shown in Fig. 2(*a*). The dark section of the line identifies the desired operating velocity range. Within this range part 1 will always be reoriented, while part 2 will never be reoriented. By cascading several of these impacts at different velocities, a set of like parts, initiallly oriented randomly can all be driven to the same orientation.

3 Vibratory Reorientation

3*a* Modelling.

The model for this section is an object in a gravity field. The table on which the object rests oscillates vertically (Fig. 3). The object is first given some initial angle, ϕ_0, possibly by a sudden motion of the table. For a proper choice of $X(t)$, the table driving function, the given object can be maintained in a rocking motion of fixed amplitude. The driving function of the table can then be changed, causing an increase in the amplitude at which the object will continue rocking. Finally, the amplitude can be increased until the object reorients.

Because of the discontinuity of motion as the part impacts the table, the solution to the equations governing the system must be broken into regions. First, the region in which there is

no impact will be examined. The equations of this motion may be generated using Lagrange's method. The potential energy function can be expressed as

$$V = mg(r \sin \phi + x), \tag{3.1}$$

where V is the potential energy, m is the part mass, r is the distance from the contact point to the center of mass, ϕ is the angle between the horizontal and the vector r, and x is the vertical position of the table. The kinetic coenergy can be expressed as:

$$T^* = \frac{1}{2} I_c \left(\frac{d\phi}{dt}\right)^2 + \frac{1}{2} m \left[\frac{dx}{dt} + r \frac{d\phi}{dt} \cos\phi\right]^2 + \frac{1}{2} m \left[r \frac{d\phi}{dt} \sin\phi\right]^2, \tag{3.2}$$

where I_c is the moment of inertia of the part about the center of mass. These two terms can be used to form the Lagrangian which is then substituted into Lagrange's equation, yielding the two equations of motion for the system,

$$(I_c + mr^2) \frac{d^2\phi}{dt^2} + mr \cos\phi \left(g + \frac{d^2x}{dt^2}\right) = 0$$
$$m \frac{d^2x}{dt^2} + mr \frac{d^2\phi}{dt^2} \cos\phi - mr \left(\frac{d\phi}{dt}\right)^2 \sin\phi + mg = F, \tag{3.3}$$

The first of the equations gives the basic motion of the object, while the second equation is an expression for F, the contact force. These equations are valid in the conservative regions without impact.

The next step in formulating the model is to represent the impact of a part with the table. At impact, the center of mass of the object has a vertical velocity which is reflected with a coefficient of restitution, ϵ. In a perfectly elastic impact, ϵ is equal to one. Physical experiments were performed to determine the coefficient of restitution for aluminum parts rocking on an aluminum plate. The coefficient of restitution was determined to be between 0.7 and 0.75 for these tests. For all of the simulations in this paper, an ϵ of 0.5 was used to be conservative.

As a part hits the table, the impact will also affect the horizontal velocity. If the table is frictionless, the part will slide. In this derivation the table was assumed to be rough enough so that the contact surface of the part will not slide on the table and after impact the part will continue to rock by rotating about the new contact point.

The next step in creating steady rocking motion was to determine table acceleration profiles which would cause a part to oscillate at fixed amplitude. The amplitude of vibration is coupled with the frequency of oscillation. As the part gains energy, both the amplitude and the period of oscillation become greater. This response makes analysis difficult; few tools are available for predicting the behavior of this type of system (Meirovitch, 1975).

An analytical solution to the system equations for time required for passage from an initial to a final angle can be derived for the case in which the driving function, d^2x/dt^2 remains constant. The derivation of this solution is based on conservation of energy in the region without impact. For a constant table acceleration, the part can be considered to be in a conservative field as shown in equation (3.3*a*). In this particular case the energy balance equation becomes

$$\frac{1}{2}(I_c + mr^2) \left(\frac{d\phi_0}{dt}\right)^2 + m\left(g + \frac{d^2x}{dt^2}\right) r \sin\phi_0 = \frac{1}{2}(I_c + mr^2) \left(\frac{d\phi}{dt}\right)^2 + m\left(g + \frac{d^2x}{dt^2}\right) r \sin\phi, \tag{3.4}$$

where ϕ_0 is the starting angle of the object, and ϕ is some angle of interest. Note that the derivative of this expression is the result obtained in equation (3.3*a*). Equation (3.4) can be solved for $d\phi/dt$,

$$\frac{d\phi}{dt} = \sqrt{\frac{2m\left(g + \frac{d^2x}{dt^2}\right) r}{I_c + mr^2}} \times \sqrt{\frac{(I_c + mr^2)\left(\frac{d\phi_0}{dt}\right)^2}{2m\left(g + \frac{d^2x}{dt^2}\right) r} + \sin\phi_0 - \sin\phi}. \tag{3.5}$$

This equation can in turn be solved by separation of variables and then integration with respect to time. For convenience the constant A is defined as

$$A = \frac{(I_c + mr^2)\left(\frac{d\phi_0}{dt}\right)^2}{2m\left(g + \frac{d^2x}{dt^2}\right) r} + \sin\phi_0, \tag{3.6}$$

and the result is

$$\int_{\phi_0}^{\phi_e} \frac{d\phi}{\sqrt{A - \sin\phi}} = \int_0^T \sqrt{\frac{2m\left(g + \frac{d^2x}{dt^2}\right) r}{I_c + mr^2}}\, dt, \tag{3.7}$$

where ϕ_e is the ending angle of interest. The right side is a simple integral; the left side is an elliptic integral. The solution to this equation will be presented as a combination of elliptic functions. Though elliptic functions cannot be expressed in terms of elementary functions, they are well understood and are easily computed. In order to show that equation (3.7) becomes an elliptic integral, a change of variables must be made. Let

$$w = \phi + \frac{\pi}{2}. \tag{3.8}$$

so that

$$\sin\phi = -\cos w, \tag{3.9}$$

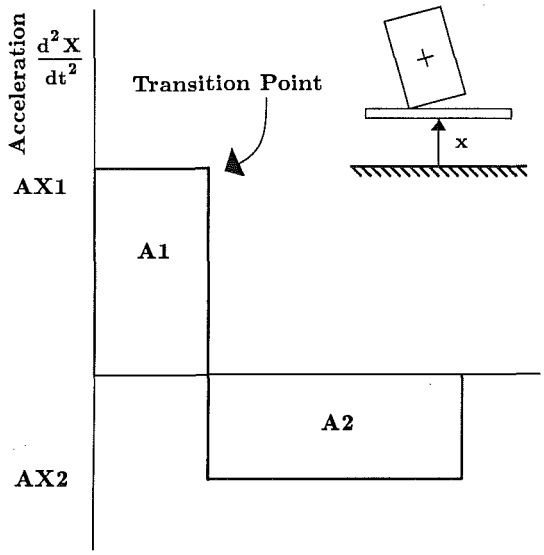


Fig. 4 Square wave acceleration input

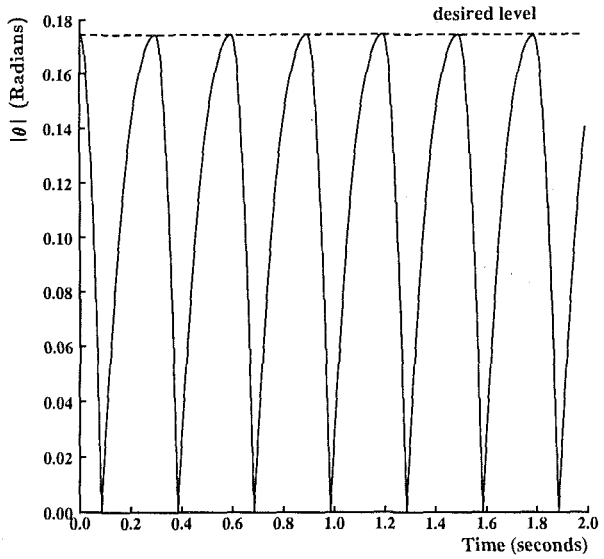


Fig. 5(a) Response of a part to input in Fig. 5(b)

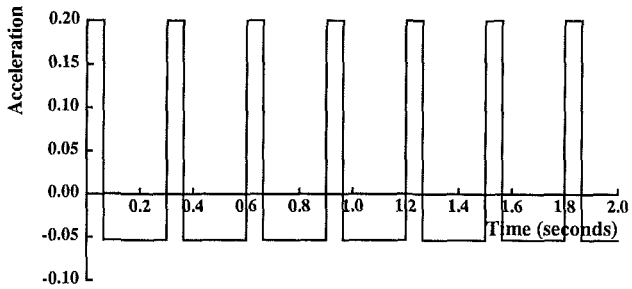


Fig. 5(b) Square wave input

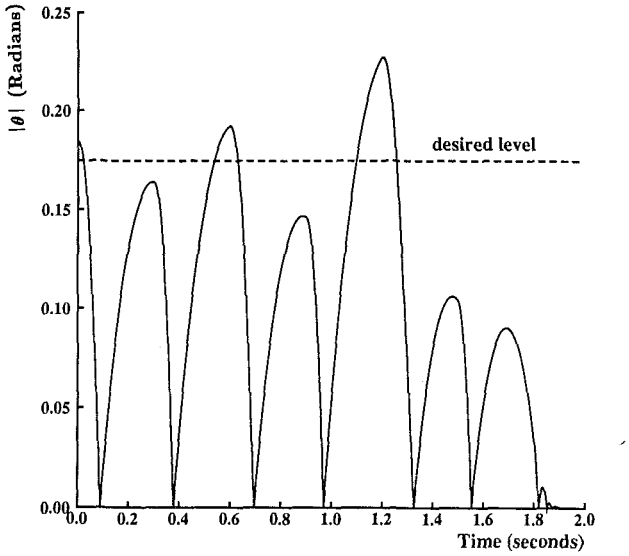


Fig. 6 Unstable response for a part starting 6 percent above the desired level

and

$$\cos\phi d\phi = \sin w dw, \tag{3.10}$$

From these we get

$$d\phi = \frac{\sin w dw}{\sqrt{1-\cos^2 w}} = dw. \tag{3.11}$$

Substituting equations (3.11) and (3.8) into the left-hand side of equation (3.7) produces the following expression

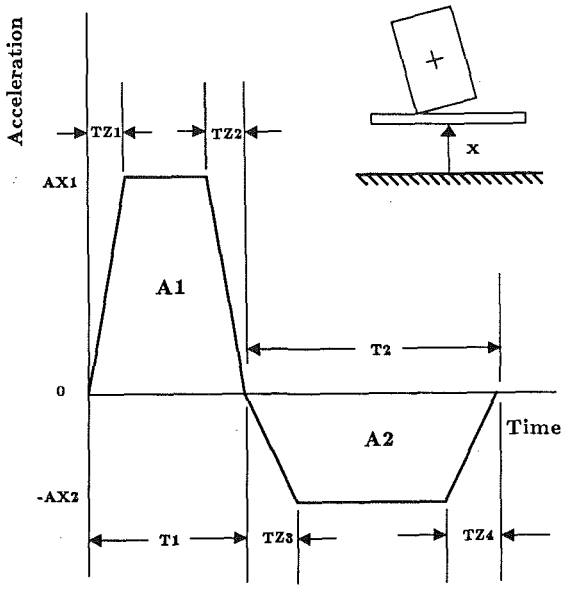


Fig. 7 Ramped square wave input

$$\int_{\phi_0}^{\phi_e} \frac{d\phi}{\sqrt{A-\sin\phi}} = \int_{\phi_0+\pi/2}^{\phi_e+\pi/2} \frac{dw}{\sqrt{A+\cos w}}, \tag{3.12}$$

In the case of impact

$$\phi_e = \psi,$$

where ψ is defined as the angle between the bottom of the part and the vector r from the contact point to the center of mass. If the integral is then rewritten as two integrals starting from zero, equation (3.12) can be expressed in terms of the elliptic integral of the first kind (Gradshteyn and Ryzhik, 1980):

$$\int_0^{\phi_e+\pi/2} \frac{dw}{\sqrt{A+\cos w}} - \int_0^{\phi_0+\pi/2} \frac{dw}{\sqrt{A+\cos w}}$$

$$= \begin{cases} \sqrt{2}(F(x_1, k_1) - F(x_2, k_1)) & A<1 \\ \sqrt{\dfrac{2}{1+A}}(F(x_3, k_2) - F(x_4, k_2)) & A>1 \end{cases}. \tag{3.13}$$

The elliptic integral of the first kind expressed in normal trigonometric form is

$$F(x, k) = \int_0^{\arcsin x} \frac{d\alpha}{\sqrt{1-k^2\sin^2\alpha}}, \tag{3.14}$$

and the limits are

$$x_1 = \sqrt{\frac{1+\sin\phi_0}{1+A}}$$

$$x_2 = \sqrt{\frac{1+\sin\phi_e}{1+A}}$$

$$k_1 = \sqrt{\frac{1+A}{2}}$$

$$x_3 = \frac{\phi_0}{2} + \frac{\pi}{4}$$

$$x_4 = \frac{\phi_e}{2} + \frac{\pi}{4}$$

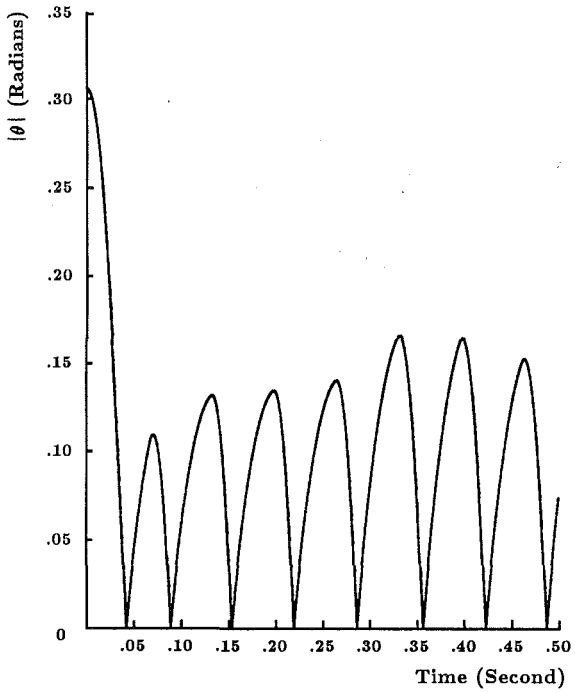


Fig. 8(*a*) Stable response to input in Fig. 8(*b*)

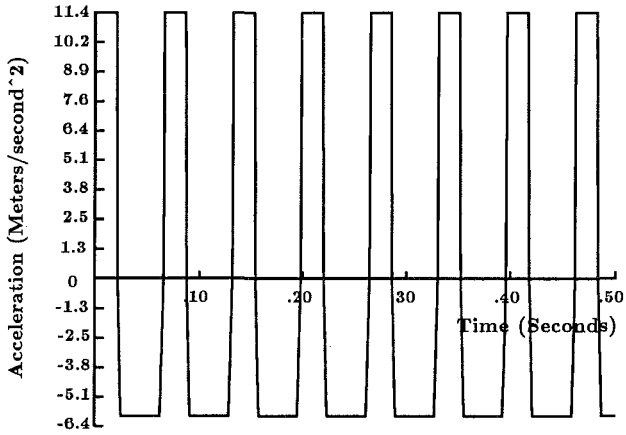


Fig. 8(*b*) Ramped square wave input

$$k_2 = \sqrt{\frac{2}{1+A}} \tag{3.15}$$

in terms of the system parameters.

3*b* Choice of Driving Function.

As mentioned previously, this analysis is correct only for constant values of d^2x/dt^2. The first driving function to be considered was a periodic rectangular wave in acceleration (Fig. 4). Four parameters must be selected to fully specify the waveform—the acceleration upward, $AX1$; the acceleration downward, $AX2$; and the switch times, $T1$ and $T2$. The selection of these parameters is constrained by several relations. First, the areas $A1$ and $A2$ must be equal in order for the table to return to its initial position and velocity after each cycle. Second, the part must start and finish the cycle at the same angle (the stability requirement). Lastly, the acceleration downward must never exceed gravity. This constraint is set so that the part does not lose contact with the table.

Equations (3.13) and (3.15) were substituted into a nonlinear equation solving program which calculated accelerations and switch times that brought a chosen part back to its initial angle in exactly one cycle. Figure 5 shows a time history of the acceleration of the table, d^2x/dt^2, and of the angle ϕ as determined by numerical simulation of equation (3.3) including the model for impact. For this case, the rocking motion is stable. However, small parameter variations cause an instability. Figure 6 shows the rocking motion for an object given a slightly larger initial amplitude than that of Fig. 5. After the first impact, the peak amplitude is too low; after the second, it is too high. Eventually, the phasing between input and output is lost and the part ceases to rock.

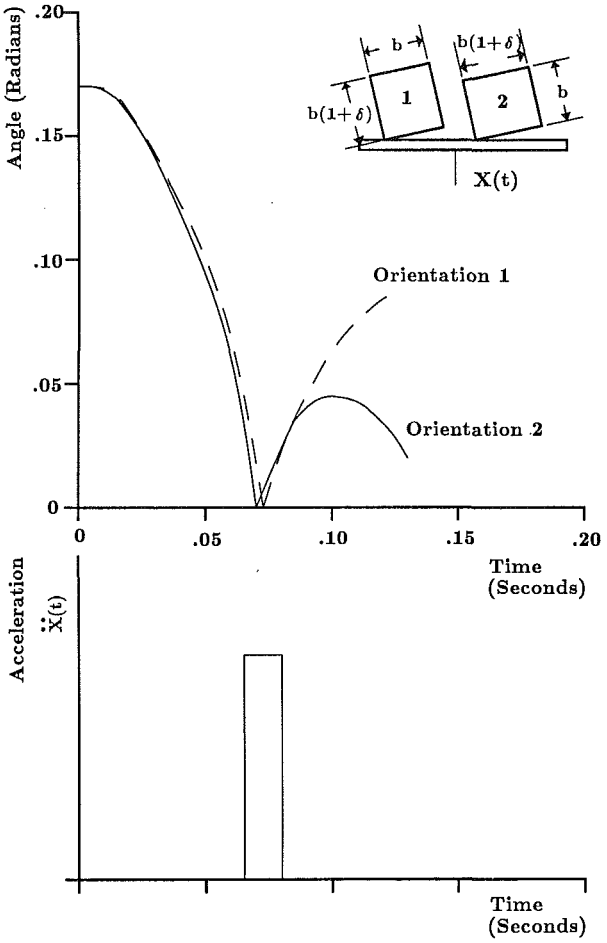


Fig. 9 Ability to separate orientations

Because our goal is to bring parts to a stable rocking motion from a range of initial conditions, the rectangular acceleration waveform will not do the job. A variety of waveforms were tested. One which exhibited good performance is shown in Fig. 7. It is simply the rectangular waveform with the edges sloped. Figure 8 shows results from a numerical simulation using this forcing function as input. A part is started with an initial angle which is twice the steady state angle yet it still reaches a stable equilibrium motion. This suggests that by using the proper waveform, a great range of initial conditions of the part can be tolerated.

If this method is to be used to selectively reorient parts, different parts should exhibit significantly different responses. Figure 9 shows variation in response for two similar parts given the same initial angle and subjected to the same acceleration profile. For these simulations, upward acceleration was set to occur around the time of impact. Several simulations were executed with various length "windows" of upward acceleration. The ability to discriminate degraded as the window became very large in comparison with the difference in impact times.

An attempt was made to design an experimental apparatus for study of the behavior described above. No satisfactory configuration could be created at a reasonable cost. To drive reasonably large parts with periods of oscillation greater than 0.5 seconds would require a table with a range of motion substantially greater than that of available electromagnetic shakers. Hydraulic shakers are designed to deliver large forces; they have servo valves which are too small to supply the flow rates necessary to achieve the required velocities. The ideal driving system would be a small diameter hydraulic actuator with a relatively large servo valve.

4 Conclusion

This study of part motion has demonstrated that small variations in inertial properties from one part to another cause significant differences in the dynamic behavior of the two parts. Two example techniques which capitalize on this property were presented. Practical implementation of the first technique, impact reorientation, is feasible given current design technology. The second technique, vibratory reorientation, will require construction of oscillating tables with performance characteristics available only through custom design.

Acknowledgments

The first author was supported through the Office of Naval Research Graduate Fellowship Program. The research described in this paper was performed at the Massachusetts Institute of Technology Artificial Intelligence Laboratory. The laboratory's research is funded in part by the Office of Naval Research under ONR contract N00014-81-K-0494 and in part by the Defense Advanced Research Projects Agency of the United States Department of Defense under ONR contract N00014-85-K-0124.

References

Boothroyd, G., 1975, "Automatic Handling of Small Parts," *Annals of the International Institute for Production Engineering Research* (Ann CIRP), Vol. 24, No. 1, pp. 393–398.

Boothroyd, G., and Ho, C., 1977, "Natural Resting Aspects of Parts for Automatic Handling," *Transactions of the ASME*, May, pp. 314–317.

Boothroyd, G., and Murch, L. E., 1970, "Performance of an Orienting Device Employed in Vibratory Bowl Feeders," *Transactions of the ASME*, Aug. pp. 694–698.

Boothroyd, G., Poli, C., and Murch, L. E., 1982, *Automatic Assembly*, Marcel Dekker, New York.

Boothroyd, G., Poli, C. R., and Murch, L. E., 1977, "Handbook of Feeding and Orienting Techniques for Small Parts," Automation Project, Department of Mechanical Engineering, University of Massachusetts, Amherst, MA.

Gradshteyn, I. S., Ryzhik, I. M., 1980, *Table of Integrals, Series, and Products*, Academic Press, New York, pp. 460–464.

Klepner, D., Kolenkow, R. J., 1973, *An Introduction To Mechanics*, McGraw-Hill, New York.

Lozano-Pérez, T., 1986, "Motion Planning and the design of Orienting Devices for Vibratory Parts Feeders," To be published in *IEEE Journal Of Robotics And Automation*, MIT AI Laboratory.

Meirovitch, L., 1975, *Elements of Vibration Analysis*, McGraw-Hill, New York.

Murch, L. E., 1977, "Feeding and Orienting Parts Automatically," SME Technical Paper, AD77-707.

Murch, L. E., and Boothroyd, G., 1975, "Feeding Small Parts for Assembly," *American Machinist*, Oct., pp. 106–110.

Murch, L. E., and Poli, C., 1977, "Analysis of Feeding and Orienting Systems for Automatic Assembly, Part 2," *Transactions of the ASME*, May, pp. 308–313.

Redford, A. H., Lo, E. K., and Killeen, P., 1983a, "Parts Feeder for A Multi-Arm Assembly Robot," Department of Aeronautical and Mechanical Engineering, University of Salford, U. K., unpublished copy of a manuscript presented at the 15th CIRP International Seminar on Manufacturing Systems, Amherst, MA, June, 1983.

Redford, A. H., Lo, E. K., and Killeen, P., 1983b, "Parts Presentation To Multi-Arm Assembly Robots," Department of Aeronautical and Mechanical Engineering, University of Salford, U. K., unpublished copy of a manuscript published in the *CIRP Annals*, Vol. 32/1, presented at the CIRP General Assembly, Harrogate, Aug. 1983.

Singer, N., 1985, "Utilizing Dynamic and Static Stability to Orient Parts," SM Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Suzuki, T., and Kohno, M., 1981, "The Flexible Parts Feeder which Helps a Robot Assemble Automatically," *Assembly Automation*, Feb., pp. 86–92.

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 Critical Speed of a Rotating Circular Plate

S. Chonan[1]

Introduction

It is well known that the response amplitude of a rotating circular plate subjected to a stationary lateral load increases rapidly as the rotation speed approaches a certain value, generally referred to as the critical speed. There are many papers which are concerned with the critical speed of rotating plates (Mote, Jr., 1965, 1967; Eversman and Dodson, 1969; Iwan and Stahl, 1973; Iwan and Moeller, 1976; Benson and Bogy, 1978). In this note a simple arithmetic equation is presented for the determination of the critical speed of plates with a clamped inner boundary and free outer boundary.

Governing Equations

Consider a plate of inner radius b and outer radius a rotating freely with a constant angular speed Ω about its axis. Denoting the lateral displacement of the plate by w the equation of motion is, with respect to the coordinate frame (r,θ) rotating with the plate,

$$D[\partial^2/\partial r^2+(1/r)\partial/\partial r+(1/r)^2\partial^2/\partial\theta^2]^2 w + \rho h[\partial/\partial t+\Omega\partial/\partial\theta]^2 w - h[(1/r)(\partial/\partial r)(r\sigma_r\partial w/\partial r) + (1/r)^2(\partial/\partial\theta)(\sigma_\theta\partial w/\partial\theta)]=0, \quad (1)$$

where $D=Eh^3/12(1-\nu^2)$; E is the Young's modulus, ν the Poisson's ratio, and h the plate thickness. The plate is fully clamped at the hub radius $r=b$ while free along the outer edge $r=a$. The inplane boundary conditions in this case are

$$v(b,\theta,t)=\sigma_r(a,\theta,t)=0, \quad (2)$$

where v is the in-plane radial displacement. The out-of-plane boundary conditions are the usual clamped-free boundary conditions (Iwan and Stahl, 1973). The in-plane stresses σ_r and σ_θ in a rotating annular plate satisfying equation (2) are found in the paper by Srinivasan and Ramamurti (1980).

One assumes the solution in the form

$$w(r,\theta,t)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}[C_{mn}(t)\cos(n\theta)+S_{mn}(t)\sin(n\theta)]R_{mn}(r). \quad (3)$$

In the present analysis the mode function of a nonrotating plate is introduced as R_{mn}, i.e.,

$$R_{mn}(r)=J_n(k_{mn}r/a)+A_{mn}Y_n(k_{mn}r/a)+B_{mn}I_n(k_{mn}r/a)+C_{mn}K_n(k_{mn}r/a), \quad (4)$$

where J_n, Y_n, I_n, and K_n are the Bessel and the modified Bessel functions of order n; A_{mn} through C_{mn} and k_{mn} are determined from the clamped-free boundary conditions.

The frequency equation is derived from these equations using a Galerkin formulation (Iwan and Moeller, 1976), separately for each value of n.

Numerical Results and Discussion

Numerical results that follow were obtained by taking the series terms up to $m=5$. It is well known that some of the natural frequencies decrease and finally come to zero as the rotation speed of the plate increases. Physically this means that a resonance appears in the system when the plate is subjected to a stationary (zero frequency) lateral load. The speed of rotation which brings about this type of instability is generally referred to as the critical speed.

Figure 1 shows the nondimensional critical speeds $(\Omega_0)_{cr}=[\rho ha^4/D]^{1/2}(\Omega)_{cr}$ as functions of the radii ratio of plate b/a. In the figure the symbol (m,n), m, $n=0, 1, \ldots$, denotes the vibration mode with m nodal circles and n nodal diameters. It is noted that (0,0) and (0,1) do not have a critical speed. The vibration mode which brings the lowest critical speed is transferred from (0,2) mode to the higher modes as b/a increases. The dashed line is the envelope of the curves which gives an approximate value of the lowest critical speed. Applying the least square method to this curve and further introducing the speed n_{cr} in rpm, the lowest critical speed is approximated as (with $\nu=0.33$)

$$n_{cr}(\text{rpm})=[Eh^2/\rho a^4]^{1/2}[42.3(b/a)^2+2.4(b/a)+11.0]. \quad (5)$$

Here, $E=$ Young's modulus of plate (N/m^2), $\rho=$ mass density (kg/m^3), $h=$ plate thickness (m), $a=$ outer radius of plate (m), $b=$ hub radius (m).

Incidentally, the critical speed experimentally obtained for a sawblade of $E=1.93\times10^{11}$ N/m^2, $\rho=7.87\times10^3$ kg/m^3, $h=1.65\times10^{-3}$m, $a=3.05\times10^{-1}$m, and $b=4.3\times10^{-2}$m was $n_{cr}=1100$ rpm. The approximate value from equation (5) is

[1]Associate Professor, Department of Mechanical Engineering, Tohoku University, Sendai, Japan 980.

Manuscript received by ASME Applied Mechanics Division, February 25, 1987; final revision May 8, 1987.

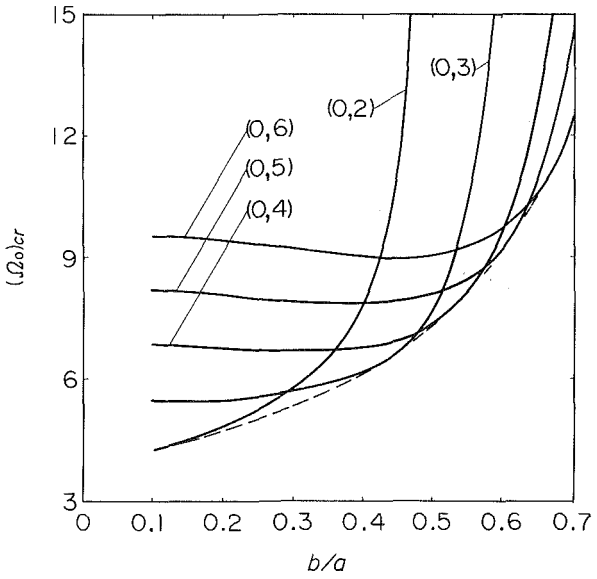


Fig. 1 Critical speeds versus inner-to-outer radius ratio; $\nu = 0.33$

$n_{cr} = 1070$ rpm. Further, for a floppy disk of $E = 4.9 \times 10^9$ N/m^2, $\rho = 1.3 \times 10^3$ kg/m^3, $h = 7.8 \times 10^{-5}$m, $a = 1.0 \times 10^{-1}$m, and $b = 2.5 \times 10^{-2}$m, the theoretical value of n_{cr} based on FEM was 230 rpm (Ono et al., 1986), while it is 216 rpm from equation (5). These examples show the effectiveness of equation (5) in predicting the critical speed of rotating clamped-free plates.

References

Benson, R. C., and Bogy, D. B., 1978, "Deflection of a Very Flexible Spinning Disk Due to a Stationary Transverse Load," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 636–642.

Eversman, W., and Dodson, Jr., R. O., 1969, "Free Vibration of a Centrally Clamped Spinning Circular Disk," *AIAA Journal*, Vol. 7, pp. 2010–2012.

Iwan, W. D., and Moeller, T. L., 1976, "The Stability of a Spinning Elastic Disk With a Transverse Load System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, pp. 485–490.

Iwan, W. D., and Stahl, K. J., 1973, "The Response of an Elastic Disk With a Moving Mass System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, pp. 445–451.

Mote, Jr., C. D., 1965, "Free Vibration of Initially Stressed Circular Disks," *ASME Journal of Engineering for Industry*, Vol. 87, pp. 258–264.

Mote, Jr., C. D., 1967, "Natural Frequencies in Annuli With Induced Thermal Membrane Stresses," *ASME Journal of Engineering for Industry*, Vol. 89, pp. 611–618.

Ono, K., Maeno, T., and Ebihara, T., 1986, "A Study of Head to Media Interface in Flexible Disks (2nd Reports, Dynamic Response Characteristics Due to a Concentrated Force)," in Japanese, *Transactions of the Japan Society of Mechanical Engineers*, Vol. 52, pp. 326–335.

Srinivasan, V., and Ramamurti, V., 1980, "Stability,and Vibration of an Annular Plate With Concentrated Edge Load," *Computer and Structures*, Vol. 12, pp. 119–129.

An Approximate Method in Plane-Stress Small Strain Plasticity

D. Durban[2,3]

Introduction

The mathematical formulation of a axially-symmetric plane stress elastoplastic problems, even when restricted to small strains and simplified by the use of the deformation theory, implies a nonlinear differential equation that can be solved, in general, only by direct numerical methods. Attempts to produce closed-form analytical solutions for stress concentration problems proved successful only for particular uniaxial stress-strain curves. Budiansky (1971) has obtained elegant expressions for the stress concentration factor at a circular hole embedded in an infinite sheet under remote uniform tension. That analysis is for Hill's (1948) orthotropic material in conjunction with a pure power law in the post yield range. Further results for the hole problem are presented in Budiansky (1984) for a whole class of anisotropic plastic solids along with a pure power law characteristic. Similar studies for the stress concentration around rigid fastening rings were given by Yang (1969) and Durban (1987a) for the power hardening law, and in Durban (1987b) for elastic/linear-hardening solids with plastic orthotropy.

In this note we suggest a new approximate method aimed at a quick, yet reliable, assessment of stress concentration factors. Material behavior is modeled by the small strain J_2 deformation theory with arbitrary hardening characteristics. We show that the governing equation for the stress components admits a solution by quadratures when the operative part of the yield locus is approximated by a suitably chosen straight line. Comparison with numerical results for the stress concentration factor at the circular hole, obtained from the exact stress equation, indicates a very good agreement over the entire elastoplastic range. The applicability of the method for other problems is demonstrated by evaluating the stress concentration factor at a rigid fastening ring.

Extension of the proposed approximation technique to orthotropic materials based on Hill's (1948) theory is straightforward.

The Method as Illustrated by the Hole Problem

Using the standard axially-symmetric plane-stress notation, we have the elastoplastic constitutive relations, of the small strain deformation theory, in the form

$$\epsilon_r = \frac{1}{E}(\sigma_r - \nu\sigma_\theta) + \left(\frac{1}{E_s} - \frac{1}{E}\right)\left(\sigma_r - \frac{1}{2}\sigma_\theta\right) \tag{1}$$

$$\epsilon_\theta = \frac{1}{E}(\sigma_\theta - \nu\sigma_r) + \left(\frac{1}{E_s} - \frac{1}{E}\right)\left(\sigma_\theta - \frac{1}{2}\sigma_r\right) \tag{2}$$

where $(\epsilon_r, \epsilon_\theta)$ are the strains, $(\sigma_r, \sigma_\theta)$ the stresses, (ν, E) the elastic constants, and E_s is the secant modulus and a known function of the effective stress

$$\sigma_e^2 = \sigma_r^2 + \sigma_\theta^2 - \sigma_r\sigma_\theta \tag{3}$$

Following the common practice of axially-symmetric plane-stress problems we combine the equilibrium equation with the compatibility condition to obtain the differential identity

$$(\sigma_r - \sigma_\theta)d\epsilon_\theta + (\epsilon_r - \epsilon_\theta)d\sigma_r = 0 \tag{4}$$

Substituting now the constitutive relations (1) and (2) in equation (4) gives the differential relation

$$d\sigma_r + d\sigma_\theta + \left(\sigma_\theta - \frac{1}{2}\sigma_r\right)\left(\frac{E_s}{E_T} - 1\right)\frac{d\sigma_e}{\sigma_e} = 0 \tag{5}$$

where E_T is the tangent modulus and a known function of the effective stress σ_e.

At this stage we introduce an approximation whereby the operative part of the yield locus (3) is replaced by the corresponding straight line. Taking the hole problem, as an example, we have at infinity $\sigma_r = \theta_\theta = \sigma_\infty$, where σ_∞ is the applied uniform tension, while at the hole $\sigma_r = 0$. The operative part is therefore the *curved line AB* (Fig. 1) and the appropriate approximation of equation (3) is given here by the *straight line*

[2]Department of Engineering, University of Cambridge, Cambridge, England CB2 1PZ.

[3]On sabbatical leave until October, 1987. Permanent address: Department of Aeronautical Engineering, Technion, Haifa 32000, Israel.

Manuscript received by ASME Applied Mechanics Division, April 16, 1987.

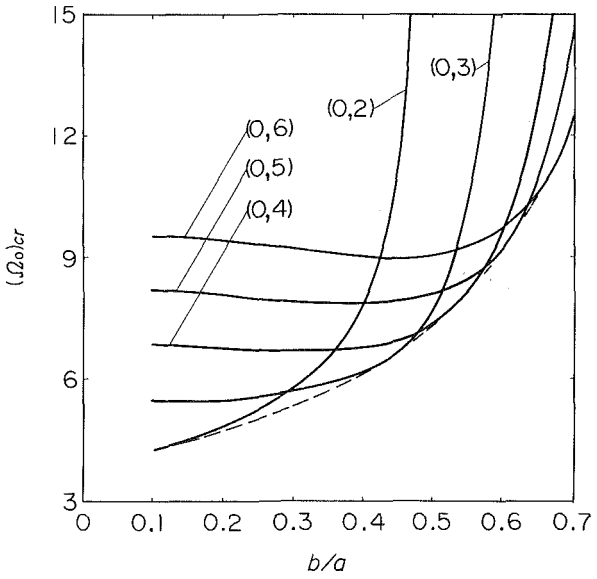


Fig. 1 Critical speeds versus inner-to-outer radius ratio; $\nu = 0.33$

$n_{cr} = 1070$ rpm. Further, for a floppy disk of $E = 4.9 \times 10^9$ N/m^2, $\rho = 1.3 \times 10^3$ kg/m^3, $h = 7.8 \times 10^{-5}$m, $a = 1.0 \times 10^{-1}$m, and $b = 2.5 \times 10^{-2}$m, the theoretical value of n_{cr} based on FEM was 230 rpm (Ono et al., 1986), while it is 216 rpm from equation (5). These examples show the effectiveness of equation (5) in predicting the critical speed of rotating clamped-free plates.

References

Benson, R. C., and Bogy, D. B., 1978, "Deflection of a Very Flexible Spinning Disk Due to a Stationary Transverse Load," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 636–642.

Eversman, W., and Dodson, Jr., R. O., 1969, "Free Vibration of a Centrally Clamped Spinning Circular Disk," *AIAA Journal*, Vol. 7, pp. 2010–2012.

Iwan, W. D., and Moeller, T. L., 1976, "The Stability of a Spinning Elastic Disk With a Transverse Load System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, pp. 485–490.

Iwan, W. D., and Stahl, K. J., 1973, "The Response of an Elastic Disk With a Moving Mass System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, pp. 445–451.

Mote, Jr., C. D., 1965, "Free Vibration of Initially Stressed Circular Disks," *ASME Journal of Engineering for Industry*, Vol. 87, pp. 258–264.

Mote, Jr., C. D., 1967, "Natural Frequencies in Annuli With Induced Thermal Membrane Stresses," *ASME Journal of Engineering for Industry*, Vol. 89, pp. 611–618.

Ono, K., Maeno, T., and Ebihara, T., 1986, "A Study of Head to Media Interface in Flexible Disks (2nd Reports, Dynamic Response Characteristics Due to a Concentrated Force)," in Japanese, *Transactions of the Japan Society of Mechanical Engineers*, Vol. 52, pp. 326–335.

Srinivasan, V., and Ramamurti, V., 1980, "Stability,and Vibration of an Annular Plate With Concentrated Edge Load," *Computer and Structures*, Vol. 12, pp. 119–129.

An Approximate Method in Plane-Stress Small Strain Plasticity

D. Durban[2,3]

Introduction

The mathematical formulation of a axially-symmetric plane stress elastoplastic problems, even when restricted to small strains and simplified by the use of the deformation theory, implies a nonlinear differential equation that can be solved, in general, only by direct numerical methods. Attempts to produce closed-form analytical solutions for stress concentration problems proved successful only for particular uniaxial stress-strain curves. Budiansky (1971) has obtained elegant expressions for the stress concentration factor at a circular hole embedded in an infinite sheet under remote uniform tension. That analysis is for Hill's (1948) orthotropic material in conjunction with a pure power law in the post yield range. Further results for the hole problem are presented in Budiansky (1984) for a whole class of anisotropic plastic solids along with a pure power law characteristic. Similar studies for the stress concentration around rigid fastening rings were given by Yang (1969) and Durban (1987a) for the power hardening law, and in Durban (1987b) for elastic/linear-hardening solids with plastic orthotropy.

In this note we suggest a new approximate method aimed at a quick, yet reliable, assessment of stress concentration factors. Material behavior is modeled by the small strain J_2 deformation theory with arbitrary hardening characteristics. We show that the governing equation for the stress components admits a solution by quadratures when the operative part of the yield locus is approximated by a suitably chosen straight line. Comparison with numerical results for the stress concentration factor at the circular hole, obtained from the exact stress equation, indicates a very good agreement over the entire elastoplastic range. The applicability of the method for other problems is demonstrated by evaluating the stress concentration factor at a rigid fastening ring.

Extension of the proposed approximation technique to orthotropic materials based on Hill's (1948) theory is straightforward.

The Method as Illustrated by the Hole Problem

Using the standard axially-symmetric plane-stress notation, we have the elastoplastic constitutive relations, of the small strain deformation theory, in the form

$$\epsilon_r = \frac{1}{E}(\sigma_r - \nu\sigma_\theta) + \left(\frac{1}{E_s} - \frac{1}{E}\right)\left(\sigma_r - \frac{1}{2}\sigma_\theta\right) \tag{1}$$

$$\epsilon_\theta = \frac{1}{E}(\sigma_\theta - \nu\sigma_r) + \left(\frac{1}{E_s} - \frac{1}{E}\right)\left(\sigma_\theta - \frac{1}{2}\sigma_r\right) \tag{2}$$

where $(\epsilon_r, \epsilon_\theta)$ are the strains, $(\sigma_r, \sigma_\theta)$ the stresses, (ν, E) the elastic constants, and E_s is the secant modulus and a known function of the effective stress

$$\sigma_e^2 = \sigma_r^2 + \sigma_\theta^2 - \sigma_r\sigma_\theta \tag{3}$$

Following the common practice of axially-symmetric plane-stress problems we combine the equilibrium equation with the compatibility condition to obtain the differential identity

$$(\sigma_r - \sigma_\theta)d\epsilon_\theta + (\epsilon_r - \epsilon_\theta)d\sigma_r = 0 \tag{4}$$

Substituting now the constitutive relations (1) and (2) in equation (4) gives the differential relation

$$d\sigma_r + d\sigma_\theta + \left(\sigma_\theta - \frac{1}{2}\sigma_r\right)\left(\frac{E_s}{E_T} - 1\right)\frac{d\sigma_e}{\sigma_e} = 0 \tag{5}$$

where E_T is the tangent modulus and a known function of the effective stress σ_e.

At this stage we introduce an approximation whereby the operative part of the yield locus (3) is replaced by the corresponding straight line. Taking the hole problem, as an example, we have at infinity $\sigma_r = \theta_\theta = \sigma_\infty$, where σ_∞ is the applied uniform tension, while at the hole $\sigma_r = 0$. The operative part is therefore the *curved line AB* (Fig. 1) and the appropriate approximation of equation (3) is given here by the *straight line*

[2]Department of Engineering, University of Cambridge, Cambridge, England CB2 1PZ.

[3]On sabbatical leave until October, 1987. Permanent address: Department of Aeronautical Engineering, Technion, Haifa 32000, Israel.

Manuscript received by ASME Applied Mechanics Division, April 16, 1987.

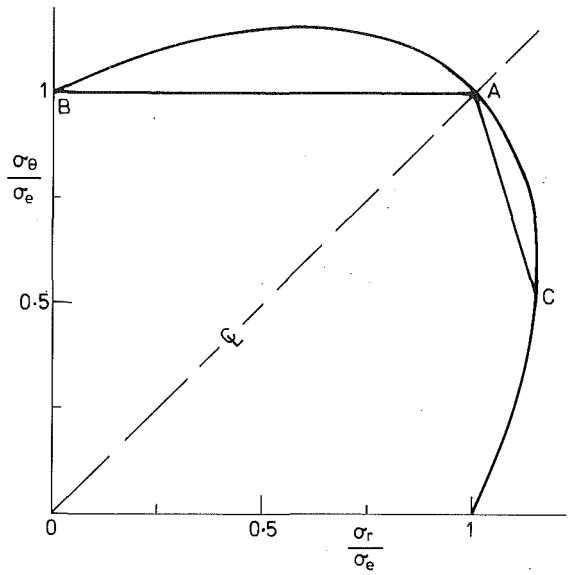


Fig. 1 The operative parts of the Mises yield locus are replaced by the appropriate straight lines approximations; *AB* for the hole problem and *AC* for the rigid ring problem

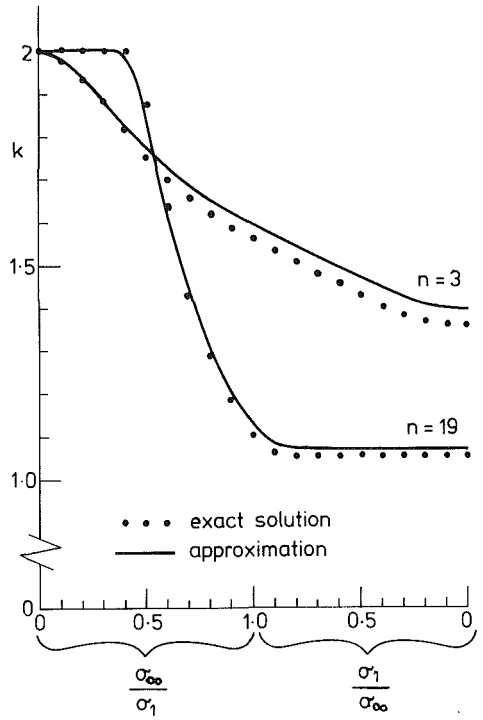


Fig. 2 The stress concentration factor at a circular hole for the Ramberg-Osgood hardening rule. The exact solution is taken from Budinansky and Mangasarian (1960).

AB. Thus, the exact definition of the effective stress (3) is now replaced by the approximation

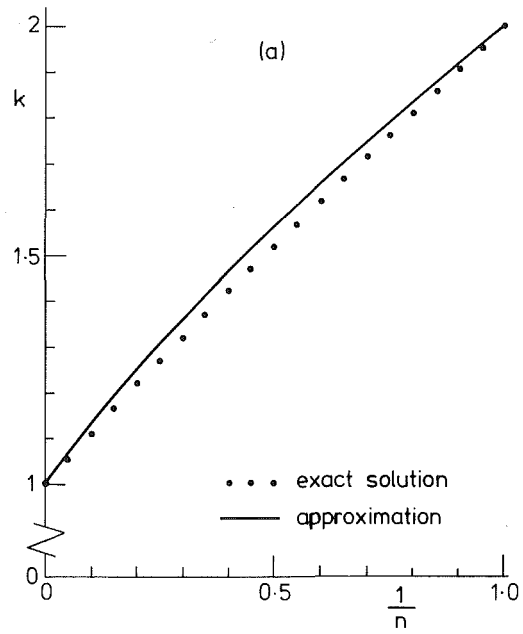
$$\sigma_\theta = \sigma_e \tag{6}$$

Inserting this approximation in equation (5) and taking σ_e as the independent variable we get the linear differential equation

$$\frac{d\sigma_r}{d\sigma_e} - \frac{1}{2}\left(\frac{E_s}{E_T} - 1\right)\frac{\sigma_r}{\sigma_e} + \frac{E_s}{E_T} = 0 \tag{7}$$

with the solution

$$\sigma_r = \left(\frac{E_{s\infty}}{E_s}\right)^{1/2}\sigma_\infty - E_s^{-1/2}\int_{\sigma_\infty}^{\sigma_e}\frac{E_s^{3/2}}{E_T}\,d\sigma_e \tag{8}$$



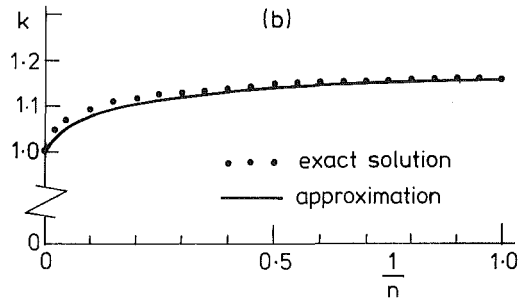


Fig. 3 Stress concentration factors for pure power hardening materials: (*a*) the hole problem; (*b*) the rigid ring problem

where $E_{s\infty}$ is the value of the secant modulus at infinity. Note that the effective stress at infinity is the same with equations (3) and (6) and is equal to σ_∞. Specifying (8) at the hole where $\sigma_r = 0$ and $\sigma_e = k\sigma_\infty$, k being the stress concentration factor, gives the transcendental equation

$$E_{s\infty}^{1/2}\,\sigma_\infty = \int_{\sigma_\infty}^{k\sigma_\infty}\frac{E_s^{3/2}}{E_T}\,d\sigma_e \tag{9}$$

which is easily solved, with any specified hardening characteristic, to obtain the dependence of the stress concentration factor on the applied load.

A numerical solution of equation (5) in conjunction with equation (3) has been given by Budinansky and Mangasarian (1960) for the Ramberg-Osgood hardening rule

$$\epsilon = \frac{\sigma}{E}\left[1 + \frac{3}{7}\left(\frac{\sigma}{\sigma_1}\right)^{n-1}\right] \tag{10}$$

where ϵ is the total uniaxial strain and (σ_1, n) are material constants. To find the corresponding approximate solution we insert equation (10) in (9) and solve the resulting equation for k

$$\left(1 + \frac{3}{7}\lambda^{n-1}\right)^{-1/2} = \int_1^k \left(1 + \frac{3}{7}n\lambda^{n-1}\Sigma^{n-1}\right)\left(1 + \frac{3}{7}\lambda^{n-1}\right)^{-3/2} d\Sigma \tag{11}$$

where $\lambda = \sigma_\infty / \sigma_1$ is the nondimensionalized load and $\Sigma = \sigma_e / \sigma_\infty$. Comparison of the results for k from equation (11) with the numerical solution of the exact equation shows (Fig. 2) a very good agreement, with an error of just a few percent.

Consider next the pure power hardening relation $\epsilon = (\sigma/\sigma_1)^n$. The solution of equation (9) is then simply

$$k = \left(\frac{2n}{n+3}\right)^{\frac{2}{n-3}} \tag{12}$$

The exact result for that case (Budiansky and Mangasarian, 1960) reads

$$k = \left[\left(\frac{n+3}{2n}\right)^{\frac{n+3}{n^2+3}}\right] \exp\left[\frac{\pi(n-1)}{\sqrt{3}(n^2+3)}\right] \tag{13}$$

Comparison of equation (12) with equation (13) shows again (Fig. 3(*a*)) an agreement, to within a few percent, over the entire range of n.

Stress Concentration at a Rigid Ring

The conditions at infinity are the same as in the hole problem with $\sigma_r = \theta_\theta = \sigma_\infty$. At the ring we have that $\epsilon_\theta = 0$ or, by equation (2), that $\sigma_\theta = \nu_s \sigma_r$ where ν_s is the secant Poisson ratio

$$\nu_s = \frac{1}{2} - \left(\frac{1}{2} - \nu\right)\frac{E_s}{E} \tag{14}$$

It follows from equation (3) that the operative part of the yield locus (Fig. 1) extends from A to point C whose coordinates are

$$\frac{\sigma_r}{\sigma_e} = \zeta_a^{-1} \qquad \frac{\sigma_\theta}{\sigma_e} = \nu_{sa}\zeta_a^{-1} \qquad \zeta_a = \sqrt{1 - \nu_{sa} + \nu_{sa}^2} \tag{15}$$

where ν_{sa} is the value of ν_s at the ring. The corresponding straight line approximation AC is here given by

$$(1 - \zeta_a)\sigma_\theta = (\nu_{sa} - \zeta_a)\sigma_r + (1 - \nu_{sa})\sigma_e \tag{16}$$

Note, that unlike the hole problem, the location of point C is here dependent on the external load σ_∞. Combining equation (16) with equation (5), we get again a linear differential equation

$$(2\zeta_a - \nu_{sa} - 1)\frac{d\sigma_r}{d\sigma_e} + \frac{1}{2}(\zeta_a - 2\nu_{sa} + 1)$$
$$\left(\frac{E_s}{E_T} - 1\right)\frac{\sigma_r}{\sigma_e} - (1 - \nu_{sa})\frac{E_s}{E_T} = 0 \tag{17}$$

that admits a solution by quadratures. The stress concentration factor is here defined as the ratio of the effective stress at the hole to the applied load at infinity. Put differently, we have at the hole

$$\sigma_r = \zeta_a^{-1} k \sigma_\infty \tag{18}$$

The resulting transcendental equation for k is similar to equation (9) but somewhat more complicated due to the coefficients of equation (17). On balance, however, it is much simpler to determine k by the present approximate method in comparison with the full numerical solution.

For the pure power law material with $\nu_s \equiv 1/2$, $\zeta_a = \sqrt{3}/2$, $E_s/E_T = n$, we find

$$k = \left(\frac{n+3}{2\sqrt{3}}\right)^{\frac{2}{(2+\sqrt{3})n - \sqrt{3}}} \tag{19}$$

while the exact result (Yang, 1969; Durban, 1987a) reads

$$k = \left[\left(\frac{n+3}{2\sqrt{3}}\right)^{\frac{n+3}{n^2+3}}\right] \exp\left[-\frac{\pi(n-1)}{\sqrt{3}(n^2+3)}\right] \tag{20}$$

Comparison of equation (19) with equation (20) reveals (Fig. 3(*b*)) a very close agreement.

Extension to Orthotropic Materials

There is no difficulty in applying the same method of approximation to Hill's (1948) family of anisotropic materials where definition (3) is generalized by

$$\sigma_e^2 = \sigma_r^2 + \sigma_\theta^2 - \frac{2R}{1+R}\sigma_r\sigma_\theta \tag{21}$$

R being the measure of transverse orthotropy.

Taking again the hole problem as an example, with a pure power law characteristic, we find the stress concentration factor

$$k = 2Q\left[\frac{nQ}{(n-1)Q^2 + 1}\right]^{\frac{1}{(n-1)Q-1}} \qquad Q = (2 + 2R)^{-1/2} \tag{22}$$

For the Mises material with $R = 1$ we recover from equation (22) relation (12).

A comparison of equation (22) with the exact solution (equation (25) in Budiansky, 1971) shows again an excellent agreement over the entire range of parameters R and n. The two results become identical for $n = 1$ ($k = 2$) and for $n = \infty$ ($k = 2Q$).

Acknowledgment

Thanks are due to the kind assistance and hospitality of the Engineering Department, Cambridge University, England.

References

Budiansky, B., 1971, "An Exact Solution to an Elastic-Plastic Stress Concentration Problem," *PMM*, Vol. 35, pp. 40–48.

Budiansky, B., 1984, "Anisotropic Plasticity of Plane-Isotropic Sheets," *Mechanics of Material Behaviour*, Elsevier Science Publishers, pp. 15–29.

Budiansky, B., and Mangasarian, O. L., 1960, "Plastic Stress Concentration at a Circular Hole in an Infinite Sheet Subjected to Equal Biaxial Tension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 82, pp. 59–64.

Durban, D., 1987a, "Plastic Stresses Induced by a Rigid Ring Embedded in a Thin Anisotropic Plate Under Uniform Tension," *Quart. Appl. Math.*, Vol. XLV, pp. 341–347.

Durban, D., 1987b, "On Two Stress Concentration Problems in Plane Stress Anisotropic Plasticity," *Int. J. Solids Struct.*, Vol. 23, pp. 469–484.

Hill, R., 1948, "A Theory of the Yielding and Plastic Flow of Anisotropic Metals," *Proc. Roy. Soc. London*, Vol. A193, pp. 281–297.

Yang, W. H., 1969, "Axisymmetric Plane Stress Problems in Anisotropic Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 7–14.

The Apparent Fracture Energy for Dynamic Crack Growth With Fine Scale Periodic Fraction Resistance

L. B. Freund[4]

Introduction

In the analysis of dynamic fracture phenomena or in the interpretation of experiments, it is commonly assumed that the crack tip speed varies in phase with the intensity of loading on the crack tip region. In other words, the speed increases or decreases as the load intensity increases or decreases, respectively. This is the case if the fracture resistance of the material appears to be spatially uniform. Suppose that, on a finer scale of observation, the fracture resistance of the material is nonuniform, so that the crack speed will be influenced by this

[4]Professor, Division of Engineering, Brown University, Providence, RI 02912. Fellow ASME.

Manuscript recieved by ASME Applied Mechanics Division, January 22, 1987; final revision July 30, 1987.

where $\lambda = \sigma_\infty / \sigma_1$ is the nondimensionalized load and $\Sigma = \sigma_e / \sigma_\infty$. Comparison of the results for k from equation (11) with the numerical solution of the exact equation shows (Fig. 2) a very good agreement, with an error of just a few percent.

Consider next the pure power hardening relation $\epsilon = (\sigma/\sigma_1)^n$. The solution of equation (9) is then simply

$$k = \left(\frac{2n}{n+3}\right)^{\frac{2}{n-3}} \tag{12}$$

The exact result for that case (Budiansky and Mangasarian, 1960) reads

$$k = \left[\left(\frac{n+3}{2n}\right)^{\frac{n+3}{n^2+3}}\right] \exp\left[\frac{\pi(n-1)}{\sqrt{3}(n^2+3)}\right] \tag{13}$$

Comparison of equation (12) with equation (13) shows again (Fig. 3(*a*)) an agreement, to within a few percent, over the entire range of n.

Stress Concentration at a Rigid Ring

The conditions at infinity are the same as in the hole problem with $\sigma_r = \theta_\theta = \sigma_\infty$. At the ring we have that $\epsilon_\theta = 0$ or, by equation (2), that $\sigma_\theta = \nu_s \sigma_r$ where ν_s is the secant Poisson ratio

$$\nu_s = \frac{1}{2} - \left(\frac{1}{2} - \nu\right)\frac{E_s}{E} \tag{14}$$

It follows from equation (3) that the operative part of the yield locus (Fig. 1) extends from A to point C whose coordinates are

$$\frac{\sigma_r}{\sigma_e} = \zeta_a^{-1} \qquad \frac{\sigma_\theta}{\sigma_e} = \nu_{sa}\zeta_a^{-1} \qquad \zeta_a = \sqrt{1 - \nu_{sa} + \nu_{sa}^2} \tag{15}$$

where ν_{sa} is the value of ν_s at the ring. The corresponding straight line approximation AC is here given by

$$(1 - \zeta_a)\sigma_\theta = (\nu_{sa} - \zeta_a)\sigma_r + (1 - \nu_{sa})\sigma_e \tag{16}$$

Note, that unlike the hole problem, the location of point C is here dependent on the external load σ_∞. Combining equation (16) with equation (5), we get again a linear differential equation

$$(2\zeta_a - \nu_{sa} - 1)\frac{d\sigma_r}{d\sigma_e} + \frac{1}{2}(\zeta_a - 2\nu_{sa} + 1)$$

$$\left(\frac{E_s}{E_T} - 1\right)\frac{\sigma_r}{\sigma_e} - (1 - \nu_{sa})\frac{E_s}{E_T} = 0 \tag{17}$$

that admits a solution by quadratures. The stress concentration factor is here defined as the ratio of the effective stress at the hole to the applied load at infinity. Put differently, we have at the hole

$$\sigma_r = \zeta_a^{-1} k \sigma_\infty \tag{18}$$

The resulting transcendental equation for k is similar to equation (9) but somewhat more complicated due to the coefficients of equation (17). On balance, however, it is much simpler to determine k by the present approximate method in comparison with the full numerical solution.

For the pure power law material with $\nu_s \equiv 1/2$, $\zeta_a = \sqrt{3}/2$, $E_s/E_T = n$, we find

$$k = \left(\frac{n+3}{2\sqrt{3}}\right)^{\frac{2}{(2+\sqrt{3})n - \sqrt{3}}} \tag{19}$$

while the exact result (Yang, 1969; Durban, 1987a) reads

$$k = \left[\left(\frac{n+3}{2\sqrt{3}}\right)^{\frac{n+3}{n^2+3}}\right] \exp\left[-\frac{\pi(n-1)}{\sqrt{3}(n^2+3)}\right] \tag{20}$$

Comparison of equation (19) with equation (20) reveals (Fig. 3(*b*)) a very close agreement.

Extension to Orthotropic Materials

There is no difficulty in applying the same method of approximation to Hill's (1948) family of anisotropic materials where definition (3) is generalized by

$$\sigma_e^2 = \sigma_r^2 + \sigma_\theta^2 - \frac{2R}{1+R}\sigma_r\sigma_\theta \tag{21}$$

R being the measure of transverse orthotropy.

Taking again the hole problem as an example, with a pure power law characteristic, we find the stress concentration factor

$$k = 2Q\left[\frac{nQ}{(n-1)Q^2 + 1}\right]^{\frac{1}{(n-1)Q-1}} \qquad Q = (2 + 2R)^{-1/2} \tag{22}$$

For the Mises material with $R = 1$ we recover from equation (22) relation (12).

A comparison of equation (22) with the exact solution (equation (25) in Budiansky, 1971) shows again an excellent agreement over the entire range of parameters R and n. The two results become identical for $n = 1$ ($k = 2$) and for $n = \infty$ ($k = 2Q$).

Acknowledgment

Thanks are due to the kind assistance and hospitality of the Engineering Department, Cambridge University, England.

References

Budiansky, B., 1971, "An Exact Solution to an Elastic-Plastic Stress Concentration Problem," *PMM*, Vol. 35, pp. 40–48.

Budiansky, B., 1984, "Anisotropic Plasticity of Plane-Isotropic Sheets," *Mechanics of Material Behaviour*, Elsevier Science Publishers, pp. 15–29.

Budiansky, B., and Mangasarian, O. L., 1960, "Plastic Stress Concentration at a Circular Hole in an Infinite Sheet Subjected to Equal Biaxial Tension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 82, pp. 59–64.

Durban, D., 1987a, "Plastic Stresses Induced by a Rigid Ring Embedded in a Thin Anisotropic Plate Under Uniform Tension," *Quart. Appl. Math.*, Vol. XLV, pp. 341–347.

Durban, D., 1987b, "On Two Stress Concentration Problems in Plane Stress Anisotropic Plasticity," *Int. J. Solids Struct.*, Vol. 23, pp. 469–484.

Hill, R., 1948, "A Theory of the Yielding and Plastic Flow of Anisotropic Metals," *Proc. Roy. Soc. London*, Vol. A193, pp. 281–297.

Yang, W. H., 1969, "Axisymmetric Plane Stress Problems in Anisotropic Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 7–14.

The Apparent Fracture Energy for Dynamic Crack Growth With Fine Scale Periodic Fraction Resistance

L. B. Freund[4]

Introduction

In the analysis of dynamic fracture phenomena or in the interpretation of experiments, it is commonly assumed that the crack tip speed varies in phase with the intensity of loading on the crack tip region. In other words, the speed increases or decreases as the load intensity increases or decreases, respectively. This is the case if the fracture resistance of the material appears to be spatially uniform. Suppose that, on a finer scale of observation, the fracture resistance of the material is nonuniform, so that the crack speed will be influenced by this

[4]Professor, Division of Engineering, Brown University, Providence, RI 02912. Fellow ASME.

Manuscript recieved by ASME Applied Mechanics Division, January 22, 1987; final revision July 30, 1987.

local variation in resistance as well as by the remote loading. This appears to be the case in brittle fracture of polycrystalline materials or composite materials, or the cleavage of a single crystal with dispersed particles or cavity in the path of the crack front (Forwood and Forty, 1965; Forwood, 1968). Even in pure cleavage of a single crystal, the advancing crack experiences varying resistance due to periodicity of the lattice, although this may be pushing a continuum model beyond acceptable limits. In any case, the actual crack motion on the finer scale fluctuates, and the comparatively smooth motion on the macroscopic scale is termed the apparent motion in the present discussion. In this note, the following question is addressed: If a crack is perceived to be advancing at a uniform rate in a material, but it is known that the material has some structure on a smaller scale that leads to local variation of the fracture resistance, then what features of the physical process of fracture can be discerned at the macroscopic level? The question is considered on the basis of a simple model of crack advance through a material with fine scale periodic resistance, and some implications for the apparent features of the process at the macroscopic level are examined. The simple analysis is based on the premise that the size of the crack tip process zone is small compared to the scale of the variation of fracture resistance, so that a semi-infinite crack idealization may be employed.

Consider rapid growth of a planar crack through a nominally elastic isotropic material under plane strain conditions. The tensile opening mode of deformation, or mode I, is assumed. The elastic modulus and Poisson's ratio of the material are E and ν, respectively. Attention is focused on points close to the crack tip compared to the in-plane dimensions of the body in which the crack grows, that is, the case of a semi-infinite crack in an otherwise unbounded body is considered. At any instant, the mechanical field surrounding the crack tip is characterized by the stress intensity factor and the speed of the crack tip. The instantaneous crack length measured with respect to some arbitrary reference point is denoted by $\ell(t)$ and the instantaneous crack tip speed is $\dot{\ell}(t)$.

Irwin's generalization of the Griffith fracture energy concept is adopted, and it is assumed that the crack grows according to an overall energy balance criterion. Thus, if Γ is the energy per unit crack advance in the plane that must be supplied to sustain growth, then the equation of motion of the crack tip is

$$\frac{E}{1-\nu^2}\frac{\Gamma}{K_o^2}=g(\dot{\ell}) \tag{1}$$

where K_o is the instantaneous value that the stress intensity factor would have due to the applied loading if the crack tip were not moving and $g(\dot{\ell})$ is a universal function of crack tip speed given by Freund (1972). Although the mathematical expression for g is complicated, the behavior of the function is simple. For all practical purposes, this function may be approximated by

$$g(\dot{\ell})=1-\dot{\ell}/c_r \tag{2}$$

where c_r is the Rayleigh wave speed of the material. The condition (1) ensures that the rate of energy flow into the crack tip through the region in which the stress intensity factor field dominates the full mechanical field is equal to the specific fracture energy of the material. The specific fracture energy Γ may depend on crack tip position, speed, or other system parameters, in general.

The main purpose here is to examine a crack propagation situation in which the nature of the resistance Γ is fundamentally different for two different levels of observation of the same process. On a macroscopic scale of observation, the material appears to have spatially homogeneous fracture resistance Γ_o, a constant. When observed on a fine scale, however, the material actually has a periodic (not necessarily sinusoidal) resistance to crack growth, say

$$\Gamma=\Gamma_m\gamma(\ell) \tag{3}$$

where Γ_m is a constant and $\gamma(\ell)$ is a periodic function of crack tip position with spatial period λ, defined for all ℓ. Suppose $\gamma(\ell)$ has the properties that

$$\max_{0<\ell\le\lambda}\gamma(\ell)=1 \qquad \min_{0<\ell\le\lambda}\gamma(\ell)\ge 0. \tag{4}$$

Thus, the constant Γ_m is the maximum value of the periodic specific fracture energy Γ.

Analysis

The *actual crack speed* as observed at the finer level of observation is determined by examining the crack tip equation of motion (1) for the periodic specific fracture energy (3). Incorporating the approximation (2), the result is

$$\frac{\dot{\ell}}{c_r}=1-\frac{E}{1-\nu^2}\frac{\Gamma_m}{K_o^2}\gamma(\ell). \tag{5}$$

The applied stress intensity factor is assumed to have little variation over distances equal to many times λ or during times equal to many times t_λ, so it is taken to be a constant for simplicity. The quantity t_λ is the time required for the crack tip to travel a distance equal to one wave length λ, and its dependence on loading level is obtained by integration as

$$c_r t_\lambda=\int_0^\lambda \frac{d\ell}{1-B\gamma(\ell)} \tag{6}$$

where B is the dimensionless combination of parameters $E\Gamma_m/(1-\nu^2)K_o^2$ which has the range $0<B<1$. The extreme values of $B\to 1^-$ and $B\to 0^+$ correspond, respectively, to the applied stress intensity factor at a level just large enough to push the crack tip past the peaks in the fracture resistance and at a level many times greater than this minimum level.

Consider now the same process at a macroscopic level at which the crack tip appears to move along steadily at constant speed, say, v_o, under the action of the uniform applied stress intensity factor K_o. This crack speed must be the average speed in the periodic fine scale fracture resistance, that is,

$$v_o(B)=\frac{\lambda}{t_\lambda}=c_r\left\{\int_0^1\frac{d\xi}{1-B\gamma(\lambda\xi)}\right\}^{-1}. \tag{7}$$

If the crack tip equation of motion is applied at the *macroscopic level*, then

$$\frac{E}{1-\nu^2}\frac{\Gamma_o}{K_o^2}=1-\frac{v_o(B)}{c_r} \tag{8}$$

where Γ_o is the macroscopically uniform specific fracture energy. The parameters on the left side of (8) are conveniently expressed in terms of B, so that the ratio of the *apparent macroscopic fracture energy* to the maximum in the fine scale periodic variation is

$$\frac{\Gamma_o}{\Gamma_m}=\frac{1-v_o(B)/c_r}{B}=\frac{1-v_o/c_r}{B(v_o)}. \tag{9}$$

This is the main result for the apparent fracture energy in terms of the details of the periodic variation of the fine scale resistance and the applied load level. It is written in two ways to make clear that the ratio may be expressed in terms of the load parameter B or the average crack speed v_o, these two quantities being related through (7).

The limiting values of the ratio (9) for very slow and very fast crack growth are easily deduced for arbitrary $\gamma(\ell)$. Recall that for very slow crack growth, $B\to 1^-$. In view of the proper-

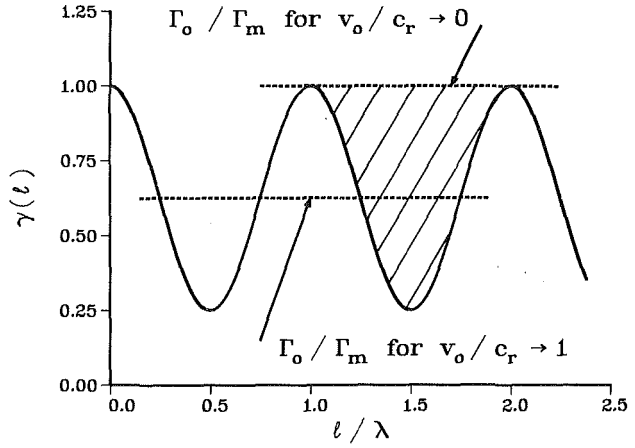


Fig. 1 A representative fine scale periodic fracture resistance, showing the limits of apparent macroscopic fracture resistance for very slow and very fast crack growth. The shaded area represents the radiated energy for macroscopically slow crack growth.

ty $(4)_1$ it is clear that $t_\lambda \to \infty$ as $B \to 1^-$. This, in turn, implies that $v_o(B) \to 0$ so that

$$\Gamma_o/\Gamma_m \to 1 \quad \text{as} \quad v_o/c_r \to 0 \quad \text{or} \quad B \to 1. \tag{10}$$

On the other hand, for severely over-driven crack growth, $B \ll 1$. Thus, to first order in small values of B,

$$c_r t_\lambda \approx \int_0^\lambda \{1 + B\gamma(\ell)\} d\ell = \lambda(1 + B\bar{\gamma}) \tag{11}$$

where $\bar{\gamma} < 1$ is the average value of $\gamma(\ell)$ within one wave length λ. The uniform macroscopic speed is then approximately $(1 - B\bar{\gamma})c_r$, so that

$$\Gamma_o/\Gamma_m \to \bar{\gamma} \quad \text{as} \quad v_o/c_r \to 1 \quad \text{or} \quad B \to 0. \tag{12}$$

Discussion

These extreme cases (10) and (12) may be interpreted as follows. Consider first the case when B is only slightly greater than unity or $v_o/c_r \ll 1$. Suppose that the variation in fracture energy has the qualitative features shown in Fig. 1. Thus, the crack tip driving force is barely large enough to push the crack tip past the maxima in the fracture energy and, according to the equation of motion, the crack tip moves slowly as it passes a maximum in fracture energy. Between the maxima, the crack tip rapidly accelerates to a relatively high speed and decelerates as it approaches the next maximum in $\gamma(\ell)$. Of the total time required for the crack tip to traverse a wave length λ in fracture energy variation, the crack tip spends most of its time in regions where the fracture energy is Γ_m. If this result is now *interpreted* in terms of a uniform crack tip speed v_o, then the *apparent* time rate of energy flux into the crack tip is $\Gamma_m v_o$ for most of the time. This is the conclusion reached in (10). Of course, the actual energy absorbed per wave length λ in the fracture process differs from the apparent energy absorbed by the amount represented as a shaded area in Fig. 1. This amount of energy, which is included in the macroscopic energy flow into the crack tip region, is radiated outward from the crack tip as it accelerates and decelerates. The spectrum of this radiation is dominated by wave lengths on the order of λ and it is *not* accounted for in the macroscopic crack tip energy flux.

The other extreme case, which is perhaps less interesting from a physical point of view, is when B has a value only slightly greater than zero. This is the situation when the driving force is far greater than the minimum required to push the crack tip past the maxima in the fracture resistance. In this case, the crack speed is always near the ideal upper limit c_r and there is little acceleration or deceleration due to the variation in the fracture energy. The crack motion is essentially uniform with the crack tip spending about equal time at all resistance levels within a wave length. Thus, the macroscopic interpretation of the process occurring with uniform crack speed v_o leads to the conclusion that the apparent fracture energy is the wave length average of fine scale periodic fracture energy, as in (12).

In general, the difference between the apparent rate of energy flow into the crack tip and the actual energy consumed in the fracture process is radiated as high frequency wave motion. The amount of energy radiated per unit crack advance is

$$E_R = \frac{1}{\lambda}\int_0^\lambda \{\Gamma_o - \Gamma_m \gamma(\ell)\} d\ell \tag{13}$$

which is the algebraic difference in area under graphs of the apparent fracture energy and the actual fracture energy versus distance.

Although the main results are evident in (10) and (12) for arbitrary periodic variation of the specific fracture energy, it is instructive to consider some special cases of $\gamma(\ell)$. For example, suppose that $\gamma(\ell) = \cos^2(\pi\ell/\lambda)$. This variation obviously satisfies the conditions in (4). The integral in (6) may be evaluated with the result that

$$\frac{\Gamma_o}{\Gamma_m} = \frac{1 - \sqrt{1 - B}}{B} = \frac{1}{1 + v_o/c_r}. \tag{14}$$

It is evident that this special case has the general properties in (10) and (12). A second instructive example is the piecewise constant function $\gamma(\ell) = 1$ for $0 < \ell < n\lambda$ and $\gamma(\ell) = q$ for $n\lambda < \ell < \lambda$, where $0 < n, q < 1$. Evaluation of the integral in (6) for the particular case with $n = q = \frac{1}{2}$ yields

$$\frac{\Gamma_o}{\Gamma_m} = \frac{3 - 2B}{4 - 3B} = \frac{6(v_o/c_r) + 2\sqrt{4 - 4(v_o/c_r) + 9(v_o/c_r)^2}}{9(v_o/c_r) - 2 + 3\sqrt{4 - 4(v_o/c_r) + 9(v_o/c_r)^2}}. \tag{15}$$

Again, the general features already discussed are evident in the result. A piecewise constant fracture resistance was used by Das and Aki (1977) in their study of strength inhomogenieties on a crustal fault in the earth during seismic slip.

It is noted that the general features of the analysis here would be the same if a stress intensity factor criterion for crack growth had been adopted, instead of the energy criterion. One issue raised by this simple calculation is whether or not the macroscopic fracture energy of a material perceived through brittle crack propagation experiments has a direct relationship with the true "surface energy" of the material. If the fracture energy of the material is inhomogeneous on a fine scale, then the answer would appear to be negative. Instead, the macroscopically perceived fracture energy represents a maximum, rather than average, of the fine scale fracture resistance. Finally, it is noted that this effect is quite different from the "lattice trapping" effect introduced by Thomson et al. (1971) and discussed by Lawn and Wilshaw (1975).

Acknowledgments

It is a pleasure to acknowledge fruitful discussion with R. J. Clifton on the subject matter of this note. Ideas similar to those represented here were discussed by J. R. Rice (1978) in an unpublished article which he kindly made available to me. The research support of the National Science Foundation, Solid Mechanics Program and the NSF Materials Research Laboratory at Brown University is gratefully acknowledged.

References

Das, S., and Aki, K., 1977, "Fault Plane with Barriers: A Versatile Earthquake Model," *Journal of Geophysical Research*, Vol. 82, pp. 5658–5670.

Forwood, C. T., and Forty, A. J., 1965, "The Interaction of Cleaveage Cracks with Inhomogenieties in Sodium Chloride Crystals," *Philosophical Magazine*, Vol. 11, pp. 1067–1082.

Forwood, C. T., 1968, "The Work of Fracture in Crystals of Sodium Chloride Containing Cavities," *Philosophical Magazine*, Vol. 17, pp. 657–667.

Freund, L B., 1972, "Crack Propagation in an Elastic Solid Subjected to General Loading – II. Non-Uniform Rate of Extension," *Journal of the Mechanics and Physics of Solids*, Vol. 20, pp. 141–152.

Lawn, B. R., and Wilshaw, T. R., 1975, *Fracture of Brittle Solids*, Cambridge University Press, pp. 144–151.

Rice, J. R., 1978, "Radiant Energy Dissipation in Crack Propagation," in *Proceedings of the 1977 ARPA Materials Research Council Meeting*, University of Michigan, pp. 60–77.

Thomson, R., Hsieh, C., and Rana, V., 1971, "Lattice Trapping of Fracture Cracks," *Journal of Applied Physics*, Vol. 42, p. 3154.

Eigenvalues for Interface Cracks in Linear Elasticity

M. F. Symington[5]

Two useful techniques for obtaining the elastic fields near the tip of a crack in a homogeneous medium are the eigenfunction method used by Williams (1957) and the analytic function method introduced by Muskhelishvili (1933). The first of these yields a solution in the form of a power series in r, the radial distance from the crack tip. The latter method gives a closed form solution which, when expanded in a series, is identical to the first solution.

Both of these methods can be used to find the linear elastic fields near the tip of a crack on the interface between two dissimilar media. However, the near-tip fields that have been derived for this problem (Williams, 1959; Erdogan, 1963; Sih and Rice, 1964) are not complete. The solutions properly characterize the singular ($O(r^{-1/2})$) behavior of the stresses at the crack tip but, unlike the homogeneous case, the series have no terms of integer order. The second order term ($O(1)$) not being included, the solution is asymptotically correct only to the first term.

Consider the stress state in a cracked composite medium due to two uniform stresses applied parallel to the crack (one above the bond line and one below it). Assuming the stresses are related by the jump condition required to maintain compatibility along the bond (Rice and Sih, 1965), a constant stress state, $O(1)$ throughout, would result. The homogeneous analog of that problem was examined by Williams (1957). The bimaterial solutions mentioned above cannot be matched to such a stress state. Using an analytic function formulation, Rice (1987) has recently given a complete series solution that includes integer order terms. It will be shown in this note that the complete series can also be obtained using Williams' method (Williams, 1959). The problem can be posed as follows.

Consider a semi-infinite crack along the interface of two bonded dissimilar media subject to arbitrary remote load. Let the medium be oriented in a Cartesian coordinate system so that the crack lies on the x axis. Position a polar coordinate system at the crack tip with the bond line along $\theta = 0$ so that material 1 occupies the upper half plane ($0 < \theta \leq \pi$) and material 2 occupies the lower half plane ($-\pi \leq \theta < 0$). Poisson's ratio and the shear modulus for each material are ν_i and μ_i, $i = 1, 2$.

Two stress functions, one for each medium, are required in order to pose this bimaterial problem. Allowing the eigenvalues λ_c to be complex, the stress functions given by Williams (1959) are

$$U_i = \Re e\left\{r^{\lambda_c+1} F_i(\theta; \lambda_c)\right\}, \quad i=1,2 \tag{1}$$

where

$$F_i(\theta; \lambda_c) = [a_i \sin(\lambda_c+1)\theta + b_i \cos(\lambda_c+1)\theta + c_i \sin(\lambda_c-1)\theta + d_i \cos(\lambda_c-1)\theta]. \tag{2}$$

The boundary conditions—traction free crack surfaces and continuity of traction and displacement across the bond line—provide eight equations in eight unknowns (a_i, b_i, c_i, and d_i, $i = 1, 2$) (Williams, 1959). Because the equations are homogeneous, a solution exists only for λ_c such that the determinant of the matrix of coefficients for the system of linear equations vanishes. Evaluating this determinant, the characteristic equation is

$$(\lambda_c+1) P \sin^2 \lambda_c \pi = 0 \tag{3}$$

where

$$P = [\mu_1(\kappa_2+1) + \mu_2(\kappa_1+1)]^2 - 4(\mu_1\kappa_2 + \mu_2)(\mu_2\kappa_1 + \mu_1)\sin^2\lambda_c\pi \tag{4}$$

and

$$\kappa_i = \begin{cases} 3-4\nu_i & \text{for plane strain} \\ \dfrac{3-\nu_i}{1+\nu_i} & \text{for plane stress.} \end{cases} \tag{5}$$

The set of all eigenvalues is the set of all λ_c that satisfy the characteristic equation. Therefore, consider possible solutions to equation (3):

(*a*) $(\lambda_c + 1) = 0$ does not give an admissible eigenvalue because bounded displacements at the crack tip and bounded strain energy in finite regions about the crack tip require $\Re e\{\lambda_c\} > 0$.

(*b*) The condition $P = 0$ is equivalent to the characteristic equation given by Williams (1959), for which the eigenvalues are $\lambda_c = \lambda + i\epsilon$ such that

$$\lambda = n + 1/2, \quad n = 0,1,2, \ldots \tag{6}$$

$$\epsilon = \frac{1}{2\pi} \ln\left[\frac{\mu_2\kappa_1 + \mu_1}{\mu_1\kappa_2 + \mu_2}\right]. \tag{7}$$

(*c*) The last possibility, $\sin \lambda_c \pi = 0$, provides the purely real integer eigenvalues

$$\lambda_c = n+1, \quad n = 0,1,2, \ldots .$$

Thus the complete series solution includes eigenfunctions given by two different sets of coefficients. The coefficients for the eigenfunctions associated with the complex eigenvalues (6, 7) have been given explicitly by Sih and Rice (1964). They can be conveniently written as (Symington, 1987)

$$\begin{aligned} a_1 &= i\frac{e^{\pi\epsilon} + \lambda_c e^{-\pi\epsilon}}{\lambda_c+1}\eta & a_2 &= i\frac{e^{-\pi\epsilon} + \lambda_c e^{\pi\epsilon}}{\lambda_c+1}\eta \\ b_1 &= \frac{e^{\pi\epsilon} - \lambda_c e^{-\pi\epsilon}}{\lambda_c+1}\eta & b_2 &= \frac{e^{-\pi\epsilon} - \lambda_c e^{\pi\epsilon}}{\lambda_c+1}\eta \\ c_1 &= -ie^{-\pi\epsilon}\eta & c_2 &= -ie^{\pi\epsilon}\eta \\ d_1 &= e^{-\pi\epsilon}\eta & d_2 &= e^{\pi\epsilon}\eta \end{aligned} \tag{8}$$

where η is a different undetermined complex constant for each complex eigenvalue. The coefficients associated with the integer eigenvalues are

[5]Graduate Student, Division of Engineering, Brown University, Providence, RI 02912.

Manuscript received by ASME Applied Mechanics Division, June 15, 1987; final revision August 4, 1987.

References

Das, S., and Aki, K., 1977, "Fault Plane with Barriers: A Versatile Earthquake Model," *Journal of Geophysical Research*, Vol. 82, pp. 5658–5670.

Forwood, C. T., and Forty, A. J., 1965, "The Interaction of Cleaveage Cracks with Inhomogenieties in Sodium Chloride Crystals," *Philosophical Magazine*, Vol. 11, pp. 1067–1082.

Forwood, C. T., 1968, "The Work of Fracture in Crystals of Sodium Chloride Containing Cavities," *Philosophical Magazine*, Vol. 17, pp. 657–667.

Freund, L B., 1972, "Crack Propagation in an Elastic Solid Subjected to General Loading – II. Non-Uniform Rate of Extension," *Journal of the Mechanics and Physics of Solids*, Vol. 20, pp. 141–152.

Lawn, B. R., and Wilshaw, T. R., 1975, *Fracture of Brittle Solids*, Cambridge University Press, pp. 144–151.

Rice, J. R., 1978, "Radiant Energy Dissipation in Crack Propagation," in *Proceedings of the 1977 ARPA Materials Research Council Meeting*, University of Michigan, pp. 60–77.

Thomson, R., Hsieh, C., and Rana, V., 1971, "Lattice Trapping of Fracture Cracks," *Journal of Applied Physics*, Vol. 42, p. 3154.

Eigenvalues for Interface Cracks in Linear Elasticity

M. F. Symington[5]

Two useful techniques for obtaining the elastic fields near the tip of a crack in a homogeneous medium are the eigenfunction method used by Williams (1957) and the analytic function method introduced by Muskhelishvili (1933). The first of these yields a solution in the form of a power series in r, the radial distance from the crack tip. The latter method gives a closed form solution which, when expanded in a series, is identical to the first solution.

Both of these methods can be used to find the linear elastic fields near the tip of a crack on the interface between two dissimilar media. However, the near-tip fields that have been derived for this problem (Williams, 1959; Erdogan, 1963; Sih and Rice, 1964) are not complete. The solutions properly characterize the singular ($O(r^{-1/2})$) behavior of the stresses at the crack tip but, unlike the homogeneous case, the series have no terms of integer order. The second order term ($O(1)$) not being included, the solution is asymptotically correct only to the first term.

Consider the stress state in a cracked composite medium due to two uniform stresses applied parallel to the crack (one above the bond line and one below it). Assuming the stresses are related by the jump condition required to maintain compatibility along the bond (Rice and Sih, 1965), a constant stress state, $O(1)$ throughout, would result. The homogeneous analog of that problem was examined by Williams (1957). The bimaterial solutions mentioned above cannot be matched to such a stress state. Using an analytic function formulation, Rice (1987) has recently given a complete series solution that includes integer order terms. It will be shown in this note that the complete series can also be obtained using Williams' method (Williams, 1959). The problem can be posed as follows.

Consider a semi-infinite crack along the interface of two bonded dissimilar media subject to arbitrary remote load. Let the medium be oriented in a Cartesian coordinate system so that the crack lies on the x axis. Position a polar coordinate system at the crack tip with the bond line along $\theta = 0$ so that material 1 occupies the upper half plane ($0 < \theta \leq \pi$) and material 2 occupies the lower half plane ($-\pi \leq \theta < 0$). Poisson's ratio and the shear modulus for each material are ν_i and μ_i, $i = 1, 2$.

Two stress functions, one for each medium, are required in order to pose this bimaterial problem. Allowing the eigenvalues λ_c to be complex, the stress functions given by Williams (1959) are

$$U_i = \Re e\left\{r^{\lambda_c+1} F_i(\theta; \lambda_c)\right\}, \quad i=1,2 \tag{1}$$

where

$$F_i(\theta; \lambda_c) = [a_i \sin(\lambda_c+1)\theta + b_i \cos(\lambda_c+1)\theta + c_i \sin(\lambda_c-1)\theta + d_i \cos(\lambda_c-1)\theta]. \tag{2}$$

The boundary conditions—traction free crack surfaces and continuity of traction and displacement across the bond line—provide eight equations in eight unknowns (a_i, b_i, c_i, and d_i, $i = 1, 2$) (Williams, 1959). Because the equations are homogeneous, a solution exists only for λ_c such that the determinant of the matrix of coefficients for the system of linear equations vanishes. Evaluating this determinant, the characteristic equation is

$$(\lambda_c+1)P \sin^2 \lambda_c \pi = 0 \tag{3}$$

where

$$P = [\mu_1(\kappa_2+1) + \mu_2(\kappa_1+1)]^2 - 4(\mu_1\kappa_2 + \mu_2)(\mu_2\kappa_1 + \mu_1)\sin^2 \lambda_c \pi \tag{4}$$

and

$$\kappa_i = \begin{cases} 3-4\nu_i & \text{for plane strain} \\ \dfrac{3-\nu_i}{1+\nu_i} & \text{for plane stress.} \end{cases} \tag{5}$$

The set of all eigenvalues is the set of all λ_c that satisfy the characteristic equation. Therefore, consider possible solutions to equation (3):

(*a*) $(\lambda_c + 1) = 0$ does not give an admissible eigenvalue because bounded displacements at the crack tip and bounded strain energy in finite regions about the crack tip require $\Re e\{\lambda_c\} > 0$.

(*b*) The condition $P = 0$ is equivalent to the characteristic equation given by Williams (1959), for which the eigenvalues are $\lambda_c = \lambda + i\epsilon$ such that

$$\lambda = n + 1/2, \quad n = 0,1,2, \ldots \tag{6}$$

$$\epsilon = \frac{1}{2\pi} \ln\left[\frac{\mu_2\kappa_1 + \mu_1}{\mu_1\kappa_2 + \mu_2}\right]. \tag{7}$$

(*c*) The last possibility, $\sin \lambda_c \pi = 0$, provides the purely real integer eigenvalues

$$\lambda_c = n+1, \quad n = 0,1,2, \ldots .$$

Thus the complete series solution includes eigenfunctions given by two different sets of coefficients. The coefficients for the eigenfunctions associated with the complex eigenvalues (6, 7) have been given explicitly by Sih and Rice (1964). They can be conveniently written as (Symington, 1987)

$$\begin{aligned} a_1 &= i\frac{e^{\pi\epsilon} + \lambda_c e^{-\pi\epsilon}}{\lambda_c+1}\eta & a_2 &= i\frac{e^{-\pi\epsilon} + \lambda_c e^{\pi\epsilon}}{\lambda_c+1}\eta \\ b_1 &= \frac{e^{\pi\epsilon} - \lambda_c e^{-\pi\epsilon}}{\lambda_c+1}\eta & b_2 &= \frac{e^{-\pi\epsilon} - \lambda_c e^{\pi\epsilon}}{\lambda_c+1}\eta \\ c_1 &= -ie^{-\pi\epsilon}\eta & c_2 &= -ie^{\pi\epsilon}\eta \\ d_1 &= e^{-\pi\epsilon}\eta & d_2 &= e^{\pi\epsilon}\eta \end{aligned} \tag{8}$$

where η is a different undetermined complex constant for each complex eigenvalue. The coefficients associated with the integer eigenvalues are

[5]Graduate Student, Division of Engineering, Brown University, Providence, RI 02912.

Manuscript received by ASME Applied Mechanics Division, June 15, 1987; final revision August 4, 1987.

$$a_i = -i\,\frac{\lambda_c - 1}{\lambda_c + 1}\,\alpha_i \xi$$
$$b_i = -\alpha_i \xi \qquad i = 1,2 \qquad (9)$$
$$c_i = i\alpha_i \xi$$
$$d_i = \alpha_i \xi$$

where ξ is a different undetermined complex constant for each integer eigenvalue and

$$\alpha_i = \frac{2\mu_i(\kappa_j + 1)}{\mu_i(\kappa_j + 1) + \mu_j(\kappa_i + 1)} \qquad (10)$$

with $j = 2$, 1 when $i = 1,2$. To make contact with the homogeneous problem, the above coefficients have been scaled so that they do not depend on material constants if the two media are the same.

The linear combination of all the eigenfunctions (1, 2) associated with the admissible complex and integer eigenvalues constitutes the complete series solution. This solution is identical to the one given by Rice (1987) when written in terms of two complex functions, $\Phi_i(z)$ and $\Lambda_i(z)$. Define these functions such that the stresses are given by

$$\sigma_{yy} + \sigma_{xx} = 4\Re\mathrm{e}\{\Phi_i(z)\}$$
$$\sigma_{yy} - \sigma_{xx} + 2i\sigma_{xy} = 2[(\bar{z} - z)\Phi_i'(z) - \Phi_i(z) + \Lambda_i(z)] \qquad (11)$$

then (Rice, 1987),

$$\Phi_i(z) = e^{-\pi\epsilon\,\mathrm{sgn}\theta} z^{-1/2 - i\epsilon} f(z) + \alpha_i g(z)$$
$$\Lambda_i(z) = e^{\pi\epsilon\,\mathrm{sgn}\theta} z^{-1/2 + i\epsilon} \bar{f}(z) - \alpha_i \bar{g}(z) \qquad i = 1,2 \qquad (12)$$

where $f(z)$ and $g(z)$ are power series with complex coefficients,

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \qquad g(z) = \sum_{n=0}^{\infty} b_n z^n. \qquad (13)$$

The undetermined constants of the complex stress functions, a_n and b_n, are related to the undetermined constants of the Airy stress function by

$$a_n = (n + 1/2 - i\epsilon)\bar{\eta}$$
$$b_n = (n+1)\xi \qquad n = 0,1,2\ldots \qquad (14)$$

where η and ξ are different complex constants for each value of n.

The inclusion of the integer order terms allows the constant stress state discussed earlier to be found exactly. It also allows the domain of validity of the series solution to be increased by matching integer order terms to the remote stress field.

Acknowledgments

I would like to acknowledge the helpful discussions I have had with Professor C. F. Shih. The research was supported by an ONR grant on Mechanics of Interface Cracks.

References

Erdogan, F., 1963, "Stress Distribution in a Nonhomogeneous Elastic Plane with Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 30, pp. 232-236.

Muskhelishvili, N. I., 1933, *Some Basic Problems of the Mathematical Theory of Elasticity*, English translation, P. Noordhoff Ltd., Groningen, Holland, 1953.

Rice, J. R., 1987, "Elastic Fracture Mechanics Concepts for Interfacial Cracks," Harvard University Technical Report MECH-101, May, to appear in ASME JOURNAL OF APPLIED MECHANICS.

Rice, J. R., and Sih, G. C., 1965, "Plane Problems of Cracks in Dissimilar Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 418-423.

Sih, G. C., and Rice, J. R., 1964, "The Bending of Plates of Dissimilar Materials with Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 477-482.

Symington, M. F., 1987, "A Review of Some Aspects of the Linear Elastic Analysis of Bimaterial Cracks," Brown University Master's Thesis, May.

Williams, M. L., 1957, "On the Stress Distribution at the Base of a Stationary Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 24, pp. 109-114.

Williams, M. L., 1959, "The Stresses Around a Fault or Crack in Dissimilar Media," *Bulletin of the Seismological Society of America*, Vol. 49, pp. 199-204.

Complex Singularities of the Transfer Function for Cylindrical Cavities in Elastic Media

A. N. Stokes[6] and A. F. Siggins[7]

1 Introduction

The response of a cylindrical cavity in an elastic medium to incoming disturbance has been much examined in recent years. A knowledge of the cavity transfer function is of importance in such areas as the seismic response of underground excavations, the response of geophones placed within boreholes, and various non-destructive testing applications. From some early work of Viktorov (1958), Fourier-Bessel series methods have been developed (Pao and Mow, 1973; Baron and Matthews, 1961), and "wave-sum" integral methods (Miklowitz, 1978) have been used to describe the transfer function.

The "wave-sum" method is useful at high frequencies giving a solution which can be conveniently divided into a Rayleigh wave component and various P and S waves. There are some difficulties in applying the method in the "illuminated zone." In this method an expression corresponding to a Fourier-Bessel series coefficient is evaluated by summing residues at singularities in a complex plane of spatial frequencies. Miklowitz (1978) discusses asymptotic properties which are associated with each type of wave.

The Fourier-Bessel series converge best for disturbances whose wavelength is rather large compared with the cavity diameter. In this case, for subsequent convolution of the transfer function with a specific disturbance, it is desirable to find singularities of the coefficients in the complex plane of temporal frequency. This has been done and some values for the first few coefficients are given by Pao and Mow (1973) and Baron and Matthews (1961). In the methods used there the numerical difficulty increases with increasing order of coefficients.

The singularities are studied by Pao and Mow (1973) because they are important in the long-term behavior of the solution. In this report a different method of computing them is described which does not seem to have that numerical problem, so that calculation of the poles can be extended into a region where asymptotic high frequency approximations apply.

2 Various Transfer Functions

A Fourier-Bessel series for the cavity hoop-stress response to an incoming steady state compressional plane wave is derived by Pao and Mow (1973) and is also described in Peck (1965). With a change of notation, its pth coefficient has a denominator, D_p, given by

$$D_p = xH_{p-1}(x)[(p^2-1)yH_{p-1}(y) - (p^3 - p + y^2/2)H_p(y)]$$
$$- H_p(x)[(p^3 - p + y^2/2)yH_{p-1}(y) - (p^2 + p - y^2/4)y^2 H_p(y)] \qquad (1)$$

[6]CSIRO Division of Mathematics and Statistics, Melbourne, Australia.
[7]CSIRO Division of Geomechanics, Mount Waverley, Australia.

Manuscript received by ASME Applied Mechanics Division, March 4, 1985; final revision June 22, 1986.

$$a_i = -i\,\frac{\lambda_c - 1}{\lambda_c + 1}\,\alpha_i \xi$$
$$b_i = -\alpha_i \xi \qquad i = 1,2 \qquad (9)$$
$$c_i = i\alpha_i \xi$$
$$d_i = \alpha_i \xi$$

where ξ is a different undetermined complex constant for each integer eigenvalue and

$$\alpha_i = \frac{2\mu_i(\kappa_j + 1)}{\mu_i(\kappa_j + 1) + \mu_j(\kappa_i + 1)} \qquad (10)$$

with $j = 2, 1$ when $i = 1,2$. To make contact with the homogeneous problem, the above coefficients have been scaled so that they do not depend on material constants if the two media are the same.

The linear combination of all the eigenfunctions (1, 2) associated with the admissible complex and integer eigenvalues constitutes the complete series solution. This solution is identical to the one given by Rice (1987) when written in terms of two complex functions, $\Phi_i(z)$ and $\Lambda_i(z)$. Define these functions such that the stresses are given by

$$\sigma_{yy} + \sigma_{xx} = 4\Re\mathrm{e}\{\Phi_i(z)\}$$
$$\sigma_{yy} - \sigma_{xx} + 2i\sigma_{xy} = 2[(\bar{z} - z)\Phi_i'(z) - \Phi_i(z) + \Lambda_i(z)] \qquad (11)$$

then (Rice, 1987),

$$\Phi_i(z) = e^{-\pi\epsilon\,\mathrm{sgn}\theta} z^{-1/2 - i\epsilon} f(z) + \alpha_i g(z)$$
$$\Lambda_i(z) = e^{\pi\epsilon\,\mathrm{sgn}\theta} z^{-1/2 + i\epsilon} \bar{f}(z) - \alpha_i \bar{g}(z) \qquad i = 1,2 \qquad (12)$$

where $f(z)$ and $g(z)$ are power series with complex coefficients,

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \qquad g(z) = \sum_{n=0}^{\infty} b_n z^n. \qquad (13)$$

The undetermined constants of the complex stress functions, a_n and b_n, are related to the undetermined constants of the Airy stress function by

$$a_n = (n + 1/2 - i\epsilon)\bar{\eta}$$
$$b_n = (n+1)\xi \qquad n = 0,1,2\ldots \qquad (14)$$

where η and ξ are different complex constants for each value of n.

The inclusion of the integer order terms allows the constant stress state discussed earlier to be found exactly. It also allows the domain of validity of the series solution to be increased by matching integer order terms to the remote stress field.

Acknowledgments

I would like to acknowledge the helpful discussions I have had with Professor C. F. Shih. The research was supported by an ONR grant on Mechanics of Interface Cracks.

References

Erdogan, F., 1963, "Stress Distribution in a Nonhomogeneous Elastic Plane with Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 30, pp. 232-236.

Muskhelishvili, N. I., 1933, *Some Basic Problems of the Mathematical Theory of Elasticity*, English translation, P. Noordhoff Ltd., Groningen, Holland, 1953.

Rice, J. R., 1987, "Elastic Fracture Mechanics Concepts for Interfacial Cracks," Harvard University Technical Report MECH-101, May, to appear in ASME JOURNAL OF APPLIED MECHANICS.

Rice, J. R., and Sih, G. C., 1965, "Plane Problems of Cracks in Dissimilar Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 418-423.

Sih, G. C., and Rice, J. R., 1964, "The Bending of Plates of Dissimilar Materials with Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 477-482.

Symington, M. F., 1987, "A Review of Some Aspects of the Linear Elastic Analysis of Bimaterial Cracks," Brown University Master's Thesis, May.

Williams, M. L., 1957, "On the Stress Distribution at the Base of a Stationary Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 24, pp. 109-114.

Williams, M. L., 1959, "The Stresses Around a Fault or Crack in Dissimilar Media," *Bulletin of the Seismological Society of America*, Vol. 49, pp. 199-204.

Complex Singularities of the Transfer Function for Cylindrical Cavities in Elastic Media

A. N. Stokes[6] and A. F. Siggins[7]

1 Introduction

The response of a cylindrical cavity in an elastic medium to incoming disturbance has been much examined in recent years. A knowledge of the cavity transfer function is of importance in such areas as the seismic response of underground excavations, the response of geophones placed within boreholes, and various non-destructive testing applications. From some early work of Viktorov (1958), Fourier-Bessel series methods have been developed (Pao and Mow, 1973; Baron and Matthews, 1961), and "wave-sum" integral methods (Miklowitz, 1978) have been used to describe the transfer function.

The "wave-sum" method is useful at high frequencies giving a solution which can be conveniently divided into a Rayleigh wave component and various P and S waves. There are some difficulties in applying the method in the "illuminated zone." In this method an expression corresponding to a Fourier-Bessel series coefficient is evaluated by summing residues at singularities in a complex plane of spatial frequencies. Miklowitz (1978) discusses asymptotic properties which are associated with each type of wave.

The Fourier-Bessel series converge best for disturbances whose wavelength is rather large compared with the cavity diameter. In this case, for subsequent convolution of the transfer function with a specific disturbance, it is desirable to find singularities of the coefficients in the complex plane of temporal frequency. This has been done and some values for the first few coefficients are given by Pao and Mow (1973) and Baron and Matthews (1961). In the methods used there the numerical difficulty increases with increasing order of coefficients.

The singularities are studied by Pao and Mow (1973) because they are important in the long-term behavior of the solution. In this report a different method of computing them is described which does not seem to have that numerical problem, so that calculation of the poles can be extended into a region where asymptotic high frequency approximations apply.

2 Various Transfer Functions

A Fourier-Bessel series for the cavity hoop-stress response to an incoming steady state compressional plane wave is derived by Pao and Mow (1973) and is also described in Peck (1965). With a change of notation, its pth coefficient has a denominator, D_p, given by

$$D_p = xH_{p-1}(x)[(p^2-1)yH_{p-1}(y) - (p^3 - p + y^2/2)H_p(y)]$$
$$- H_p(x)[(p^3 - p + y^2/2)yH_{p-1}(y) - (p^2 + p - y^2/4)y^2 H_p(y)] \qquad (1)$$

[6]CSIRO Division of Mathematics and Statistics, Melbourne, Australia.
[7]CSIRO Division of Geomechanics, Mount Waverley, Australia.

Manuscript received by ASME Applied Mechanics Division, March 4, 1985; final revision June 22, 1986.

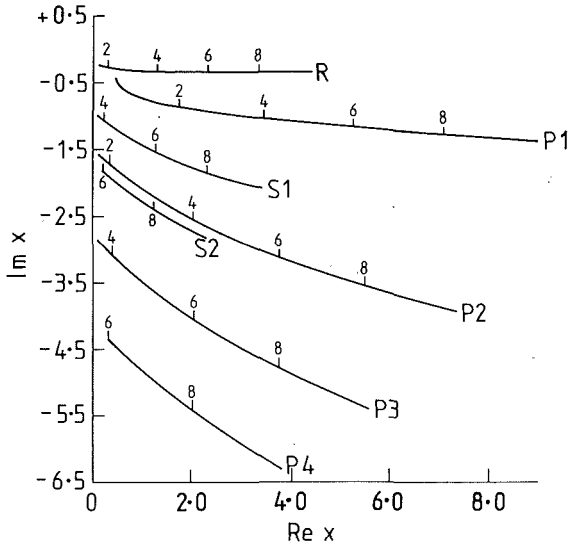


Fig. 1 Trajectories of the poles in the complex x plane

where

x is the wavenumber $\omega a(\rho/(\lambda+2\mu))^{1/2}$;

y is the wavenumber $\omega a(\rho/\mu)^{1/2}$;

λ and μ are Lamé elastic constants;

ρ is density;

ω is circular frequency;

a is cavity radius.

The omission by Pao and Mow (1973) of y^2 in the last term in equation (1) appears to be an error. The Hankel function, $H_p(x)$ is usually taken to be $H_p^{(1)}(x)$.

The steady-state solution for general frequency may be used to derive solutions for transient problems. For example Miklowitz (1978) derived the cavity response to a line-load source within the medium. The line load is impulsive in time but its Laplace or Fourier transform corresponds to a steady-state solution and an appropriate denominator, C_p, is found where

$$C_p=\frac{H_{p+2}(x)H_{p-2}(y)+H_{p+2}(y)H_{p-2}(x)+(y^2/x^2-1)H_p(x)[H_{p+2}(y)+H_{p-2}(y)]}{H_p(x)H_p(y)} \tag{2}$$

Viktorov (1958), studying Rayleigh waves on the surface of a concave cylinder, derived a frequency equation,

$$V_p=\frac{H_{p+2}(x)+H_{p-2}(x)-2(y^2/x^2-1)H_p(x)}{H_{p+2}(x)-H_{p-2}(x)}$$

$$-\frac{H_{p+2}(y)-H_{p-2}(y)}{H_{p+2}(y)+H_{p-2}(y)}=0 \tag{3}$$

In fact the expressions D_p, C_p, and V_p are closely related and have the same zeros. By using the recurrence relations

$$H_{p+1}(x)+H_{p-1}(x)=\frac{2p}{x}H_p(x) \tag{4}$$

and

$$xH_p'(x)=pH_p(x)-xH_{p+1}(x)=xH_{p-1}(x)-pH_p(x) \tag{5}$$

it is possible to reduce D_p to the form

$$D_p=(p^2-1)xyH_p'(x)H_p'(y)-\frac{y^2}{2}(xH_p'(x)H_p(y)$$

$$+yH_p'(y)H_p(x))+\left[p^2-\left(p^2-\frac{y^2}{2}\right)^2\right]H_p(x)H_p(y) \tag{6}$$

and to show that

Table 1 Comparison of calculated eigenvalues for circular cavities in elastic media. Poisson's ratio = 0.25.

p	Our value	Pao and Mow 1973	Baron and Parnes 1961
0	0.4474 - .4420 i	0.44647 - 0.44127 i	0.4464 - 0.4410 i
1	1.09272 - 0.7653 i	1.09272 - 0.7653 i	1.0929 - 0.441 i
2	1.90754 - 0.8978 i	1.90754 - 0.8978 i	1.9076 - 0.897 i
3	2.75652 - 0.9915 i	2.75652 - 0.9915 i	
4	3.63132 - 1.0666 i	3.63132 - 1.0666 i	
5	4.52440 - 1.1314 i	4.52440 - 1.1314 i	

$$C_p=-8D_p/(x^2y^2H_p(x)H_p(y))$$

$$V_p=\frac{-8D_p}{x^2y^2(H_{p+2}(x)-H_{p-2}(x))(H_{p+2}(y)+H_{p-2}(y))} \tag{7}$$

Since Hankel functions have no singularities, except at $x = y = 0$, the zeros of V_p and C_p are those of D_p.

Dividing D_p by $H_p(x)\,H_p(y)$ yields:

$$(p^2-1)F(x)F(y)-y^2/2(F(x)+F(y))$$

$$+p^2-(p^2-y^2/2)^2=0 \tag{8}$$

where $F(x) = xH_p'(x)/H_p(x)$.

The Hankel function ratios, $F(x)$, can be conveniently computed by the use of continued fraction expansions. The expansion is derived from the recurrence relation for the confluent hypergeometric function (Abramowitz and Stegun, 1965; 13.4.15); the theory is set out by Temme (1975), but in the present case is rather simpler, because a normalizing sum is not required. Then:

$$F(x)=ix-1/2+\cfrac{p-1/4}{2-2ix-\cfrac{(p-9/4)}{4-2ix-\cfrac{(p-25/4)}{6-2ix\ \ldots}}} \tag{9}$$

where $i^2=-1$.

This is an expansion suitable for large arguments analogous to the asymptotic series used in such cases but convergent. Methods for evaluating such fractions are described in Abramowitz and Stegun (1965).

3 Calculation of Poles and Their Interpretation

Being an expansion for large argument, expression (8) seems to avoid the numerical difficulties found by Pao and Mow (1973) as p increased. Convergence difficulties arose only when x came very close to the negative imaginary axis.

Expression (8) is an analytic function in both p and x (and hence y). Consequently the equation implicitly determines either p as a function of x or vice versa, and that function is generally multi-valued and analytic except at branch points. For various purposes it is convenient to hold either p or x real and consider poles in the complex domain of the other variable. Miklowitz (1978), following Peck (1965), constrains x to be real, while Pao and Mow (1973) and Baron and Mathews (1961) consider x complex with p integer.

Table 1 gives a comparison of some poles calculated by the method described here with other values in the literature for Poisson's ratio, $\upsilon = 0.25$ ($\upsilon = \lambda/2(\lambda+\mu)$). Agreement with Pao and Mow (1973) in particular is complete except for a puzzling small discrepancy at $p = 0$. Our value was checked carefully and seems correct, notwithstanding that two other published values are closer to each other than to ours.

Continuous trajectories of the poles in the complex $x-$*plane for real positive* p up to $p=10$ were calculated and

Table 2 Table of zeros of $D_p(x)$ for small integer p, and complex x. The derivatives of D with respect to x at the zeros are also given; these are needed in calculating residues.

p	x value at zero of D		Derivative	
	Real	Imag	Real	Imag
		P1 branch		
0.000	0.4474	-0.4420	-0.4999	-0.2448
1.000	1.0927	-0.7654	-4.9668	7.4510
2.000	1.9075	-0.8978	-30.0801	36.8774
3.000	2.7565	-0.9915	-99.0362	105.8696
4.000	3.6313	-1.0666	-241.8189	235.9453
5.000	4.5244	-1.1314	-492.8187	451.8991
6.000	5.4306	-1.1899	-889.7499	781.5193
7.000	6.3467	-1.2440	-1473.0596	1255.3524
8.000	7.2702	-1.2950	-2285.5579	1906.5410
9.000	8.1998	-1.3435	-3372.2041	2770.6892
10.000	9.1342	-1.3900	-4779.8574	3885.7664
		R branch		
2.000	0.2862	-0.2786	-4.9601	1.4331
3.000	0.7946	-0.3245	-26.0566	-6.9026
4.000	1.2979	-0.3437	-71.0670	-30.3484
5.000	1.8027	-0.3525	-151.0883	-74.4564
6.000	2.3098	-0.3559	-278.2668	-144.4714
7.000	2.8192	-0.3560	-465.3952	-245.1723
8.000	3.3307	-0.3542	-725.7965	-380.8448
9.000	3.8440	-0.3510	-1073.1565	-555.3453
10.000	4.3588	-0.3469	-1521.5138	-772.0622
		P2 branch		
2.000	0.4041	-1.7852	64.5148	-77.1535
3.000	1.2325	-2.2281	381.3022	-23.6993
4.000	2.0782	-2.5798	874.4141	395.4999
5.000	2.9359	-2.8760	1414.0438	1339.6630
6.000	3.8033	-3.1336	1860.6656	2926.2717
7.000	4.6787	-3.3625	2074.7217	5252.8540
8.000	5.5612	-3.5692	1917.3248	8407.5664
9.000	6.4501	-3.7580	1249.5570	12474.1729
10.000	7.3446	-3.9322	-68.2529	17534.6934
		S1 branch		
4.000	0.2580	-1.1025	-85.9125	193.4471
5.000	0.7711	-1.3573	-428.2726	306.4424
6.000	1.2858	-1.5530	-1015.4634	306.9080
7.000	1.8036	-1.7144	-1868.3984	129.3510
8.000	2.3246	-1.8528	-3012.7083	-288.1648
9.000	2.8486	-1.9745	-4474.5874	-1009.1045
10.000	3.3752	-2.0835	-6279.9517	-2099.0527
p	x value at zero of D		Derivative	
		P3 branch		
4.000	0.4155	-3.1005	496.9440	-1230.6393
5.000	1.2554	-3.6245	2376.0454	-1691.7261
6.000	2.1053	-4.0664	5293.3813	-1130.9962
7.000	2.9635	-4.4518	8977.6406	911.3301
8.000	3.8289	-4.7958	13116.4980	4775.0220
9.000	4.7005	-5.1076	17393.8184	10721.2168
10.000	5.5777	-5.3938	21501.0898	18962.2012
		S2 branch		
6.000	0.2542	-1.8734	-321.9348	1263.5731
7.000	0.7625	-2.1746	-1445.1500	1955.0776
8.000	1.2721	-2.4264	-3231.8972	2488.4888
9.000	1.7840	-2.6448	-5729.9917	2727.8069
10.000	2.2983	-2.8387	-8987.1289	2553.8904
		P4 branch		
6.000	0.4199	-4.4230	1669.4871	-6051.8496
7.000	1.2660	-4.9842	7091.9971	-8432.7812
8.000	2.1185	-5.4749	15154.9092	-8953.7119
9.000	2.9771	-5.9140	25458.0840	-6767.7256
10.000	3.8410	-6.3131	37521.6367	-1206.1699

are shown in Fig. 1 for Poisson's ratio = 0.25. As noted in Pao and Mow (1973), for positive integer values $p = n$, there are n poles if n is odd and $n+1$ if n is even. So most of the trajectories must cover a range of p values which is bounded below, and it is of interest to know how this bound is attained. In fact all but one of the trajectories approach the negative imaginary axis as p diminishes, cross at $p = 2n - 1/2$ for some positive integer n, then continue as a mirror image reflected about the imaginary axis, with p increasing again. The exception is the single trajectory which reaches $p = 0$; since the expression in equation (8) is an even function in p, the trajectory here meets its negative p counterpart at a branch point.

The p values on the trajectories are marked where p is an integer; each is plotted as far as $p = 10$. Asymptotically the behavior of the various branches is as described by Peck and Miklowitz (1969) with x regarded as the independent variable instead of p. Most of the poles arise when either $H_p(x)$ or $H_p(y)$ enter the transition region for Bessel functions where the argument approximately equals the order, and an expansion in Airy functions applies. So if $p \approx x$, then

$$x \approx p - a_j (p/2)^{1/3} \omega \quad j = 1,2 \ldots . \tag{10}$$

where $\omega = \exp(-2\pi i/3)$ and a_j is the jth zero of the Airy function Ai.

The phase velocity approaches that for P waves, and the trajectories are labelled $P1$, $P2$, ...

Similary there are trajectories for which

$$y \approx p - a_j (p/2)^{1/3} \omega \quad j = 1,2 \ldots . \tag{11}$$

These have a phase velocity approaching that for S waves and are labelled $S1$, $S2$, . . .

Finally there is one pole which is obtained from the Debye expansion for $H(x)$ where $|x|, |y| < |p|$. This is the Rayleigh wave pole, with

$$x \approx (c_R/c_D)p + \text{constant} \tag{12}$$

Here c_R and c_D are the phase velocities for Rayleigh and P waves (dilational waves), respectively. That trajectory is labelled R.

It is interesting, but perhaps not surprising, that it is the P_1 trajectory which penetrates $p = 0$. In the Fourier-Bessel series the $p = 0$ term determines the radially symmetric radiation component in the far-field. The two trajectories which turn about at $p = 1.5$ are P_2 and R, at $p = 3.5$ they are P_3 and S_1, and at $p = 5.5$ they are P_4 and S_2.

In Table 2 the zeros for integer p are given, together with the derivatives of D_p with respect to x. These are needed in order to calculate the residues, if the integral arising from the use of a Fourier-Bessel series as a transfer function is to be evaluated by residue calculus.

4 Conclusion

Essentially the same denominator function, with the same zeros, occurs in the small time "wave sum" method of Miklowitz (1978), and the long-time Fourier-Bessel method described by Pao and Mow (1973). The continued fraction (9) gives an improved method of calculation of the zeros. Regarding the denominator as a function of two complex variables, determining either as an analytic multi-valued function of the other, enables physical interpretation of the zeros, or branches, according to their asymptotic behavior.

5 References

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

Baron, M. L., and Mathews, A. T., 1961, "Diffraction of a Pressure Wave by a Cylindrical Cavity in an Elastic Medium," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, pp. 347-354.

Blair, D. P., 1984, "Rise Times of Attenuated Seismic Pulses Detected in Both Empty and Fluid-Filled Boreholes," *Geophysics*, Vol. 48, No. 4, pp. 398-410.

Miklowitz, J., 1978, *Theory of Elastic Waves and Waveguides*, North Holland, Amsterdam.

Pao, Y. H., and Mow, C. C., 1973, *Diffraction of Elastic Waves of Dynamic Stress Concentrations*, Crane, Russack and Company, New York.

Peck, J. C., 1965, "Plane-Strain Diffraction of a Transient Elastic Wave by a Circular Cavity," Ph.D. thesis, California Institute of Technology, Pasadena.

Peck, J. C., and Miklowitz, J., 1969, "Shadow Zone Response in the Diffraction of a Plane Compressional Pulse by a Circular Cavity," *Int. J. Solids and Structures*, Vol. 5, pp. 437-454.

Temme, N. M., 1975, "On the Numerical Evaluation of the Modified Bessel Function of the Third Kind," *J. Computational Phys.*, Vol. 19, pp. 324-337.

Viktorov, I. A., 1958, "Rayleigh-Type Waves on a Cylindrical Surface," *Soviet Physics-Acoustics*, Vol. 4, pp. 131-136.

Journal of Applied Mechanics

Book Reviews

Acoustic and Electromagnetic Waves, by D. S. Jones. Oxford University Press, New York, 1986. 745 pages. Price: $145.00.

REVIEWED BY J. D. ACHENBACH[1]

Acoustic and electromagnetic wave motions are governed by essentially the same equations, and the same mathematical techniques are used to solve wave propagation problems in the two fields. This book discusses the common features of methods to analyze the two kinds of wave motion. Where necessary it draws attention to differences, which may be due to different boundary conditions or due to polarization effects of electromagnetic waves.

This is a book for the advanced reader, who has had exposure to the basic material. The book is some 700 pages long, but it seems to contain the information of a volume twice that length. The material is presented in a very concise form, little detail is offered. Typically the author states a problem and a few steps in the solution process, then he gives the expression for the solution and concludes with a brief discussion. This reviewer would have liked to see a larger list of selected referneces for the details of the mathematical developments and for related work. The number of relevant publications in the areas covered by this book makes anything resembling complete referencing totally impossible, but the author has gone to the other extreme of including a very short list.

The first two Chapters are general in nature. Chapter 1 deals with the general properties of solutions of the equations governing acoustics and electromagnetism in matter which has certain macroscopic properties. Chapter 2 deals with relativistic effects. The remainder of the book is primarily concerned with solutions to specific problems. Chapter 3 deals with radiation, and Chapter 4 with resonators. The theory of waveguides is discussed in Chapter 5, while Chapter 6 discusses refraction. Surface waves are the topic of Chapter 7. The next two chapters deal with scattering by smooth objects (Chapter 8) and diffraction by edges (Chapter 9). The last Chapter is concerned with transient waves. Seven appendices give details of a variety of mathematical techniques and special functions. Each chapter concludes with a set of exercises.

The book presents a wealth of interesting and valuable material. It is an excellent contribution to the literature on wave phenomena. The sophisticated reader will find most he or she may ever want (or need) to know about analytical methods to solve linear acoustic and electromagnetic wave problems.

[1]Professor, Department of Civil Engineering, Northwestern University, Evanston, IL 60208.

Control of Structures, by H. H. E. Leipholz and M. Abdel-Rohman. Martinus Nijhoff Publishers, The Netherlands, 1986. 413 pages. Price: $105.50.

REVIEWED BY W. L. HALLAUER, JR.[2]

Many conference proceedings and doctoral dissertations have beeen devoted to structural control in recent years, but this is apparently the first published monograph on the subject. The authors intend it to serve as a reference for researchers, practicing engineers, and students in advanced courses on structural dynamics and control. The book consists primarily of short summaries of theory, extensive numerical simulations of control for simple beam structures, and discussion of the simulation results. Much of the material is drawn from the authors' previous publications. There is very little reference to experiments and in-service implementations, and no measured results are presented.

Although this book includes much general material on control theory, its title suggests a substantially broader subject area than is considered. A title such as "Active Control of the Dynamic Response of Civil Structures" would more accurately describe the actual contents. Most of the book is not directly relevant to applications outside of civil engineering, such as control of acoustically excited structural response and control of static and dynamic deformation of flexible aircraft and spacecraft structures.

Chapter 1—Introduction to Structural Control (35 pages)—defines and illustrates the major classes of structural control (active, passive, open-loop and feedback), discusses active feedback control of a one-degree-of-freedom oscillator, and conceptually describes some control devices. The appendix summarizes the principal concepts and definitions of classical control theory.

Chapter 2—Morphology of Structural Control (29 pages)—considers several issues with the use of partial differential operators, functional analysis, and examples involving strings and beams. The authors' main point is that conceptual design of structural control generally allows a wide variety of options for control operators and optimization objectives, so that the design for a particular situation is not simply a well defined mathematical process, but rather an art based on experience. Also, discretization by modal analysis is introduced, and the authors discuss the problems of high order and spillover instability associated with a large number of modes.

Chapter 3—Automatic Active Control of Simple Span Bridges (160 pages)—considers: classical control; modern control by pole assignment, optimal regulation, and optimal tracking; and state estimation by pole assignment. A bridge is idealized as a uniform simply-supported beam, and the distur-

[2]Professor of Aerospace Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.